DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

A generalized exponential time series regression model for electricity prices

N. Haldrup, O. Knapik and T. Proietti

CREATES Research Paper 2016-8

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

A generalized exponential time series regression model for electricity prices ☆

N. Haldrup[a,*], O. Knapik[a], T. Proietti[a,b]

[a] *Aarhus University, Department of Economics and Business Economics, CREATES - Center for Research in Econometric Analysis of Time Series, Fuglesangs Alle 4, 8210 Aarhus V, Denmark*
[b] *Dipartimento di Economia e Finanza, via Columbia 2, 00133 Rome, Italy*

Abstract

We consider the issue of modeling and forecasting daily electricity spot prices on the Nord Pool Elspot power market. We propose a method that can handle seasonal and non-seasonal persistence by modelling the price series as a generalized exponential process. As the presence of spikes can distort the estimation of the dynamic structure of the series we consider an iterative estimation strategy which, conditional on a set of parameter estimates, clears the spikes using a data cleaning algorithm, and reestimates the parameters using the cleaned data so as to robustify the estimates. Conditional on the estimated model, the best linear predictor is constructed. Our modeling approach provides good fit within sample and outperforms competing benchmark predictors in terms of forecasting accuracy. We also find that building separate models for each hour of the day and averaging the forecasts is a better strategy than forecasting the daily average directly.

Keywords: Robust estimation, long-memory, seasonality, electricity spot prices, Nord Pool power market, forecasting, robust Kalman filter, generalized exponential model
JEL: C1, C5, C53, Q4

1. Introduction

The daily spot prices from the Nord Pool market exhibit persistent (long memory) features combined with a periodic behaviour, related to weekly, monthly and yearly seasonality, which is also rather persistent. The fractional noise model (see Granger, Joyeux, 1980, and Hosking, 1981) is suitable for capturing the persistence of the series at the long run frequency. Periodic patterns reverting slowly to deterministic cycles can be for instance modelled by the class of generalized fractional, or Gegenbauer, processes, introduced by Hosking (1981) in his seminal paper, and analyzed by Gray, Zhang and Woodward (1989). After applying a cascade of (generalized) fractional filters, a popular approach is to assume that the filtered series is a short memory autoregressive moving average (ARMA) process.

☆Acknowledgment: The authors appreciate support from Center for Research in Econometric Analysis of Time Series, CREATES funded by the Danish National Research Foundation with grant number DNRF78.
*Corresponding author
Email address: `nhaldrup@creates.au.dk` (N. Haldrup)

These processes have been used widely in many empirical applications. Ferrara and Guegan (2001) used Gegenbauer processes to analyze cointegration in the Nikkei spot index. Smallwood and Beaumont (2003) analyzed IBM trading volume. Woodward, Cheng, and Grey (1998) considered atmospheric CO_2 data. There is also evidence of fractional seasonal integration in macroeconomic time series (Arteche, Robinson (2000), Gil-Alana, Robinson (2001).

Gegenbauer processes have also been used in the energy market data analysis. Soares and Souza (2006) used Gegenbauer ARMA (GARMA) processes with explanatory variables to forecast electricity demand. Diongue et al. (2004) introduced the k-factor Gegenbauer integrated generalized autoregressive conditional heteroskedasticity (GIGARCH) for modelling energy prices and subsequently Diongue et al. (2009) considered this process for forecasting EEX electricity spot market prices in Germany.

An alternative semiparametric approach is based on the generalized exponential model for the spectrum, as in Hsu and Tsai (2009). According to this approach, the logarithm of the spectrum of the short memory filtered series is represented by a finite trigonometric polynomial. Bloomfield (1973) introduced the exponential (EXP) model and discussed its maximum likelihood estimation, relying on the distributional properties of the periodogram of a short memory process, based on Walker (1964) and Whittle (1953). The model was then generalised to processes featuring long range dependence by Robinson (1991) and Beran (1993), originating the fractional EXP model (FEXP). Janacek (1982) proposed a method of moments estimator of the long memory parameter based on the log-periodogram. Maximum likelihood estimation of the FEXP model has been dealt with recently by Narukawa and Matsuda (2011). Hurvich (2002) addressed the issue of predicting with it.

Electricity spot prices are very complex data. They exhibit strong seasonality at the annual, weekly and daily levels and very high volatility and abrupt, short-lived and generally unanticipated extreme price changes, known as spikes (or jumps); (see Janczura et al. 2013, Weron, 2006, Serati et al., 2008, De Jong, 2006, among others). Characteristics of spot electricity prices include also mean-reversion and (seasonal) long memory; see Haldrup and Nielsen (2006a, 2006b), Weron (2008), Diongue, Guegan and Vingal (2004, 2009), Koopman et al. (2007).

The unique characteristics of electricity prices and related derivatives contracts have boosted the demand on econometric models that can precisely capture their dynamics (see Benth and Koakebakker, 2008, Möst and Keles, 2010, Raviv, Bouwman and van Dijk, 2015). Electricity price models are of huge importance in areas such as forecasting, derivative pricing and risk management. The recent reviews on forecasting electricity prices in Zareipour (2012) and Weron (2014) document the huge interest in that topic.

Most of the models are built for daily average prices, which play a key role in the electricity market. The average daily price is widely used to approximate other spot electricity prices and is used as a reference price for derivatives contracts e.g. futures and forwards. In electricity markets such as the Nord Pool Spot system, the average price is determined on a day-ahead market. Therefore, forecasts of these prices are highly relevant for market participants, as they can help in more efficient trading and consequently increase profits and control risk.

The paper contributes to the above literature by proposing a time series regression which relates prices to a set of explanatory variables with errors that follow a generalized exponential model, which accounts for the most prominent features, such as long range dependence in the mean and the weekly cycle. Differently from related research on Nord Pool daily electricity spot price modelling (see Weron, Simonsen and Wilman, 2004), we build our model directly for the levels, rather than for the first differences. We also conduct a rolling forecasting exercise in order to compare the daily average electricity spot prices forecasts arising from our model with several benchmarks.

Estimation is carried out in the frequency domain by maximising an asymptotic approximation to the likelihood function, known as the Whittle likelihood. The paper also introduces a new robust estimation procedure for dealing with price spikes, based on a robustification of the Kalman filter, which takes into account possible long-memory as well as the presence of outliers in the data.

The article is organized as follows. Section 2 describes the data and provides motivation for our modeling approach. Section 3 introduces the seasonal fractional exponential model and presents the concept of generalized long-memory and relevant theory on Gegenbauer processes. Section 4 discusses robust inference and forecasting with this model. Section 5 presents empirical results for modelling and forecasting daily average electricity spot prices from the Nord Pool power market. Section 6 concludes.

2. Data and time series descriptives of electricity spot prices from Nord Pool

Our data set refers to the Nordic power exchange, Nord Pool Spot, owned by the Nordic and Baltic transmission system operators, one of the leading power markets in Europe. About 380 companies from 20 countries trade on the Nord Pool Spot's markets, with participants including both producers and large consumers, for a trading volume of approximately 500 terawatt hours in 2014. The market includes Norway, Sweden, Finland, Denmark (since 2000), Estonia (since 2010), Lithuania (since 2012) and Latvia (since 2013). A detailed review of the operation of the market is given in NordPool (2015).

Within the Nord Pool Spot, Elspot is the auction market for day-ahead electricity delivery. The Nord Pool Spot web-based trading system enables participants to submit bids and offers for each individual hour of the next day. Orders can be made between 08:00 and 12:00 a.m. Central European Time (CET). The aggregated buy and sell orders form demand and supply curves for each delivery hour of the next day. The intersection of the curves constitutes the system price for each hour (quoted for megawatt hour, MWh) and is the equilibrium price that would exist in the absence of transmission constraints within the grid. We will use the words *system* and *spot* prices interchangeably to indicate this price. The hourly prices are announced to the market at 12:42 CET and contracts are invoiced between buyers and sellers between 13:00 and 15:00. All 24 prices on day $t + 1$ are determined on a given day t and released simultaneously (Weron, 2004, Raviv et al. 2015). The system price is not necessarily the price that is paid in the single areas of the power grid since the areas can be subject to transmission congestion. Hence, different area spot prices are likely to be determined as well as part of the day ahead auction. The reason why the (daily average) system price is important is that it serves as the reference price and clearing of most financial contracts and hence it is crucial for derivatives pricing and risk management.

The data set consists of the twenty-four hourly spot electricity system prices from the Nord Pool Elspot power exchange for each day from January 1, 2000 - January 30, 2014 covering 5144 observation days. In 2006 the trading currency in Nord Pool has changed from local (Nordic) currencies into the euro. Therefore, all prices are exchanged into the same currency.

Figure 1 plots the daily electricity log system price, $y_t = \ln Y_t$, where Y_t is the average of 24 hourly prices, which is the main object of interest in our empirical modelling and forecasting exercises. While the daily average (based on hourly prices) represents our target variable, we will also consider building separate models for the hourly prices, with the purpose of forecasting the daily average by aggregating the forecasts of the individual hours. We adopt the logarithmic transformation in order to stabilize the variation of the series. Looking at the data, we observe potentially nonstationary and persistent behavior which might be caused by long-memory features of the electricity spot prices. System prices vary over the week and over the year and are characterized by persistent level changes and spikes.

