
Munich Personal RePEc Archive

Modeling hourly Electricity Spot Market Prices as non stationary functional times series

Liebl, Dominik

University of Cologne

September 2010

Online at https://mpra.ub.uni-muenchen.de/25017/
MPRA Paper No. 25017, posted 15 Sep 2010 10:58 UTC

Modeling hourly Electricity Spot Market Prices
as non stationary functional times series

Dominik Liebl*

University of Cologne

Abstract

The instantaneous nature of electricity distinguishes its spot prices from spot prices for equities and other commodities. Up to now electricity cannot be stored economically and therefore demand for electricity has an untempered effect on electricity prices. In particular, hourly electricity spot prices show a vast range of dynamics which can change rapidly. In this paper we introduce a robust version of functional principal component analysis for sparse data. The functional perspective interprets spot prices as functions of demand for electricity and allows to estimate a single price curve for each day. Variations in market fundamentals such as commodity prices are absorbed by the first principal components.

Keywords: Functional principal component analysis, non stationary functional time series data, sparse data, electricity spot market prices, European Electricity Exchange (EEX).

1. Introduction

Spot prices for electricity are peculiar. Up to now electricity cannot be stored economically therefore the amount of electricity that can be used for arbitrage over time can be neglected and demand for electricity has an untempered effect on electricity prices. The pricing in the power market is based on marginal generation costs of the last power plant that is required to cover the demand. The supply curve is based on the increasing generation costs of the installed power

*Earlier versions of this paper have been presented at the DAGStat 2010 in Dortmund and the Statistische Woche 2010 in Nürnberg. I want to thank especially Prof. Alois Kneip for fruitful discussions, Prof. Pascal Sarda, who encouraged me constructively on my way to write this paper during my stay at the working group STAPH in Toulouse, and Prof. Wolfgang Härdle for constructive discussion at the DATGStat 2010.

Corresponding Author: Dominik Liebl (liebl@wiso.uni-koeln), Seminar für Wirtschafts– und Sozialstatistik, Lehrstuhl Prof. Mosler, Universität zu Köln, Albertus–Magnus–Platz, 50923 Köln Germany

plants, with a steeply increasing, exponential shape. Usually base load plants such as nuclear and lignite plants cover the minimal load, i.e. demand for electricity, throughout the year. Load following is mostly done by medium and peak load plants such as hard-coal and gas-fired power plants. Pricing above marginal costs, so called peak load pricing, is typical in markets with non sortable products. The reason behind this are opportunity costs and incremental costs from constraints, such as those arising from emission and capacity limits, which then become marginal costs relevant (Cramton, 2004). This deviation from variable cost pricing is also observable on the low demand side. Plant operators try to avoid shutting off power plants in order to avoid ramp up costs and therefore bid occasionally below variable costs, although latter becomes only visible if it is allowed to sell negative prices are allowed to trade.

In micro economic theory it is common to set prices equal to the marginal cost of production and to determine the equilibrium prices by the interaction of demand and supply curves. Particularly, electricity spot prices are usually modeled by the Cournot model or the Supply Function Equilibrium model (Klemperer and Meyer, 1989). A recent example is the paper of Willems et al. (2009) an other well known example is the paper of Green and Newbery (1992). Functional data analysis (fda) is able to share this perspective, since its atomic statistical units are functions rather than points and/or vectors. The books of Ramsay and Silverman (2005) and Ferraty and Vieu (2006) give a broad overview to functional data analysis. As explained above, we have to distinguish between pricing based on marginal generation costs and pricing based on opportunity costs. The functional approach can model the marginal cost system well since marginal generation costs are strongly connected to the demand for electricity. The opportunity costs system has to be modeled separately and both models may be connected by a kind of regime switching mechanism.

Markov regime-switching models, such as in Mount et al. (2006) and Kosater and Mosler (2006), became one of the most applied approaches in modeling and forecasting electricity prices. These try to divide the series into regimes with own mean and covariance structures. But the supply curve induces a continuum of mean and covariance regimes and it is therefore difficult to assign prices to certain regimes, even for the often used less volatile daily or half-daily averages of spot prices. One of the few high frequency analysis is done by Karakatsani and Bunn (2008); Karakatsani and Bunn model and forecast electricity spot prices from the UK-Power Exchange. They divide their data into sub samples for the 48 half hourly trading periods and additionally separate weekdays from weekends. Within such sub samples, daily demand values for electricity have got a clear smooth sinoidal pattern over the year with higher/lower demand values during the winter/summer months. The demand pattern is translated into a price pattern by the time invariant shape of the supply curve. This translation causes additional distortions and that may be the reason why the authors present their results for the trading periods 25 (12:30pm) and 35 (17:30pm) with low varying demand patterns. Furthermore, electricity from volatile renewable

energy sources like wind and solar is often provided with a purchase guarantee. If this is the case, then not the smooth demand patterns are relevant for the price patterns but the rougher adjusted demand patterns with netted out hourly electricity infeeds from renewable energy sources. Regarding the role of the supply curve as a diffuser of the (adjusted) demand pattern, we propose to focus on the estimation of its shape rather than on the estimation of the price pattern directly.