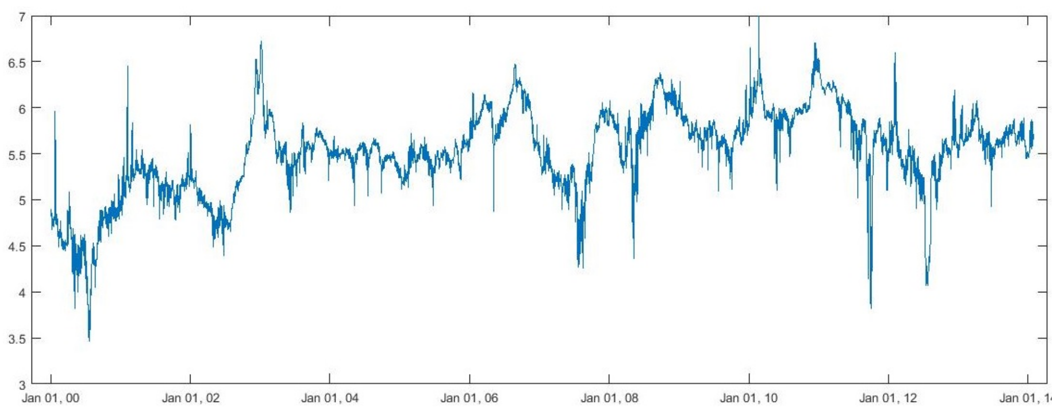

Figure 1: Daily log average spot system price for Nord Pool power market (Norwegian kroners (NOK) per MWh), y_t. Sample: January 1, 2000 - January 30, 2014

Table 1 contains descriptive statistics for the hourly and daily average log spot prices y_t. Log electricity spot prices are on average higher during the day than during the night. The highest values are in the morning for the hours from 08:00 to 12:00 CET and early evening, 17:00-19:00 CET. Those peak hours also have relative high skewness and kurtosis reflecting that these hours experience more extreme prices and price spikes. This also explains the relative high volatility during peak hours caused by a very elastic power supply curve and high marginal costs during peak hours. For each hour we have negative skewness and positive kurtosis, suggesting an empirical distribution which is skewed to the left with heavy tails and outliers caused by price spikes and drops.

Table 1: Descriptive statistics for log spot prices y_t in Nord Pool

	Minimum	Maximum	1. Quartile	3. Quartile	Mean	Median	Stdev	Skewness	Kurtosis
1	3.08	6.72	5.21	5.77	5.45	5.51	0.49	-0.85	1.43
2	2.94	6.70	5.17	5.73	5.41	5.48	0.52	-0.96	1.73
3	2.14	6.71	5.13	5.71	5.37	5.45	0.54	-1.03	1.94
4	1.47	6.69	5.10	5.70	5.35	5.44	0.56	-1.10	2.23
5	-1.71	6.70	5.11	5.71	5.35	5.45	0.57	-1.40	6.23
6	1.14	6.74	5.17	5.74	5.40	5.48	0.55	-1.18	2.89
7	1.14	6.80	5.23	5.80	5.47	5.54	0.52	-1.19	3.23
8	2.51	7.39	5.30	5.87	5.55	5.60	0.50	-0.84	1.83
9	3.20	7.81	5.35	5.91	5.60	5.63	0.48	-0.55	1.45
10	3.51	7.78	5.37	5.91	5.60	5.64	0.46	-0.54	1.09
11	3.59	7.44	5.38	5.90	5.61	5.64	0.45	-0.57	0.88
12	3.60	7.22	5.36	5.89	5.60	5.63	0.44	-0.59	0.82
13	3.59	7.11	5.35	5.88	5.58	5.61	0.45	-0.62	0.81
14	3.56	7.06	5.33	5.87	5.57	5.60	0.45	-0.65	0.83
15	3.58	7.04	5.31	5.86	5.55	5.59	0.45	-0.66	0.86
16	3.56	6.87	5.30	5.85	5.55	5.59	0.46	-0.67	0.90
17	3.56	7.12	5.31	5.86	5.55	5.59	0.47	-0.61	0.94
18	3.58	7.45	5.33	5.89	5.58	5.61	0.48	-0.50	1.05
19	3.62	7.38	5.33	5.89	5.58	5.61	0.47	-0.58	0.97
20	3.59	7.17	5.32	5.86	5.57	5.60	0.46	-0.68	1.02
21	3.58	6.94	5.31	5.83	5.55	5.59	0.45	-0.72	1.04
22	3.56	6.76	5.30	5.82	5.54	5.58	0.45	-0.72	1.05
23	3.58	6.73	5.28	5.82	5.52	5.57	0.45	-0.69	0.90
24	3.17	6.73	5.23	5.78	5.47	5.52	0.47	-0.78	1.21
Average price	3.46	6.99	5.29	5.83	5.52	5.57	0.46	-0.69	0.96

Figure 2 presents the sample autocorrelation function (ACF) and log periodogram of the logarithm of the daily average spot electricity price and its first differences. The log periodogram of electricity log prices has characteristics often found for time series processes with long memory. There are evident peaks around the harmonic frequencies: 0, $2\pi/7$, $4\pi/7$, $6\pi/7$ corresponding to the long run and weekly cycles present in the data. The persistence of the autocorrelation function at the seasonal lags 7, 14 and so on is also pronounced as can be seen from Figure 2. The seasonal persistence is more pronounced when we look at the first differences. In order to model the persistence it is reasonable to choose a model which allows for fractional integration at the zero as well as the seasonal frequencies.

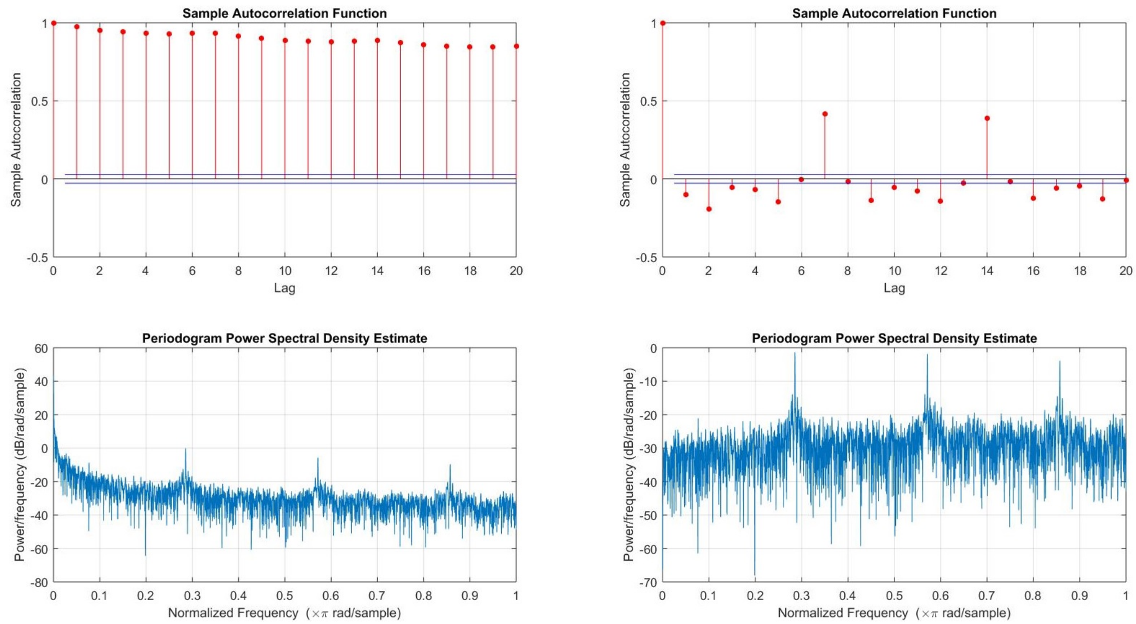

Figure 2: Autocorrelation function and log periodogram for daily average log-prices y_t and their first differences Δy_t. Sample: January 1, 2000 - January 30, 2014

3. The seasonal fractional exponential model

The crucial step in constructing a model for electricity price dynamics consists of finding an appropriate description of the seasonal pattern. There are different suggestions for that task in the literature. In order to account for seasonality we build our model on the Gegenbauer processes as in Hsu and Tsai (2009). We extend their approach by proposing a time series regression model with generalized FEXP disturbances, which is particularly well suited to account for most features characterizing electricity prices and allows to incorporate additional explanatory variables as covariates in the analysis. Our model is similar to the time series regression model based on Gegenbauer ARMA (GARMA) processes proposed by Soares and Souza (2006) for electricity demand prediction, but we take a flexible semiparametric approach.