We propose to use (functional) principal component analysis to fit a low dimensional factor model to the daily supply curves. From the methodological perspective, the Dynamic Semiparametric Factor Model (DSFM) of Park et al. (2009) and the follow up application to electricity spot prices Härdle and Trück (2010) are very close to our approach. Park et al. use an iterating optimization algorithm to fit an orthogonal factor model to the data and argue that (functional) principal component analysis may not be able to handle sparse and non stationary (functional) time series data. We extend the procedure of Yao et al. (2005), that is able to handle sparse data, to the context of non stationary time series data. Härdle and Trück estimate a factor model to the daily N dimensional electricity spot price vectors $Y_t = (Y_1, \ldots, Y_{tN})'$ and report that a three dimensional factor model explains about 80% of the variation in hourly spot prices at the European Electricity Exchange (EEX). As the above cited papers, this application of the DSFM model focuses on modeling and forecasting the price pattern directly and has to use a regime switching mechanism in order to cope with the vast price-diffusions from the interactions of demand patterns with the supply curve.

The estimation of the supply curves has a limitation that is imposed by the auction design. Even though for each day t there are N supply curves, one for each trading period $h = 1, \ldots, N$, we can only estimate the mean supply curve of day t. The general auction design is similar for the greatest electricity markets places like in the Netherlands, Germany, Austria, Scandinavian countries, France, and California. It is a two-sided single-price auction, which means that there are bids from the purchase and sell side, that are matched at a singe market clearing price. The price for each trading period, $h = 1, \ldots, 24$, of a day t is settled by a separate auction, and all period specific auctions of a day t are conducted simultaneously the day before. The traders register the amounts of electricity they are willing to sell/purchase for individually selectable price intervals in a trading tool, where each trading period is represented by a new input line. Note that they base their bid-decisions on the same information set for all period specific auctions. The price settlement mechanism determines for each trading period, h, the spot price, Y_{th}, by the point of intersection of the (over all bids) aggregated supply and demand curves. More details about auction designs at power exchanges can be found among others in the book of Rafal Weron (2006). The horizontally shifts of the N demand curves reveal the shape a daily mean supply curve by the price vector, $Y_t = (Y_1, \ldots, Y_{tN})'$, and the according (residual) demand values. Figure 1 shows the raw data of

three consecutive days of two different (arbitrary) weeks – obviously, the supply curves possess depend on former supply curves and form a (functional) time series data set.

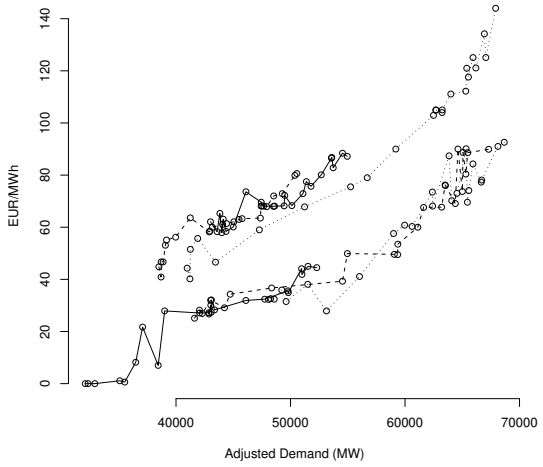

Figure 1: Three consecutive days from two different arbitrary weeks.

Since we estimate the supply curves from the hourly spot prices, Y_{th}, at the European Electricity Exchange (EEX), we will generally refer to the curves as price curves. The market of the EEX has got a high share of producers of electricity from volatile renewable energy sources (mainly wind), who feed their electricity directly into the grid and receive a certain guaranteed price. Adjusted demand, u, shall reflect the price relevant residual amount of electricity that is demanded from the conventional market participants. We assume that the hourly spot prices for electricity at the EEX come from an underlying smooth process, such that

$$Y_{ti} \quad = \quad X_t(u_{ti}) + \varepsilon_{ti}, \tag{1}$$

where $X_t(.)$ is a smooth monotone function of adjusted demand $u \in \mathcal{U}$, where $\mathcal{U}$ is a closed and bounded subspace of $\mathbb{R}$, we will set, without loss of generaltiy, $\mathcal{U} = [0, 1]$. The index $i = 1, \ldots, 24$ in u_{ti} refers to the i-th order statistic of the hourly adjusted demand values defined as $u_{th} = d_{th} - p_{th}$, where d_{th} is the gross demand for electricity and p_{th} is the corresponding infeed of electricity from wind energy at day $t = \{0, \pm, 1, \pm 2, \ldots\}$ in hour $h = 1, \ldots, N$, with $N = 24$. The noise term, ε_{ti}, is assumed to be independently distributed within and between each day t, with $E(\varepsilon_{ti}) = 0$ and a heteroscedastic $Var(\varepsilon_{ti}) = \sigma^2_\varepsilon(u_{it})$. The within independence is realistic since the hourly prices of the day t are determined contemporaneously at 12 o'clock at day $t-1$. The between independence

and a model for the heteroscedasticity follows from the error-in-variables discussion in the next paragraph.

As it can be seen in figure 1 the model from equation (1) is supported by the raw data. There are some remarkable strong deviations, especially for high and low values of adjusted demand. This comes from an inherent inaccuracy of adjusted demand values, u_{ti}. The spot market of electricity is actually an one-day-ahead future market, and the market participants (i.e. the traders) base their decisions on their own hourly forecast values of adjusted demand. Instead of these price relevant but unobservable forecast values we have to form the adjusted demand values, u_{ti}, from the actual realized values of gross demand, d_{ti}, and actual production of electricity from renewable energy sources, p_{ti}. Formally, we have to deal with an error-in-variables problem and formalize the noisy covariates as $u_{ti} = w_{ti} + \nu_{ti}$, where w_{ti} are the unobservable price relevant adjusted demand values, and the noise term ν_{ti} is assumed to have $E(\nu_{ti}) = 0$ and $Var(\nu_{ti}) = \sigma_\nu^2$ for all $t = \{0, \pm 1, \pm 2, \dots\}$ and $i = \{1, \dots, 24\}$. This inaccuracy causes stronger distortions at low and high values of adjusted demand where the price curves have got higher slopes than for moderate values of adjusted demand. Actually, this is a degenerated case of an error-in-variable problem, since the dependent variable, Y_{ti}, is observed nearly without noise. We assume that the noise in the observations of Y_{ti} is negligible and that we can translate the error-in-variables problem into an usual estimation problem with heteroscedastic error terms. A more sophisticated estimation procedure is beyond the scope of our paper, but might be a topic for future studies.

Our aim is to estimate the daily price functions X_t from the discrete data vector $\mathbf{Y}_t = (Y_{t1}, \dots, Y_{tN_t})'$. We use N_t to refer to the amount of prices per day t that are used to estimate the function X_t since some prices are assigned to the opportunity regime. We use a parsimonious ex-post assignment, price above a certain threshold are classified to the opportunity regime, all others are classified to the marginal cost regime. A reasonable threshold seems to be 145 EUR, since for prices > 145 EUR the traders lose their continuous marginal cost reference and begin to bid in clusters $(150 \text{ EUR}, 200 \text{ EUR}, 250 \text{ EUR}, \dots)$ of prices (see left panel of figure 2). The right panel of figure 2 shows all prices classified to the marginal cost regime.

Generically, we assume the daily price curves to come from a stochastic process (X_t) for $t = \{0, 1, 2, \dots\}$ with realizations in the space of square integrable functions $\mathcal{H} = L^2(\mathcal{U})$ on a compact set $\mathcal{U} \subset \mathbb{R}$. As in multivariate statistics, stationary functional stochastic process are usually described by their time invariant mean function and covariance operator. However the series of price curves has got a clear stochastic, non stationary trend. The curves inherit this property from the non stationary prices for raw materials (such as gas, coal, and Co2-certificates) that are needed to produce electricity. The (functional) random walk model,

$$X_t = \delta + X_{t-1} + e_t, \qquad \text{with } t = \{0, 1, 2, \dots\} \tag{2}$$

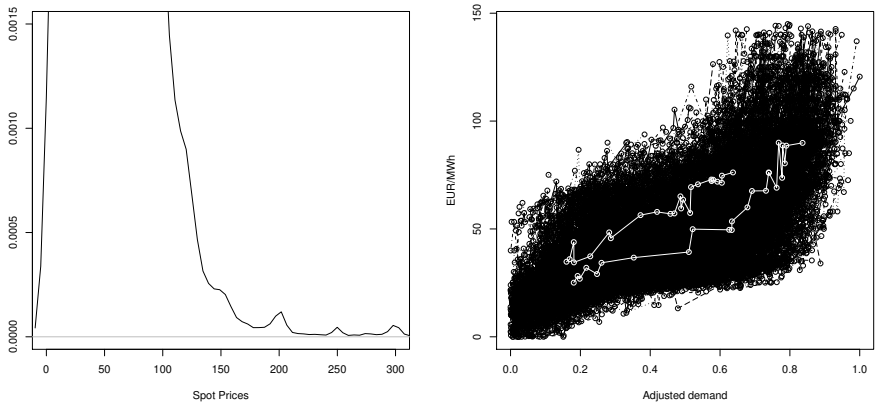

Figure 2: Left Panel: Non parametric density estimation of electricity spot prices (Note: Only prices $\leq$ 325 EUR are plotted, but there are prices up to 2437 EUR). Right Panel: Pooled spot prices classified to the marginal regime, with two highlighted price vectors from two arbitrary days.

with a linear trend, where $\delta \in \mathcal{H}$, an initial value from a random function, Z_0, that is normally distributed with mean, μ_Z, and covariance operator, Γ_Z, and a white noise process, $(e_t) \in \mathcal{H}$, with mean zero and covariance operator, Γ_e, seems most appropriate. Note that, as in the univariate case the $\mathrm{E}X_t = \mu_Z$ for all t, but the covariance operator depends on t, such that the process defined in (2) is non stationary. This is a special case of the so called ARH(1) model, i.e. an auto regressive model in $\mathcal{H}$. Note that any ARH(p), with $p > 1$, model can be transformed into an ARH(1) model, such that the above model is not necessarily a restriction with respect to the lag order (see Bosq (2000)). Furthermore, as usual practice in multivariate time series analysis we do not apply functional moving average terms in order to avoid identification problems with the AR terms.