3.1. Model specification

Let y_t denote a daily time series referring to the logarithmic prices of a particular hour of the day, or the daily average. We consider the following times series regression model:

$$y_t = x_t'\beta + u_t, \tag{1}$$

where x_t is a $k \times 1$ vector of explanatory variables and u_t is a zero mean random process. Denoting by B the lag operator, u_t is generated by the following fractionally integrated process

$$(1-B)^{d_0} \prod_{j=1}^{3} (1 - 2\cos\varpi_j B + B^2)^{d_j} u_t = z_t, \quad \varpi_j = \frac{2\pi}{7} j, j = 0, 1, 2, 3, \tag{2}$$

where $d_j, j = 0, 1, 2, 3$, are the fractional integration parameters at the Gegenbauer frequencies and z_t is a short memory process characterized by the spectral density $f_z(\omega)$, that will be defined below. The factor $(1-B)^{d_0}$ accounts for the long range dependence at the long run frequency, whereas the Gegenbauer polynomials $(1 - 2\cos\varpi_j B + B^2), j = 1, 2, 3$, account for the persistent seasonal behavior of the process at the frequencies $\varpi_j = \frac{2\pi}{7}j$, which correspond to cycles with period 7 days ($j = 1$), 3.5 days ($j = 2$, two cycles in a week) and 2.3 days ($j = 3$, three cycles per week). The process is stationary if $d_0 \in (0, 0.5)$ and $d_j \in (0, 0.25)$ (see Woodward, Gray, Elliott (2011), pg. 418, Theorem 11.5a). The Gegenbauer process was introduced by Hosking (1981) and formalized by Gray, Zhang and Woodward (1989).

The spectrum of the short memory component, z_t, follows Bloomfield's exponential model (see Bloomfield, 1973) of order q:

$$f_z(\omega) = \frac{1}{2\pi}\exp\left(c_{z0} + 2\sum_{k=1}^{q} c_{zk}\cos(\omega k)\right), \tag{3}$$

where $c_{z0}, c_{zk}, k = 1, \dots, q$ are real-valued parameters, known as the cepstral coefficients of z_t (Bogert, Healy and Tukey, 1963).

An important implication of the model specification is that the logarithm of the spectral generating function of u_t, denoted $2\pi f(\omega)$, is linear in the memory coefficients and in the (short-run) coefficients c_{zk}:

$$\ln[2\pi f(\omega)] = c_0 + 2\sum_{k=1}^{q} c_{zk}\cos(\omega k) - 2d_0\ln|2\sin(\omega/2)| - 2\sum_{j=1}^{3} d_j \ln\left|4\sin\left(\frac{\omega+\varpi_j}{2}\right)\sin\left(\frac{\omega-\varpi_j}{2}\right)\right|. \tag{4}$$

The inverse Fourier transform of the logarithmic spectrum in (4) provides the cepstral coefficients of the process u_t:

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln[2\pi f(\omega)]\exp(\imath\omega k)\mathrm{d}\omega, k = 0, 1, 2, \dots. \tag{5}$$

By trigonometric identities, see Gradshteyn and Ryzhik (2007) (formulae 1.441.2 and 1.448.2), for $k \geq 1$,

$$c_k = I(k \leq q)c_{zk} + \frac{1}{k}\left(d_0 + 2\sum_{j=1}^{3} d_j\cos(\varpi_j k)\right), \quad k = 0, 1, 2, \dots. \tag{6}$$

The sequence $\{c_k\}_{k=0,1,\dots}$ is referred to as the *cepstrum* (Bogert, Healy and Tukey, 1963) and carries all the relevant information that is needed for the linear prediction of the process u_t, and y_t thereof.

In the sequel we will refer to the specification consisting of (2) and (4), as the model for the logarithmic spectral density of u_t, as the Gegenbauer Exponential (GEXP) model.

3.2. Wold representation and linear prediction

Let $\mathcal{F}_t$ denote the information available up to time t, consisting of the past values of y_t and the current and past values of x_t. If the model is correctly specified, the one-step ahead prediction error variance (p.e.v.) of y_t, $\sigma^2 = Var(y_t|\mathcal{F}_{t-1})$, is obtained by the Szegö-Kolmogorov formula as

$$\sigma^2 = \exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln[2\pi f(\omega)]\mathrm{d}\omega\right],$$

from which it follows that $c_{z0} = \ln \sigma^2$. Moreover, if the stationarity condition is satisfied, we can obtain the Wold and autoregressive representations of the system,

$$y_t = x_t'\beta + \psi(B)\xi_t, \quad \phi(B)(y_t - x_t'\beta) = \xi_t, \quad \xi_t \sim WN(0, \sigma^2) \tag{7}$$

where $\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots, \sum_j \psi_j^2 < \infty$, and $\phi(B) = \sum_{j=0}^{\infty} \phi_j B^j = \psi(B)^{-1}$, $\sum_j \phi_j^2 < \infty$. The moving average coefficients of the Wold representation are obtained recursively from the cepstral coefficients by the formula

$$\psi_j = j^{-1} \sum_{r=1}^{j} r c_r \psi_{j-r}, \qquad j = 1, 2, \ldots, \tag{8}$$

with $\psi_0 = 1$. See Janacek (1982), Pourahmadi (1983) and Hurvich (2002) for a derivation of (8). The autoregressive coefficients are obtained according to the recursive formula

$$\phi_j = -j^{-1} \sum_{r=1}^{j} r c_r \phi_{j-r}, \qquad j = 1, 2, \ldots, \tag{9}$$

with starting value $\phi_0 = 1$.

4. Robust Estimation

4.1. Approximate (Whittle) Likelihood Estimation

Given a time series realization of length n, $\{(y_t, x_t), t = 1, 2, \ldots, n\}$, and letting $\omega_j = \frac{2\pi j}{n}$ denote the Fourier frequencies, for $j = 1, \ldots, \lfloor \frac{n-1}{2} \rfloor$, where $\lfloor \cdot \rfloor$ is the largest integer not greater than the argument, estimation of the regression parameters, β, the memory parameters, $d_j, j = 0, 1, 2, 3$, and the cepstral coefficients $c_{zk}, k = 0, 1, \ldots, q$, is carried out in the frequency domain.

Denoting the periodogram of y_t by $I_y(\omega)$, that of x_t by $I_x(\omega)$ and the cross-periodogram of x_t and y_t by $I_{xy}(\omega)$, where

$$I_y(\omega) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} y_t e^{-\imath\omega t} \right|^2, \quad I_x(\omega) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} x_t e^{-\imath\omega t} \right|^2, \quad I_{xy}(\omega) = \frac{1}{2\pi n} \left| \sum_{t=1}^{n} x_t y_t e^{-\imath\omega t} \right|^2,$$

the periodogram of $u_t = y_t - x_t'\beta$ can be written as $I(\omega) = I_y(\omega) - 2\beta' I_{xy}(\omega) + \beta' I_x(\omega)\beta$. If it is assumed that u_t is a stationary long memory process characterised by the spectral density $f(\omega)$, as implied by (4), the Whittle approximation to the likelihood is

$$\mathcal{L} = -\sum_{j=1}^{T-1} \left[\ln f(\omega_j) + \frac{I(\omega_j)}{f(\omega_j)} \right]. \tag{10}$$

The maximiser of (10) is the Whittle pseudo maximum likelihood estimator of $(\beta, d_0, d_1, d_2, d_3, c_{z0}, c_{z1}, \ldots, c_{zq})$. The parameter β can be concentrated out of the likelihood function, yielding the frequency domain generalised least squares estimate

$$\hat{\beta} = \left[\sum_{j=1}^{T-1} \frac{1}{f(\omega_j)} I_x(\omega_j) \right]^{-1} \sum_{j=1}^{T-1} \frac{1}{f(\omega_j)} I_{xy}(\omega_j).$$

After replacing into (10), the profile likelihood can be maximized with respect to the memory and cepstral parameters.

We refer to Dahlhaus (1989), Velasco and Robinson (2000), Giraitis, Koul and Surgailis (2012) and Beran et al. (2013) for the properties of the estimator in the long memory case.

4.2. Tapering

For estimation we use the tapered periodogram. Tapering aims at reducing the estimation bias that characterises the periodogram ordinates in the nonstationary case. Velasco (1999) and Velasco and Robinson (2000) show that that with adequate data tapers, the Whittle estimator of the parameters of classes of fractional integrated models, encompassing the FEXP and is consistent and asymptotically normal, when the true memory parameter is in the nonstationary region. The adoption of a data taper makes the estimates invariant to the presence of certain deterministic trends.

The tapered discrete Fourier transform of u_t is defined as the squared modulus of

$$w(\omega_j) = \left(2\pi \sum_{t=1}^{n} h_t^2\right)^{-1/2} \sum_{t=1}^{n} h_t u_t e^{i\omega_j t} \tag{11}$$

where $\{h_t\}_{t=1}^n$ is a taper sequence, i.e. a sequence of nonnegative weights that downweights the extreme values of the sequence on both sides, leaving the central part almost unchanged. Note that the raw periodogram arises in the case $h_t = 1$. As in Velasco and Robinson (2000), the $\{h_t\}_{t=1}^n$ sequence is obtained from the coefficients of the polynomial

$$\left(\frac{1 - z^{[n/p]}}{1 - z}\right)^p.$$

Typical choices are $p = 2, 3$. The tapered periodogram is then $I(\omega_j) = |w(\omega_j)|^2$.

4.3. Robust filtering and forecasting

As documented in section 2, electricity prices are characterized by abnormally high or low values, called price spikes. Their effect on the periodogram, and thus on the Whittle estimates depends on their size, pattern and recurrence. Assume that there is no periodicity (for the daily series, while their occurrence during the day may be more systematic), and recall that the frequency response function of a pulse dummy, $I_t(t = \tau)$, is constant across the frequency range. The results of the presence of a single outlier at time $t = \tau$ are the bias downwards of the memory parameters estimates and inflation of the estimate of the conditional and the unconditional variance of the series. There are several strategies for robustifying the estimates of the parameters and thus of the spectrum. Hill and McCloskey (2015) propose to replace the sample spectrum $I(\omega)$ in (10) by a robust periodogram, constructed from the Fourier transform of a robust autocovariance estimate. Our alternative strategy is based on an iterative data cleaning method extended by the robust Kalman filter introduced by Martin and Thomson (1982).