The mean function of the functional random walk is independent of t and we can investigate, without loss of generality, the properties of the demeaned process $(X_t^*) = (X_t - \mu_Z)$. This yields the same functional random walk process as in equation (2) but with an initial functional random variable Z_0^* that has the zero function as it's mean; we write $X_t^* = \delta + X_{t-1}^* + e_t$. As already noted in Park et al. (2009), we need stationary of the time series in order to use the well developed functional principal component analysis. Traditional transformation procedures such as differentiation of the time series, (X_t^*), cannot be used, because the prices are observed at non equidistant adjusted demand values, u_{ti}. We propose a new transformation procedure that decomposes the original series, (X_t^*), into its stationary functional component, $(\tilde{X}_t^*)$, and its non stationary univariate random walk component, (Θ_t). The decomposition is motivated by the unit sphere projection of functional data (see Locantore et al.

(1999) and Gervini (2008)).

In order to use a notation, that corresponds to the matrix notation of multivariate statistics, we introduce the tensor product notation defined as $\tilde{X}_t \otimes \tilde{X}_t(u,v) = \tilde{X}_t(u)\tilde{X}_t(v)$, for $(u,v) \in \mathcal{U} \times \mathcal{U} \in \mathbb{R}^2$ and $t = \{0,1,2,\dots\}$. Then, the covariance function of a stationary functional series, $(\tilde{X}_t)$, can be written as $\tilde{\rho}(u,v) = E\left[(\tilde{X}_t - \mu_t) \otimes (\tilde{X}_t - \mu_t)\right](u,v)$. The covariance operator is defined as $\Gamma_{\tilde{X}_t} f(u) = \int \tilde{\rho}(u,v) f(v) dv$, for any function $f \in \mathcal{H}$. Its spectral decomposition by its eigenvalues, $\tilde{\lambda}_1 > \tilde{\lambda}_2 > \dots$, and corresponding eigenfunctions, $\phi_1, \phi_2, \dots$, with the usual restrictions $\int_{\mathcal{U}} \phi_k^2 = 1$ and $\int_{\mathcal{U}} \phi_k \phi_m = 0$ for $m < k$, allows us to write the functionals $\tilde{X}_t$ by the well known Karhunen-Loève decomposition as

$$\tilde{X}_t(u) = \mu_Z(u) + \sum_{k=1}^{\infty} \beta_{tk} \phi_k(u).$$

Where $\mu_Z = E(\tilde{X}_t)$ and $\beta_{tk} = \int_{\mathcal{U}} (\tilde{X}_t - \mu_Z)\phi_{tk}$ are the principal component scores with $E(\beta_{tk}) = 0$ and $E(\beta_{tk}^2) = \lambda_k$. Not least because of the best basis property (in the mean square error sense) of the Karhunen-Loève decomposition, often a relatively small number of eigenfunctions provides a good fit to the sample curves.

One problematic fact regarding our data is that the adjusted demand values u_{tj} are not uniformly distributed over the whole domain $\mathcal{U}$, but may be clustered at sub-intervals within $\mathcal{U}$. This makes it difficult to approximate the integrals in $\beta_{tk} = \int_{\mathcal{U}} (\tilde{X}_t - \mu_Z)\phi_{tk}$ by traditional methods like the trapezoidal rule where $\beta_{tk} \approx \sum_{j=1}^{N_t} Y_{tj} \phi_k(u_{tj})(u_{tj} - u_{t,j-1})$, with $u_{t0} = 0$. This is the same problem as in the so called sparse data problem in functional data analysis where it is difficult to approximate integrals from insufficiently many data points (see e.g. Yao et al. (2005)). Yao et al. suggest to estimate the principal component scoresβ_{tk} by their conditional expectation given the sparse data $\{u_{t1}, \dots, u_{tN_t}\}$. This procedure works very well for our non uniform distributed u_{tj}.

In the next section we introduce a new decomposition of a non stationary functional time series into its stationary spherical component and into its non stationary scaling component. In the section 3 we discuss how to estimate the mean function, μ_Z, the covariance function, $\rho(u,v)$, which is the kernel of the covariance operator, $\Gamma_X f(u) = \int_{\mathcal{U}} \rho(u,v) f(v) dv$, and the standard deviation, σ. In subsection 3.1 we estimate the eigenfunctions of the covariance operator and discus their interpretations. The subsection 3.2 adapts the conditional estimation of principal component scores from Yao et al. (2005) in order to estimate the pc-scores based on non uniformly (on $\mathcal{U}$) distributed data, u_{ti}. Finally, we demonstrate the goodness of fit of our estimation procedure in subsection 3.3.