The proposed procedure entails iterating the following steps:

1. Estimate the parameters of the GEXP time series regression model of section 3.1 by approximate Whittle likelihood estimation, by maximising (10).
2. Obtain the AR or MA approximation of the GEXP model as described in Section 3.2.
3. Cast the approximating AR or MA model in state space form and apply the robust Kalman filter outlined in Appendix A. In order to eliminate the influence of outliers (price spikes), the filter shrinks y_t towards its one-step-ahead prediction, depending in the size of the innovation.
4. Replace the series y_t by its cleaned version (using the robust Kalman filter) and go to step 1.
5. Repeat steps 1-4 until no further outlier is found.

5. Empirical results

5.1. A time series model for Nord Pool electricity spot prices

In this section we present the empirical results for the daily Nord Pool data, whose characteristics were summarized in Section 2, with particular reference to the logarithm of the daily spot average price time series. A generalized exponential model based time series regression model like (2) may be adequate to capture the dynamics in the conditional mean of the series. The explanatory variables that we considered are the water reservoir level and dummy variables for holiday effects and week-of day effects. We set off by presenting and discussing the maximum likelihood estimates of the model parameters and by assessing its empirical adequacy. We also examine the influence of outliers (price spikes) on the parameter estimates, and with special attention to the memory parameter estimates.

Figure 3 presents the original and the transformed log electricity daily average spot prices with the use of tapering after removing the mean. As it can be seen from the plot, tapering brings the series closer to stationarity and removes potential trends from the data.

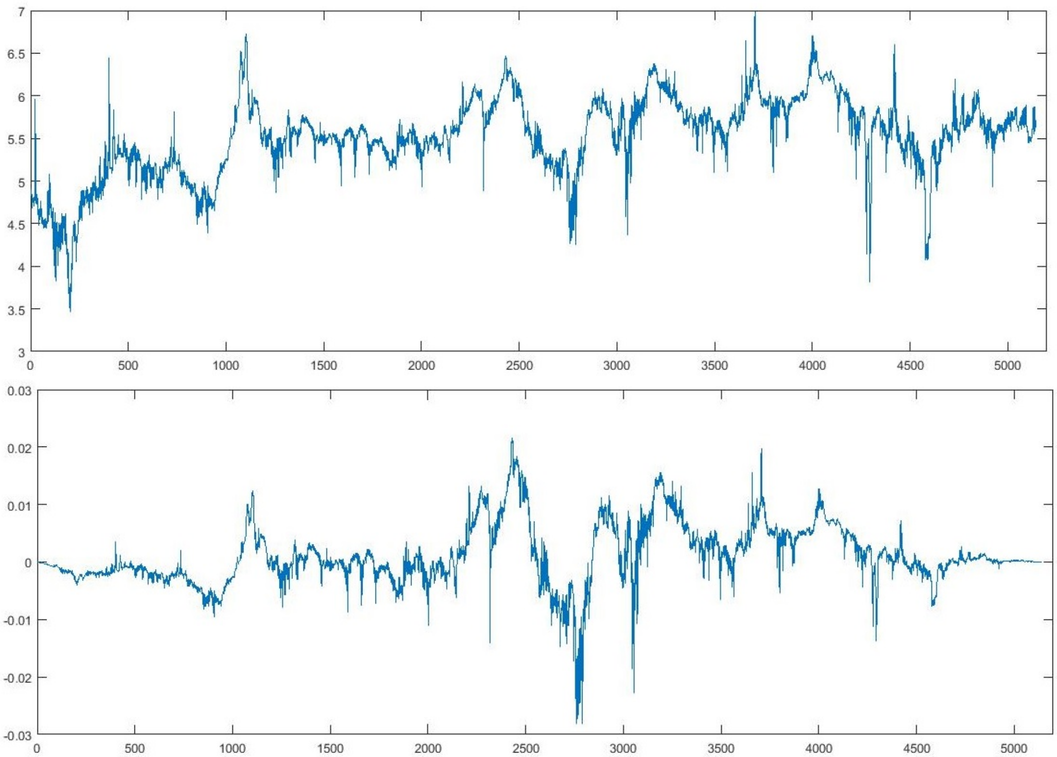

Figure 3: Original and tapered log daily average electricity spot prices

The results of fitting the GEXP model to the considered data on the log average electricity spot prices are presented in Tables 2-4. In each table we set together classical and robust estimates of the parameters in order to evaluate the influence of robustification for the parameter estimates. Table 2 presents the estimates of the memory parameters at the zero and seasonal frequencies: 0, $2\pi/7, 4\pi/7, 6\pi/7$. Electricity spot prices seem to be non-stationary, although still mean reverting, at the zero frequency, as the estimated d is significantly above 0.5. There is also a clear evidence of seasonal long-memory ($d_i > 0, i = 1, 2, 3$). The (seasonal) memory estimates are different depending on the estimation method: classic or robust. The robust estimates of memory parameters are usually higher than the classical ones. Normally the influence of outliers on memory parameters estimates will typically be a downward bias when outliers are present. It is seen, however, that robustification only affects the memory estimates to a minor extent.

Table 2: The estimates of memory parameters for the zero (d) and harmonic frequencies ($d_1 - d_3$)

	d	d_1	d_2	d_3
Classical	**0.7653**	**0.2121**	**0.1242**	0.0848
	(0.0784)	(0.0486)	(0.0500)	(0.0506)
Robust	***0.7984***	***0.1063***	***0.1024***	***0.1216***
	(0.0767)	(0.0472)	(0.0493)	(0.0505)

Note: Standard errors in brackets. The statistically significant parameters at the 5% significance level appear in bold.

Table 3 contains the estimated short-run cepstral coefficients describing the short run dynamics of the series. The number of short run cepstral coefficients is chosen according to the Akaike Information Criterion (AIC). For both estimation methods the optimal number according to AIC of the coefficients is 24. However, most of the short-run cepstral coefficients are statistically insignificant. The estimated intercept, $\hat{c}_0$, is the estimate of the logarithm of the prediction error variance. We might observe that the prediction error variance, $\hat{\sigma}^2 = \exp(\hat{c}_0)$, is lower for the robust estimation method.

Table 3: The estimates of the short-run cepstral coefficients

	c_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	c_9	c_{10}	c_{11}	c_{12}
Classical	***-13.7946***	-0.0140	0.0817	***0.1045***	0.0581	0.0182	0.0136	***0.0797***	-0.0050	0.0164	0.0096	-0.0101	-0.0017
	(0.0197)	(0.1373)	(0.0697)	(0.0468)	(0.0363)	(0.0306)	(0.0266)	(0.0304)	(0.0220)	(0.0206)	(0.0193)	(0.0185)	(0.0179)
Robust	**-14.5875**	0.2439	0.0636	0.0669	0.0532	0.0516	**0.1064**	0.0246	0.0304	-0.0007	0.0142	0.0023	-0.0003
	(0.0197)	(0.1373)	(0.0697)	(0.0439)	(0.0341)	(0.0306)	(0.0266)	(0.0304)	(0.0220)	(0.0205)	(0.0187)	(0.0179)	(0.0179)
	c_{13}	c_{14}	c_{15}	c_{16}	c_{17}	c_{18}	c_{19}	c_{20}	c_{21}	c_{22}	c_{23}	c_{24}	
Classical	-0.0210	***0.1191***	-0.0113	0.0117	-0.0112	0.0195	-0.0262	-0.0074	***0.1123***	-0.0119	-0.0307	0.0254	
	(0.0174)	(0.0194)	(0.0166)	(0.0163)	(0.0160)	(0.0158)	(0.0157)	(0.0155)	(0.0166)	(0.0152)	(0.0151)	(0.0150)	
Robust	**0.0492**	**0.0545**	0.0273	0.0016	0.0016	0.0162	-0.0242	**0.0354**	**0.0866**	-0.0203	**-0.0542**	-0.0286	
	(0.0174)	(0.0194)	(0.0166)	(0.0163)	(0.0157)	(0.0155)	(0.0156)	(0.0155)	(0.0166)	(0.0152)	(0.0151)	(0.0150)	

Note: Standard errors in brackets. The statistically significant parameter estimates at 5% significance level appear in bold.

Table 4 shows the estimation results concerning the explanatory variables. Most of the considered explanatory variables are statistically insignificant. Holidays and Sunday appear to be statistically significant when the classical estimator is used for estimation.

Table 4: The estimates of explanatory variables coefficients

	Holidays	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Water reservoir
Classical	***-0.0722***	0.0026	0.0039	-0.0045	-0.0200	-0.0265	***-0.0644***	-0.0051
	(0.0143)	(0.0307)	(0.0459)	(0.0521)	(0.0521)	(0.0478)	(0.0334)	(0.0054)
Robust	0.0096	0.0002	-0.0015	0.0003	0.0012	-0.0039	0.0067	-0.0003
	(0.0095)	(0.0132)	(0.0183)	(0.0201)	(0.0201)	(0.0203)	(0.0158)	(0.0052)

Note: The brackets contain standard errors. The statistically significant parameters at 5% significance level appear in bold.

Figure 4 displays the log-periodogram of the series and the estimated logarithmic spectrum. It is visible on the plot that the model effectively captures the spectral peaks at the seasonal frequency and at the zero frequency.

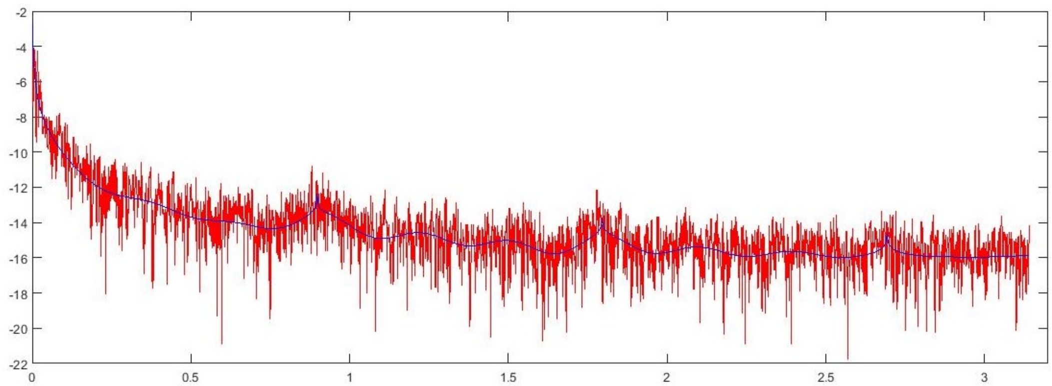

Figure 4: Goodness of fit in spectral domain for the cleaned log electricity prices

Figure 5 displays the 'cleaned' residuals from the model. The latter result from the iterative procedure outlined in section 4.3. After convergence, we obtain the coefficients of the AR representation and we truncate them at $m = 50$; the AR approximation is cast in state space form (see the appendix) and the pseudo-innovations are computed by the Kalman filter. As it can be noticed from their plot, the residuals display some volatility clustering, but are otherwise unaffected by the spike feature. In Figure 6 we plot the sample autocorrelation function of the cleaned residuals. We may conclude that the proposed modelling strategy involving tapering and robust estimation provides a reasonable fit of the series by the model.

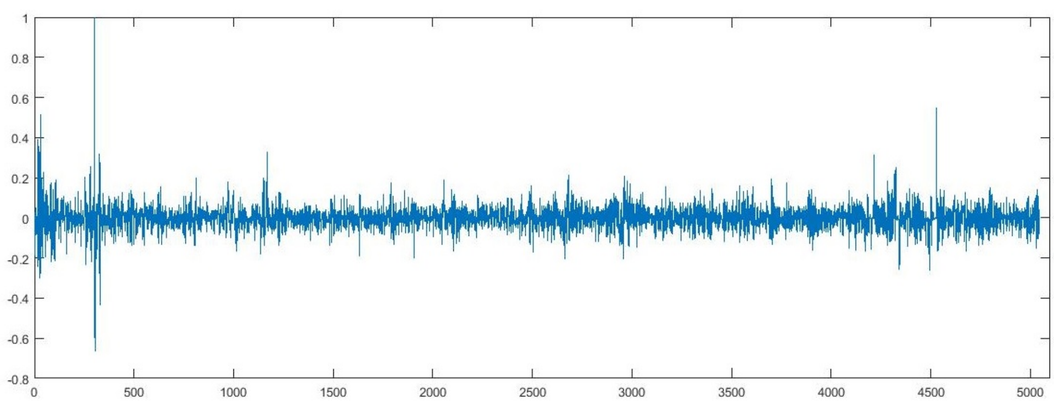

Figure 5: Cleaned residuals

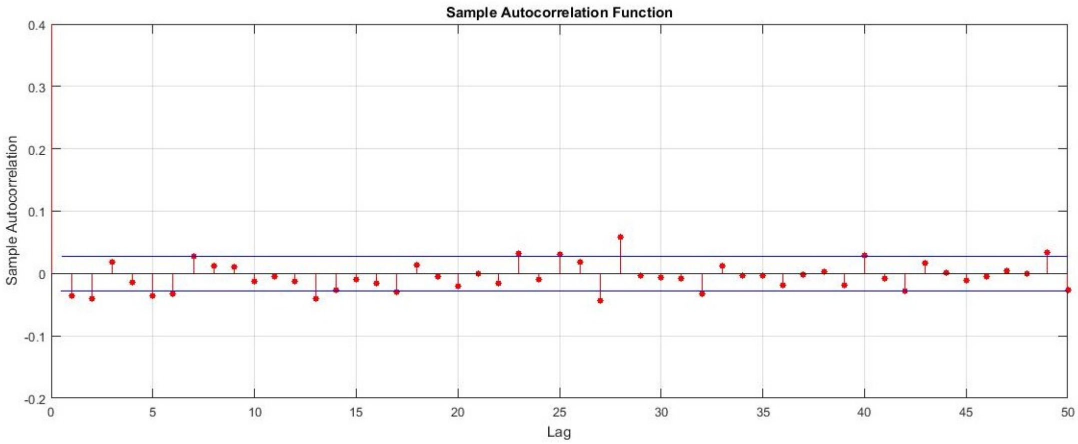

Figure 6: Sample autocorrelation for cleaned residuals

5.2. Forecasting

This section deals with predicting daily average electricity log-prices. We compare several forecasting strategies using a range of univariate time series models. The forecasts using the generalized exponential model are obtained with the finite autoregressive approximation of the model and the use of the Kalman filter as it is described in Appendix A.1.

We also address the question of whether the prices for the individual hours contain useful predictive information compared to the daily average price. We compare the forecasts obtained from separate models for each 24 hours of the day which are then averaged and compared with the forecasts from the aggregate model which is modeled directly for the daily average prices. Raviv et. al (2015), using a different model class, demonstrate that forecast averaging may be superior to direct modelling and forecasting of the (average) daily observations.

The comparative assessment of the predictive accuracy is based on a rolling forecasting experiment: starting from 30.01.2014, we estimate the model and forecast the next day average price; we then proceed by updating the sample by adding one observation, and deleting the one at the beginning of the sample, re-estimating the parameters and forecasting the next day prices, until the end of sample is reached. The evaluation sample is based on 365 one-step-ahead forecasts.

Following Raviv et al. (2015), we consider the ARX model as benchmark model[1]. The $ARX(p)$ model of order p is specified as follows:

$$y_t = \sum_{j=1}^{p} \phi_j y_{t-j} + \sum_{k=1}^{K} \psi_k d_{kt} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0, \sigma^2),$$

where d_{kt} are dummies for Saturdays, Sundays, and for each month of the year. The order p is chosen at every point in time according to the Akaike Information Criterion.

[1]The comparison has been also made with the best seasonal ARIMA benchmark, as it is reported in Sores, Sauza (2006). However this benchmark was outperformed by the ARX model.

The second model used for comparative purposes is the seasonal random walk with seasonal period 7 days, $y_t = y_{t-7} + \varepsilon_t$, according to which the next day prediction is $\hat{y}_{t+1|t} = y_{t-6}$.

Another relevant predictor is the seasonal Holt-Winters method (see Holt, 1957, Winters, 1960, and Hyndman et al., 2008); we have considered its additive formulation:

$$\begin{aligned}
&\hat{y}_{t+1|t} = l_t + b_t + s_{t-6} \\
&l_t = \alpha\,(y_t - s_{t-7}) + (1-\alpha)\,(l_{t-1} + b_{t-1}) \\
&b_t = \beta^*\,(l_t - l_{t-1}) + (1-\beta^*)\,b_{t-1} \\
&s_t = \gamma\,(y_t - l_{t-1} - b_{t-1}) + (1-\gamma)\,s_{t-7},
\end{aligned}$$

where α, β^* and γ denote smoothing parameters, taking values in [0,1], estimated by minimizing the one-step-ahead mean square prediction error within the training sample.

The statistics that are used for assessing the forecasting accuracy of the different methods and models are the mean absolute error (MAE) and the root mean squared error (RMSE),

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=T}^{H-1}\left(\hat{y}_{t+1|t} - y_{t+1}\right)^2}, \tag{12}$$

$$MAE = \frac{1}{N}\sum_{t=T}^{H-1}\left|\hat{y}_{t+1|t} - y_{t+1}\right|, \tag{13}$$

see Hyndman and Athanasopoulos (2012), where H is the total number of observations, T is the length of the estimation window, and $N = H - T$ is the number of out-of-sample forecasts.

Hence, our rolling forecast experiment compares four different predictors, including the one arising from our seasonal fractional exponential model, applied to the daily average price series. The four predictors are then applied to the 24 individual time series referring to the single hours of the day and the one-step-ahead forecasts are later aggregated into a daily average forecast. This yields a total of 8 models to be compared.

Table 5 contains different measures of forecasting accuracy for the models considered. Analogously to the work of Raviv et al. (2015), we provide the exact values of the forecasting accuracy measures only for the ARX model (benchmark). The performance of the other models is presented in relative terms, that is as the ratio of the accuracy measure for a specific model and that of the benchmark.