2. Principal component analysis for non stationary data

From a practical point of view it would be a great advantage to project the infinite dimensional ARH(1) process into a finite dimensional functional space, $\mathcal{P}$, spanned by K basis functions, $\phi_1, \ldots, \phi_K$, such that the mean squared error of the projection,

$$\sum_{t=1}^{T}\sum_{j=1}^{N_t}\left\{X_t^*(u_{tj}) - X_t^{*,K}(u_{tj})\right\}^2 \text{ with } X_t^{*,K} = \sum_{k=1}^{K}\beta_{tk}\phi_k \quad \text{and } \beta_{tk} = \int_{\mathcal{U}} X_t^*\phi_k, \tag{3}$$

is minimized, where $(X_t^*) = (X_t - \mu_Z)$. If the series (X_t^*) corresponds (at least with high accuracy) to a K dimensional functional time series, $(X_t^{*,K})$, we could transform the infinite dimensional process, (X_t^*), into K univariate time series, $(\beta_{t1}), (\beta_{t2}) \ldots, (\beta_{tK})$, that are orthogonal to each other. The well known Karhunen-Loève theorem suggests to use the K eigenfunctions that correspond to the K highest eigenvalues of the covariance operator of the process (X_t^*) as basis functions. Estimation of the eigenfunctions works perfectly for iid or stationary data, but in the case of non stationary processes we face the problem that each element of (X_t^*) has got a different covariance operator. Nevertheless, we can show that each covariance operator, $\Gamma_{X_t^*}$ of X_t^* for all $t = \{1, 2, \ldots\}$, is an element of the same space (see theorem 2.1).

Theorem 2.1. *Without loss of generality, given a de-meand version* $(X_t^*) = (X_t - \mu_Z)$ *of the random walk process in equation* (2).

a) The covariance operators, $\Gamma_{X_t^*} = EX_t^* \otimes X_t^* \quad$ *for* $t = \{1, 2, \ldots\}$, *are elements of the same space.*
As a consequence, the eigenfunctions of the covariance operators are the same for all $t = \{1, 2, \ldots\}$.

b) The covariance operators, $\Gamma_{X_t^*} = EX_t^* \otimes X_t^* \quad$ *for* $t = \{1, 2, \ldots\}$, *are asymptotically identical, apart from scale differences.*

This characteristics motivated us to transformation the process by the unit-sphere projection, that is usually used in multivariate robust statistics of iid samples (see e.g. Huber and Ronchetti (2009), Locantore et al. (1999) and Gervini (2008)). We propose to decompose the series, (X_t^*), into a functional component, $(\tilde{X}_t^*) \in \mathcal{H}$, and an univariate component, $(\Theta_t) \in \mathbb{R}$, such that $(X_t^*) = (\tilde{X}_t^* \ \Theta_t)$. We call the fist component spherical component and the latter scaling component.

Definition The spherical component of a functional random walk as in equation (2) is given by $\pi X_t^* = \pi\delta + \pi X_{t-1}^* + \pi e_t$. Where $\pi = (.)/||.||_2$, is the unit-sphere projection operator, with $||.||_2 = \sqrt{\int_{\mathcal{U}}(.)^2}$, is the L_2 norm in $\mathcal{H}$.

Definition The scaling component of a functional random walk as in equation (2) is given by $||X_t^*||_2 = ||\delta||_2 + ||X_{t-1}^*||_2 + ||e_t||_2$. With $||.||_2 = \sqrt{\int_{\mathcal{U}}(.)^2}$, the L_2 norm in $\mathcal{H}$.

It can be shown that the spherical component, compactly written as $\tilde{X}_t^* = \tilde{\delta} + \tilde{X}_{t-1}^* + \tilde{e}_t$, is stationary and that the covariance operators of each element in $(\tilde{X}_t^*)$ has got the same eigenfunctions as its non stationary counterpart in (X_t^*) (see theorem 2.2). Note that the scaling component, compactly written as $\Theta_t = \alpha + \Theta_{t-1} + \epsilon_t$, is a standard univariate random walk with drift $\alpha \in \mathbb{R}$ and white noise process $(\epsilon_t) \sim N(0, \sigma_\epsilon)$.

Theorem 2.2. *Without loss of generality, given a de-meand version* $(X_t^*) = (X_t - \mu_Z)$ *of the random walk process in equation* (2).

a) Its spherical component, $(\tilde{X}_t^*) = (\pi X_t^*)$*, is a stationary process.*

b) The covariance operators, $\Gamma_{\tilde{X}_t^*} = E\tilde{X}_t^* \otimes \tilde{X}_t^* \quad$ *for* $t = \{1, 2, \dots\}$*, are elements of the same space as the non spherical counterparts,* $\Gamma_{X_t^*} = EX_t^* \otimes X_t^*$.
As a consequence, the eigenfunctions of the covariance operators, $\Gamma_{\tilde{X}_t^*}$*, are the same as of the covariance operators,* $\Gamma_{X_t^*}$*, for all* $t = \{1, 2, \dots\}$.