The results clearly show that the GEXP model outperforms the competing models, as the lowest value of MAPE and RMSE is for the GEXP model. The prices for individual hours contain useful predictive information for the daily average price. The forecasts based on sample averaging the forecasts from the GEXP model estimated separately for each particular hour are the most accurate. Moreover the GEXP model provides more accurate results than the univariate models considered in Raviv et al. (2015, page 236, Table 2). Some further improvements of daily average price forecasting might be achieved by playing with the weights when computing the average, see Raviv et al. (2015) for further details.

Table 5: Forecasting accuracy measures

	RMSE	MAE
ARX*	0.0666	0.0475
Seasonal HW	1.0339	1.0056
Seasonal RW	1.6209	1.6746
GEXP	**0.9351**	**0.9421**
ARX: averaging	0.9962	1.0050
Seasonal HW: averaging	1.0433	1.0286
Seasonal RW: averaging	1.6209	1.6746
GEXP: averaging	**0.9115**	**0.9277**

Note: The first row contains the values of forecasting accuracy measures for ARX model (benchmark), the remaining rows contain the ratio of the accuracy measure for a specific model and for the benchmark. The smaller the value the better the forecasting performance of a given model.

We test also for the significance of the difference in forecasting performance of the models with the use of the Giacomini and White test (see Giacomini and White, 2006, for details).

Table 6 includes the test results based on the absolute error loss, which corresponds to testing the null hypothesis of equal predictive ability in terms of MAE. The absolute error loss function was chosen for comparison with the work of Raviv et al (2015) and also to decrease the influence of the outliers on the final results. Similar results can be obtained with the square error loss function. A positive value means that the MAE of the model in the column is larger than that of the row model. The results confirm that the GEXP modelling approach outperforms competing models in terms of absolute error. Also, it can be seen that averaging of forecasts from different GEXP models for each hour of the day is better in terms of the forecasting performance strategy than forecasting the average electricity price directly with the GEXP model.

Table 6: Giacomini-White test statistics based on the absolute error loss function

	ARX	Seasonal HW	Seasonal RW	GEXP	ARX: averaging	Seasonal HW: averaging	Seasonal RW: averaging
Seasonal HW	-0.30						
Seasonal RW	-7.31	-7.06					
GEXP	**2.15**	**1.99**	**8.72**				
ARX: averaging	-0.33	0.02	7.62	-2.37			
Seasonal HW: averaging	-1.00	-0.88	7.30	-2.82	-0.93		
Seasonal RW: averaging	-7.31	-7.06	0.22	-8.72	-7.62	-7.30	
GEXP: averaging	**2.14**	**2.13**	**9.25**	**0.55**	**2.65**	**3.26**	**9.25**

Note: The test statistic is computed such that a positive value means that the MAE of the row model is smaller than the MAE of the column model.

In addition to the previous analysis, we also estimated the model confidence set (MCS), according to the methodology proposed by Hansen, Lunde and Nason (2011). A MCS is the set of models that contains the best model with a given level of confidence. The procedure consists of a sequence of tests which permits to construct a Superior Set of Models (SSM), where the null hypothesis of Equal Predictive Ability is not rejected at a given confidence level.

Given the initial set of models, M^0, comprising the 8 predictors, denote the prediction error arising from method k as $v_{kt} = y_t - \hat{y}_{k,t|t-1}$, and by $d_{kl,t} = |v_{kt}| - |v_{lt}|$ the loss differential at time t between predictors k and l, for all $k, l \in M^0$, and define

$$\bar{d}_{kl} = \frac{1}{n - n_0} \sum_{t=n_0+1}^{n} d_{kl,t}, \quad \bar{d}_k = \frac{1}{K} \sum_{l=1}^{K} \bar{d}_{kl}.$$

The t-statistics associated with the null $H_0 : \mathrm{E}(\bar{d}_{kl}) = 0$ (equal forecast accuracy) versus the alternative $H_0 : \mathrm{E}(\bar{d}_{kl}) > 0$, is $t_{kl} = \frac{\bar{d}_{kl}}{\hat{\mathrm{Var}}(\bar{d}_{kl})}$, where $\hat{\mathrm{Var}}(\bar{d}_{kl})$ is an estimate of $\mathrm{Var}(\bar{d}_{kl})$. Also, the t-statistics associated with the null $H_0 : \mathrm{E}(\bar{d}_{k.}) = 0$ (equal forecast accuracy) versus the alternative $H_0 : \mathrm{E}(\bar{d}_{k.}) > 0$, is $t_k = \frac{\bar{d}_{k.}}{\hat{\mathrm{Var}}(\bar{d}_{k.})}$, where $\hat{\mathrm{Var}}(\bar{d}_{k.})$ is an estimate of $\mathrm{Var}(\bar{d}_{k.})$. If M is the current set of model under assessment, to test the hypotheses $H_{0,M} : \mathrm{E}(\bar{d}_{kl}) = 0, \forall k, l \in M$ or $H_{0,M} : \mathrm{E}(\bar{d}_{k.}) = 0, \forall k \in M$, i.e. all the models have the same predictive accuracy, we use the test statistics $T_{R,M} = \max_{kl \in M} |t_{kl}|$ and $T_{\max,M} = \max_{k \in M} t_k$. We initially set $\mathcal{M} = \mathcal{M}^0$, and test $H_{0,M}$ using the above statistics at the significance level $\alpha = 0.10$. The critical values are obtain by the block-bootstrap method. If $H_{0,\mathcal{M}}$ is accepted, then the MCS at level $1 - \alpha$ is $\widehat{M}^*_{1-\alpha} = M$; otherwise, we proceed to eliminate from the set the predictor for which the $t_{k.}$ statistic is a maximum, $k \in M$, and iterate the procedure with the surviving predictors.

The MCS p-values of the $T_{R,M}$ and $T_{max,M}$ statistics, are reported in Table 7. The estimated SSMs differ for the number of the eliminated models as well as for their compositions. We can observe that only 2 forecasting strategies: seasonal random walk and seasonal random walk averaging were eliminated by the MCS procedure and hence the superior set of models contains 6 models. This empirical finding highlights the statistical equivalence of forecasting future daily electricity spot prices with the GEXP model and simple ARX or seasonal Holt-Winter model. However, the rankings of the models based on the two test statistics are highly in favor of the GEXP model. The GEXP models has taken the two first places for both considered test statistics. The results confirm that the best forecasting strategy is the one involving averaging of forecasts from different GEXP models for every hour of the day. The second place is taken by the approach based on direct forecasting average electricity price with the GEXP model.

Table 7: Superior Set of Models on 99% confidence level (2 models eliminated)

	$Rank_{R,M}$	$T_{R,M}$	$p-value_{R,M}$	$Rank_{max,R}$	$T_{max,R}$	$p-value_{max,R}$	$Loss$
ARX	3	0.4127	1	3	1.2651	1	0.0475
Seasonal HW	5	0.5658	1	5	1.3623	1	0.0477
GEXP	2	-1.1610	1	2	0.2533	1	0.0447
ARX: averaging	4	0.5498	1	4	1.3540	1	0.0477
HW: averaging	6	1.1889	1	6	1.7612	0.0024	0.0488
GEXP: averaging	1	-1.5468	1	1	-0.2533	1	0.0440

Note: Comparison of the superior set of models. The p–values of the $T_{R,M}$ and $T_{max,M}$ statistics, are reported in the fourth and seventh columns, respectively. The *p–value* of the test statistic, is equal to the minimum of the overall p–values. The columns $Rank_{R,M}$ and $Rank_{max,M}$ report the ranking over the models belonging to the SSMs. The last column *Loss* is the average loss across the considered period.

6. Conclusions

This paper has proposed a model for daily electricity spot prices from the Nord Pool power exchange. It can be applied also to other utilities presenting similar seasonal and long-memory pattern, eg. electricity demand or loads. The model is formulated in the frequency domain and captures the long range dependence and the persistent seasonal pattern of prices as well as the effect explanatory variables (eg. calendar effects component and water reservoir levels). One of the most challenging tasks when modelling electricity prices is to deal with extreme observations such as price spikes. To handle this problem we have proposed a novel estimation strategy based on a robust Kalman filter. The GEXP model provides a statistically coherent representation of the price dynamics. Price spikes significantly influence parameter estimates and robustifying the estimates can prove valuable. Whether using classical and robust estimates, the effect on the estimated memory appears to be minor, however.

A forecasting exercise was conducted to put the predictive ability of our model under test. The empirical evidence suggests that the best strategy for daily average spot electricity price forecasting is to construct separate models for each hour and average the forecasts, in line with findings of Raviv et al (2015).

References

[1] Nord Pool Spot, http://www.nordpoolspot.com, accessed: 2015-12-16,

[2] Arteche, J. and Robinson, M. P. (2000), *Semiparametric inference in seasonal and cyclical long memory processes*, Journal of Time Series Analysis, 21(1), 1–25.

[3] Benth, F. E. and Koekebakker, S. (2008), *Stochastic modeling of financial electricity contracts*, Energy Economics, 30(3), 1116–1157.

[4] Beran, J. (1993), *Fitting long-memory models by generalized linear regression*, Biometrika, 80, 817–822.

[5] Beran, J., Feng, Y., Ghosh, S. and Kulik, R. (2013), *Long-memory processes: Probabilistic Properties and Statistical Methods*, Berlin, Springer.