As a consequence of theorem 2.2, asymptotically, the covariance operators, $\Gamma_{X_t^*}$, of the non stationary original process, (X_t^*), are the same as the covariance operators, $\Gamma_{\tilde{X}_t^*}$, of the spherical process, $(\tilde{X}_t^*)$ except for scale differences. Therefore, we can estimate the original covariance operators from the stationary spherical series, $(\tilde{X}_t)$. And rescale the estimated covariance operator by the scaling component, (Θ_t), that has absorbed the scale differences. The K eigenfunctions that belong to the first K eigenvalues, $\lambda_1, \dots, \lambda_K$, of the spherical covariance operator will fulfill the optimality criterion in (3).

3. Estimation of the mean, covariance function, and standard deviation

We estimate the mean function by local linear polynomial smoothing as proposed by Yao et al. (2005). The measurement errors are balanced when all prices are pooled and therefore the estimation of the mean function

$$\hat{\mu}(u) = Sm[u, (u_{ti}, Y_{ti}), T, N_t, h_\mu]$$

stays satisfactory, where $S[v, (u_{ti}, Y_{ti}), T, N_t, h_\mu]$ denotes the result of the local polynomial smoothing procedure of the pooled data $(u_{ti}, Y_t(u_{ti}))$, for $i = 1, \dots, N_t$, and $t = 1, \dots, T$, evaluated at $v \in \mathbb{R}$ with smoothing parameter h_μ. All smoothing parameters are determined such that they are the minimizing the generalized cross validation criterion (Silverman (1984)). Explicit formulas of the estimators of the mean function and the covariance function are given in the Appendix A. In figure 3 the estimated mean function, $\hat{\mu}$, is plotted along with all pooled data points, (u_{ti}, Y_{ti}) for $t = 1, \dots, T$ and $i = 1, \dots, N_t$, that are classified to the marginal cost regime ($Y_{ti} \leq 145$ EUR). Furthermore, all prices with

adjusted demand values smaller than 34,000 (MW) are omitted because they are not dense enough to guarantee subsequent stable local polynomial smoothing.

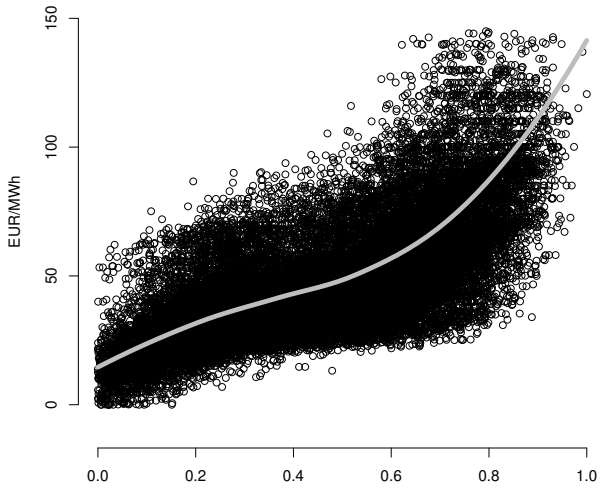

Figure 3: Estimated mean function, $\hat{\mu}$, and all pooled prices minus outliers.

The estimation of the covariance function uses the above explained spherical component, $(\tilde{X}_t^*)$, of the original series, (X_t^*). The spherical estimator of the covariance function is given in equation 4, where $Sm\left[u, v, (t, \tilde{G}_t), T, h_{\tilde{\rho}}, \text{linear}\right]$ denotes the result of the local linear polynomial surface smoothing procedure of the pooled data $(t, \tilde{G}_t)$, $t = 1, \ldots, T$, evaluated at $(u, v) \in \mathbb{R}^2$ with smoothing parameter $h_{\tilde{\rho}}$ and $||.||_E$ denotes the standard euclidean norm,

$$\tilde{\rho}_n(u,v) = Sm\left[u, v, (t, \tilde{G}_t), T, h_{\tilde{\rho}}, \text{linear}\right], \tag{4}$$

$$\text{with: } \tilde{G}_t = \left[\frac{(Y_t(u_{ti}) - \hat{\mu}(u_{ti}))\,(Y_t(u_{tj}) - \hat{\mu}(u_{tj}))}{||Y_t(u_{ti}) - \hat{\mu}(u_{ti})||_E ||Y_t(u_{tj}) - \hat{\mu}(u_{tj})||_E}\right]_{\substack{i,j=1,\ldots,N_t \\ \text{for all } i \neq j}}. \tag{5}$$

We use the subscript n to denote the estimator of $\tilde{\rho}$ in order to avoid messy superscripts. One should exercise caution in estimation of the covariance function $\tilde{\rho}$. As equation 1 indicates, we have to take the noise term into account otherwise the estimator of the diagonal, $\tilde{\rho}(v = u, u) = \tilde{\rho}(u)$, would be biased. A straight forward solution, originally proposed by Staniswalis and Lee (1998), is

to leave out the diagonal elements, as done in equation 5.