[6] Bloomfield, P. (1973), *An exponential model for the spectrum of a scalar time series*, Biometrika, 60(2), 217–226.

[7] Bogert, B., Healy, M. and Tukey, J. (1963), *The frequency analysis of time series for echoes: Spectrum, pseudo-autocovariance, cross-spectrum and shape cracking*, In Proc. Symp. on Time Series Analysis, p. 209–243.

[8] Dahlhaus, R. (1989), *Efficient parameter estimation for self-similar processes*, The Annals of Statistics, 17(4), 1749–1766.

[9] De Jong, C. (2006), *The nature of power spikes: A regime-switch approach*, Studies in Nonlinear Dynamics and Econometrics, 10(3).

[10] Diongue, K. A., Guégan, D. and Vignal, B. (2004), *A k-factor GIGARCH process: estimation and application on electricity market spot prices*, In Probabilistic Methods Applied to Power Systems, 2004 International Conference on, p. 1–7.

[11] Diongue, K. A., Guégan, D. and Vignal, B. (2009), *Forecasting electricity spot market prices with a k-factor GIGARCH process*, Applied Energy, 86(4), 505–510.

[12] Giacomini, R. and White, H. (2006), *Tests of conditional predictive ability* Econometrica, 74(6), 1545–1578.

[13] Gil-Alaña, L. A. and Robinson, P. M. (2001), *Testing of seasonal fractional integration in UK and Japanese consumption and income*, Journal of Applied Econometrics, 16(2), 95–114.

[14] Gradshteyn, I. S. and Ryzhik, I. M. (2007), *Table of integrals, series, and products*, Elsevier/Academic Press, Amsterdam, Seventh edition.

[15] Gray, L. H., Zhang, N.-F. and Woodward, A. W. (1989), *On generalized fractional processes*, Journal of Time Series Analysis, 10(3), 233–257.

[16] Haldrup, N and Nielsen, Ø. M. (2006), *Directional congestion and regime switching in a long memory model for electricity prices*, Studies in Nonlinear Dynamics & Econometrics, 10(3), September, 1–24.

[17] Haldrup, N. and Nielsen, Ø. M. (2006), *A regime switching long memory model for electricity prices*, Journal of Econometrics, 135(1-2), 349–376.

[18] Hansen, P. R., Lunde, A. and Nason, J. M. (2011), *The model confidence set*, Econometrica, 79(2), 453–497.

[19] Holt, C. C. (1957), *Forecasting seasonals and trends by exponentially weighted moving averages*, Technical report, Carnigie Institute, ONR Research Memorandum.

[20] Hosking, J.R.M. (1981), *Fractional differencing*, Biometrika, 68(1), 165–176.

[21] Hsu, N.J. and Tsai, H. (2009), *Semiparametric estimation for seasonal long-memory time series using generalized exponential models*, Journal of Statistical Planning and Inference, 139(6), 1992–2009.

[22] Hurvich, C. M. (2002), *Multistep forecasting of long memory series using fractional exponential models*, International Journal of Forecasting, 18(2), 167–179.

[23] Hyndman, R. J. and Athanasopoulos, G. (2012), *Forecasting: principles and practice.*

[24] Janacek, G. J. (1982), *Determining the degree of differencing for time series via the log spectrum*, Journal of Time Series Analysis, 3(3), 177–183.

[25] Janczura, J., Trück, S. ,Weron, R. and Wolf, R. C. (2013), *Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling*, Energy Economics, 38(0), 96– 110.

[26] Koopman, S. J., Ooms, M. and Carnero, M. A. (2007), *Periodic seasonal Reg-ARFIMA-GARCH models for daily electricity spot prices*, Journal of the American Statistical Association, 102(477), 16–27.

[27] Martin, R.D. and Thomson, D.J. (1982), *Robust-resistant spectrum estimation*, Proceedings of the IEEE, 70(9):1097–1115.

[28] Masreliez, C. and Martin, R. (1977), *Robust bayesian estimation for the linear model and robustifying the kalman filter*, Automatic Control, IEEE Transactions on, 22(3), 361–371.

[29] McCloskey, A. and Hill, J. B. (2013), *Parameter estimation robust to low-frequency contamination*, Journal of Business & Economic Statistics.

[30] Möst, D. and Keles, D. (2010), *A survey of stochastic modeling approaches for liberalised electricity markets*, European Journal of Operational Research, 207(2), 543–556.

[31] Narukawa, M. and Matsuda, Y. (2011), *Broadband semi-parametric estimation of long-memory time series by fractional exponential models*, Journal of Time Series Analysis, 32(2), 175–193.

[32] Pearlman, J. G. (1980), *An algorithm for the exact likelihood of a high-order autoregressive-moving average process*, Biometrika, 67, 232–233.

[33] Pourahmadi, M. (1983), *Exact factorization of the spectral density and its application to forecasting and time series analysis*, Communications in Statistics - Theory and Methods, 12(18), 2085–2094.

[34] Raviv, E., Bouwman, K. E. and Van Dijk, D. (2015), *Forecasting day-ahead electricity prices: Utilizing hourly prices*, Energy Economics, 50, 227–239.

[35] Robinson, P. M. (1991), *Nonparametric function estimation for long memory time series*, In Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proc. of the 5th Int. Symposium in Economic Theory and Econometrics, p. 437–457.

[36] Serati, M., Manera, M. and Plotegher, M. (2008), *Modeling electricity prices: from the state of the art to a draft of a new proposal*, Liuc papers in Economics, Cattaneo University (LIUC).

[37] Smallwood, A. and Beaumont, P. (2003), *Multiple frequency long memory models*, Working paper, Dept. of Economics, University of Oklahoma.

[38] Soares, L. J. and Souza, L. R. (2006), *Forecasting electricity demand using generalized long memory*, International Journal of Forecasting, 22(1), 17–28.

[39] Velasco, C. (1999), *Non-stationary log-periodogram regression*, Journal of Econometrics, 91(2), 325–371.

[40] Velasco, C. and Robinson, P. M. (2000), *Whittle pseudo-maximum likelihood estimation for nonstationary time series*, Journal of the American Statistical Association, 95(452), 1229–1243.

[41] Walker, A. M. (1964), *Asymptotic properties of least-squares estimates of parameters of the spectrum of a stationary non-deterministic time-series*, Journal of the Australian Mathematical Society, 4(3), 363–384.

[42] Woodward, A. W., Cheng, C. Q. and Gray, L. H. (1998), *A k-factor GARMA long-memory model*, Journal of Time Series Analysis, 19(4), 485–504.

[43] Weron, R. (2006), *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, Wiley, Chichester.

[44] Weron, R. (2008), *Market price of risk implied by asian-style electricity options and futures*, Energy Economics, 30(3), 1098–1115.

[45] Weron, R. (2014), *Electricity price forecasting: A review of the state-of-the-art with a look into the future*, International Journal of Forecasting, 30(4), 1030–1081.

[46] Weron, R., Bierbrauer, M. and Trück, S. (2004), *Modeling electricity prices: jump diffusion and regime switching*, Physica A: Statistical Mechanics and its Applications, 336(1–2), 39–48.

[47] Weron, R., Simonsen, I. and Wilman, P. (2004), *Modeling highly volatile and seasonal markets: evidence from the Nord Pool electricity market*, In Hideki Takayasu, The Application of Econophysics, p. 182–191. Springer Japan.

[48] Whittle, P. (1953), *Estimation and information in stationary time series*, Arkiv för Matematik, 2(5), 423–434.

[49] Winters, P. R. (1960), *Forecasting sales by exponentially weighted moving averages*, Management Science, p. 324–342.

[50] Woodward, W. A., Gray, H. L. and Elliott, A. C. (2011), *Applied Time Series Analysis*, CRC Press, 1 edition.

[51] Zareipour, H. (2012), *Short-term electricity market prices: A review of characteristics and forecasting methods*, In Alexey Sorokin, Steffen Rebennack, Panos M. Pardalos, Niko A. Iliadis, and Mario V. F. Pereira editors, Handbook of Networks in Power Systems I, Energy Systems, p. 89–121. Springer Berlin Heidelberg.

Appendix A. Robust filtering and forecasting

Robust spectrum estimation and robust forecasting are based on the $AR(m)$ or $MA(m)$ approximation of the GEXP process. In general, an $ARMA(p,q)$ time series model for $u_t = y_t - x_t'\beta$,

$$u_t + \phi_1 u_{t-1} + \cdots + \phi_p u_{t-p} + \xi_t + \theta_1 \xi_{t-1} + \cdots + \theta_q \xi_{t-q}, \xi_t \sim \mathrm{WN}(0, \sigma^2),$$

can be written in state space form with measurement equation,

$$u_t = Z\alpha_t + G\xi_t, \qquad t = 1, \ldots, n, \tag{A.1}$$

where α_t is a random vector with $m = \max(p,q)$ elements, $Z = [1, 0, \ldots, 0]$, $G = 1$. The evolution of the states is governed by the transition equation:

$$\alpha_{t+1} = T\alpha_t + H\xi_t, \qquad t = 1, 2, \ldots, n, \tag{A.2}$$

where

$$T = \begin{bmatrix} -\phi_1 & 1 & 0 & \cdots & 0 \\ -\phi_2 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \cdots & \cdots & 0 & 1 \\ -\phi_m & 0 & \cdots & \cdots & 0 \end{bmatrix}, H = \begin{bmatrix} \theta_1 - \phi_1 \\ \theta_2 - \phi_2 \\ \vdots \\ \vdots \\ \theta_m - \phi_m \end{bmatrix}.$$

The initial state vector, α_1, assuming stationarity (the eigenvalues of T are inside the unit circle), has a distribution with mean $\mathrm{E}(\alpha_1) = 0$ and variance $\mathrm{Var}(\alpha_1) = \sigma^2 P_{1|0}$, satisfying the matrix equation $P_{1|0} = TP_{1|0}T' + HH'$.