The variance of the curves $\tilde{X}_t^*(u)$ is usually reflected as a prominent feature along the diagonal of the covariance function $\tilde{\rho}$ and may be under estimated by the above explained estimation procedure. Following Yao et al. (2003) we rotate the coordinates, (u_{ti}, u_{tj}), of each element of $\tilde{G}_t$, clockwise by 45°,

$$\begin{pmatrix} u_{ti}^* \\ u_{tj}^* \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} u_{ti} \\ u_{tj} \end{pmatrix},$$

and estimate the surface again with a local quadratic polynomial smoother,

$$\dot{\rho}_n(u^*, v^*) = Sm\left[u^*, v^*, (t, \dot{G}_t), T, h_{\tilde{\rho}}, \text{quadratic}\right] \quad (6)$$

$$\text{with: } \dot{G}_t = \left[\frac{\left(Y_t(u_{ti}^*) - \hat{\mu}(u_{ti}^*)\right)\left(Y_t(u_{tj}^*) - \hat{\mu}(u_{tj}^*)\right)}{||Y_t(u_{ti}^*) - \hat{\mu}(u_{ti}^*)||_E ||Y_t(u_{tj}^*) - \hat{\mu}(u_{tj}^*)||_E}\right]_{\substack{i,j=1,\ldots,N_t \\ \text{for all } i \neq j}}. \quad (7)$$

The quadratic fit orthogonal to the diagonal of the covariance function approximates the variance of the functions better. The diagonal of the estimated covariance function, $\tilde{\rho}_n(u, v = u) = \tilde{\rho}_n(u)$ is set equal to $\dot{\rho}_n(0, u/\sqrt{2})$ for all u with $\tilde{\rho}_n(u) < \dot{\rho}_n(0, u/\sqrt{2})$; we denote this adjusted estimation of the covariance function classically by $\hat{\rho}$. In the left panel of figure 4 the estimated covariance function, $\hat{\rho}$, is shown; the sharpened diagonal is hardly visible but existent. The right panel of figure 4 gives a comparison of the three estimated diagonals of the covariance function, $\tilde{\rho}_n(u)$, $\dot{\rho}_n(u)$, and $\tilde{D}_n(u)$, where the latter uses only the noisy diagonal elements (see equation (8)).

$$\tilde{D}_n(u) = Sm\left[u, (t, \tilde{G}_t^D), T, h_D, \text{linear}\right] \quad (8)$$

$$\text{with: } \tilde{G}_t^D = \left[\frac{\left(Y_t(u_{ti}) - \hat{\mu}(u_{ti})\right)^2}{||Y_t(u_{ti}) - \hat{\mu}(u_{ti})||_E^2}\right]_{i=1,\ldots,N_t}.$$

The dashed line refers to the estimator $\tilde{\rho}_n(u)$, from equation (4), the dotted line refers to the estimator $\dot{\rho}_n(u)$, from equation (6) with rotated coordinates and quadratic smoothing, and the solid line is refers to the estimator, $\tilde{D}_n(u)$, from smoothing only the diagonal components. It is clearly visible that the inclusion of the diagonal elements leads to a strong distortion of the covariance diagonal, because of the noise term.

The difference between the diagonal estimation with diagonal components, $\tilde{D}$, and without diagonal components but rotated coordinates, $\dot{\rho}$, can be used to estimate the variance of the noise term, σ.

3.1. Spectral decomposition of the covariance operator

Given an estimation of the spherical covariance function, $\hat{\rho}$, we can derive the set of eigenvalues, $\{\hat{\lambda}\}_{k=1,\ldots K}$, and the set of eigenfunctions, $\{\hat{\phi}_k\}_{k=1,\ldots K}$, by the

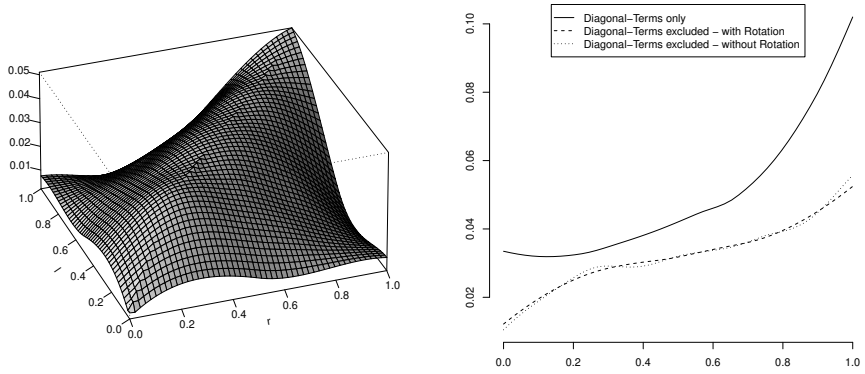

Figure 4: Left Panel: Estimate covariance function, $\hat{\rho}$. Right Panel: The results of the three versions of diagonal estimation of the covariance function.