The above state space representation is due to Pearlman (1980), and encompasses both the pure AR and MA case.

Appendix A.1. The Kalman filter

Assume that the process is Gaussian, define $\mathcal{F}_t = \{u_1, u_2, \dots, u_t\}$, the information set up to and including time t, $\tilde{\alpha}_{t|t-1} = \mathrm{E}(\alpha_t|\mathcal{F}_{t-1})$, and $\mathrm{Var}(\alpha_t|\mathcal{F}_{t-1}) = \sigma^2 P_{t|t-1}$.

The Kalman filter (KF) is the following recursive algorithm: for $t = 1, \dots, n$,

$$\begin{aligned}
\nu_t &= u_t - Z\tilde{\alpha}_{t|t-1}, & f_t &= ZP_{t|t-1}Z' + GG',\\
& & C_t &= P_{t|t-1}Z'f_t^{-1},\\
\tilde{\alpha}_{t|t} &= \tilde{\alpha}_{t|t-1} + C_t\nu_t, & P_{t|t} &= P_{t|t-1} - C_t f_t C_t',\\
& & Q_t &= HG'f_t^{-1},\\
\tilde{\alpha}_{t+1|t} &= T\tilde{\alpha}_{t|t} + Q_t\nu_t, & P_{t+1|t} &= TP_{t|t}T' + HH' - (Q_t f_t Q_t' + Q_t f_t C_t' T' + TC_t f_t Q_t').
\end{aligned} \tag{A.3}$$

The above equations compute the innovations $\nu_t = u_t - \mathrm{E}(u_t|\mathcal{F}_{t-1})$, and $\sigma^2 f_t$ is the prediction error variance at time t, that is $\mathrm{Var}(u_t|\mathcal{F}_{t-1})$; $\tilde{\alpha}_{t|t}$ are the updated, or real time, estimates of the state vector, and C_t is the gain, $C_t = \mathrm{Cov}(\alpha_t, y_t|Y_{t-1})[\mathrm{Var}(y_t|Y_{t-1})]^{-1}$ and we can prove the following result: $\alpha_t|Y_t \sim \mathrm{N}(\tilde{\alpha}_{t|t}, \sigma^2 P_{t|t})$. The vector Q_t represents $\mathrm{Cov}(\xi_t, u_t|\mathcal{F}_{t-1})[\mathrm{Var}(u_t|\mathcal{F}_{t-1})]^{-1}$, so that $\tilde{\alpha}_{t+1|t}$ is the one-step-ahead state prediction and we can write $\alpha_{t+1}|\mathcal{F}_t \sim \mathrm{N}(\tilde{\alpha}_{t+1|t}, \sigma^2 P_{t+1|t})$.

Appendix A.2. The robust filter

Masreliez, Martin (1977) and Martin, Thomson (1982) proposed to obtain a robust filter is to modify the Kalman filter updating and prediction equations (A.3) by using a bounded and continuous function of the standardised innovations that so as to control the effects of outliers on the conditional mean estimates.

Let us denote $\tilde{\nu}_t = \frac{\nu_t}{f_t^{1/2}}$. The above KF is modified as follows: for $t = 1, \dots, n$,

$$\begin{aligned}
\nu_t &= u_t - Z\tilde{\alpha}_{t|t-1}, & f_t &= ZP_{t|t-1}Z' + GG',\\
& & C_t &= P_{t|t-1}Z'f_t^{-1},\\
\tilde{\alpha}_{t|t} &= \tilde{\alpha}_{t|t-1} + C_t f_t^{1/2}\psi(\tilde{\nu}_t), & P_{t|t} &= P_{t|t-1} - w(\tilde{\nu}_t)C_t f_t C_t',\\
& & Q_t &= HG'f_t^{-1},\\
\tilde{\alpha}_{t+1|t} &= T\tilde{\alpha}_{t|t} + Q_t f_t^{1/2}\psi(\tilde{\nu}_t), & P_{t+1|t} &= TP_{t|t}T' + HH' - w(\tilde{\nu}_t)(Q_t f_t Q_t' + Q_t f_t C_t' T' + TC_t f_t Q_t').
\end{aligned} \tag{A.4}$$

Here $\psi(u)$ is the Hampel's two-part redescending function:

$$\psi(u) = \begin{cases} u, & |u| \le a,\\ \mathrm{sign}(u)\frac{a}{b-a}(b - |u|), & a < |u| \le b,\\ 0, & |u| > b, \end{cases}$$

for $a < b$. The weight function is $w(u) = \psi(u)/u$. Note that if $\psi(u) = u$, the identity function, $w(u) = 1$ and the above recursions yield the Kalman filter (A.3).

A clean estimate of y_t is then

$$\tilde{u}_{t|t} = Z\tilde{\alpha}_{t|t} + GG'f_t^{-1/2}\psi(\tilde{\nu}_t). \tag{A.5}$$

Noticing that $u_t = Z\tilde{\alpha}_{t|t} + GG' f_t^{-1}\nu_t$, we can write $\tilde{u}_{t|t} = u_t - GG' f_t^{-1}\nu_t[1 - w(\tilde{\nu}_t)]$, which shows that $\tilde{u}_{t|t} = u_t$ when $w(u) = 1$, which occurs for $|u| < a$, i.e. when the standardized innovations are small. On the contrary, $w(u) = 0$, which takes place for $|u| > b$, simple manipulations show that in the presence of a large residual, the cleaned observations is shrunk towards the one step ahead prediction:

$$\tilde{u}_{t|t} = [1 - GG' f_t^{-1}]u_t + GG f_t^{-1}\tilde{u}_{t|t-1}, \quad \tilde{u}_{t|t-1} = Z\tilde{\alpha}_{t|t-1}.$$

It can be shown that the steady state Kalman filter for the AR or MA model considered has $\lim_{t\to\infty} f_t = GG'$, and thus, after processing a large number of observations, the occurrence of a large outlier causes $\tilde{u}_{t|t} \to \tilde{u}_{t|t-1}$. Finally, when $a < |\tilde{\nu}_t| \leq b$, $\tilde{u}_{t|t}$ arises a weighted linear combination of u_t and $\tilde{u}_{t|t-1}$.

The theoretical underpinnings of the robust KF are provided in Masreliez and Martin (1977).

Research Papers 2016

2015-51: Tommaso Proietti: Exponential Smoothing, Long Memory and Volatility Prediction

2015-52: Mark Podolskij, Christian Schmidt and Mathias Vetter: On U- and V-statistics for discontinuous Itô semimartingale

2015-53: Mark Podolskij and Nopporn Thamrongrat: A weak limit theorem for numerical approximation of Brownian semi-stationary processes

2015-54: Peter Christoffersen, Mathieu Fournier, Kris Jacobs and Mehdi Karoui: Option-Based Estimation of the Price of Co-Skewness and Co-Kurtosis Risk

2015-55 Kadir G. Babaglou, Peter Christoffersen, Steven L. Heston and Kris Jacobs: Option Valuation with Volatility Components, Fat Tails, and Nonlinear Pricing Kernels

2015-56: Andreas Basse-O'Connor, Raphaël Lachièze-Rey and Mark Podolskij: Limit theorems for stationary increments Lévy driven moving averages

2015-57: Andreas Basse-O'Connor and Mark Podolskij: On critical cases in limit theory for stationary increments Lévy driven moving averages

2015-58: Yunus Emre Ergemen, Niels Haldrup and Carlos Vladimir Rodríguez-Caballero: Common long-range dependence in a panel of hourly Nord Pool electricity prices and loads

2015-59: Niels Haldrup and J. Eduardo Vera-Valdés: Long Memory, Fractional Integration, and Cross-Sectional Aggregation

2015-60: Mark Podolskij, Bezirgen Veliyev and Nakahiro Yoshida: Edgeworth expansion for the pre-averaging estimator

2016-01: Matei Demetrescum, Christoph Hanck and Robinson Kruse: Fixed-b Inference in the Presence of Time-Varying Volatility

2016-02: Yunus Emre Ergemen: System Estimation of Panel Data Models under Long-Range Dependence

2016-03: Bent Jesper Christensen and Rasmus T. Varneskov: Dynamic Global Currency Hedging

2016-04: Markku Lanne and Jani Luoto: Data-Driven Inference on Sign Restrictions in Bayesian Structural Vector Autoregression

2016-05: Yunus Emre Ergemen: Generalized Efficient Inference on Factor Models with Long-Range Dependence

2016-06: Girum D. Abate and Luc Anselin: House price fluctuations and the business cycle dynamics

2016-07: Gustavo Fruet Dias, Cristina M. Scherrer and Fotis Papailias: Volatility Discovery

2016-08 N. Haldrup, O. Knapik and T. Proietti: A generalized exponential time series regression model for electricity prices