spectral decomposition of the covariance operator, $\hat{\Gamma} f(u) = \int_{\mathcal{U}} \hat{\rho}(u,v) f(v) dv$, with $f(u) \in L^2(\mathcal{U})$, from the solutions of the eigenequations

$$\int_{\mathcal{U}} \hat{\rho}(u,v) \hat{\phi}_k(u) du = \hat{\lambda}_k \hat{\phi}_k(v),$$

with the usual restrictions $\int_{\mathcal{U}} \hat{\phi}_k^2 = 1$ and $\int_{\mathcal{U}} \hat{\phi}_k \hat{\phi}_l = 0$ for all $k < l$. The standard procedure is to discretize the covariance function, $\hat{\rho}$, at an equidistant grid $(u_1^d, \ldots, u_n^d) \times (u_1^d, \ldots, u_n^d) \in \mathcal{U} \times \mathcal{U}$ and then to use routines from the multivariate spectral decomposition of matrices (see e.g. Ramsay and Silverman (2005) for a detailed explanation).

From theorems 2.1 and 2.2 above, we know that the eigenfunctions of the spherical covariance operators, $\Gamma_{\tilde{X}_t^*}$, are the same as the eigenfunctions of all original covariance operators, $\Gamma_{X_t^*}$, such that these fulfill the best basis property of equation (3). Note that, they are not unique in this characteristic. There may be an other set of eigenfunctions that is as well efficient in the mean squared error sense but is better to interpret. Again, methods from the multivariate statistics can be used here to produce new eigenfunctions. Given a discretized set of estimated eigenfunctions, $[\phi_1^d, \ldots, \phi_K^d]$ with $\hat{\phi}_k^d = (\hat{\phi}_k(u_1^d), \ldots, \hat{\phi}_k(u_n^d))' \in \mathbb{R}^n$, every $K \times K$ rotation matrix R, with $R'R = RR' = I$, leads to a new orthonormal set of basis vectors $[\psi_1^d, \ldots, \psi_K^d] = R[\phi_1^d, \ldots, \phi_K^d]$.

Often, eigenfunctions are only interpretable after a suitable rotation scheme. The well known VARIMAX rotation tries to maximize the variance of each discretized eigenfunction $\hat{\phi}^d = (\hat{\phi}(u_1^d), \ldots, \hat{\phi}(u_n^d))'$ by either scaling the values $\hat{\phi}(u_i^d)$ against zero or against very high absolute values. Figure 5 shows the four rotated eigenfunctions that belong to the four highest eigenvalues. The

different types of power plants become apparent surprisingly well. The greatest part (70.88%) of the total variations of the spot prices, which belong to the marginal cost regime, is between 50,200 MW and 62,000 MW. This region is generally supplied by coal power plants, which face the most price volatile resource commodities (hard coal, brow coal, and CO2 certificates). The second greatest part (22.16%), between 62,000 MW and 73,720 MW, generally can be assigned to gas and oil power plants that are often used in hours with peak demands for electricity. The third part (5.49%) of the total variance, between 34,520 MW and 50,200 MW, can be assigned to the base power plants, mostly nuclear power plants.

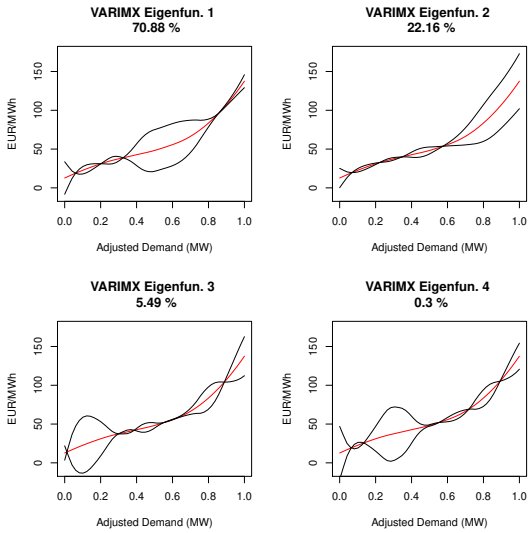

Figure 5: VARIMAX rotated eigenfuncitons.

3.2. Conditional estimation of the principal component scores

From the theorem 2.2 we know that the spherical covariance operator, $\Gamma_{\tilde{X}_t^*}$, is asymptotically the same as the covariance operator, $\Gamma_{X_t^*}$, except for scale differences. This allows us to model, first, the spherical sample curves, $(\tilde{X}_t^*)$, as a K dimensional process (see equation (3)),

$$\tilde{X}_t^{*,K} = \sum_{k=1}^{K} \tilde{\xi}_{tk}\phi_k,$$

and then to rescale them to their original size, (X_t^*), by their scaling component, (θ_t). For simplicity, we do not distinguish notationally between sample and generic versions of (X_t). The usual estimation of the pc-scores approximates the integral, $\tilde{\xi}_{tk} = \int_{\mathcal{U}} \phi_k \tilde{X}_t^*$. Given the non uniformly distributed data,

$u_{t1}, \ldots, u_{tN_t}$, over the domain, $\mathcal{U}$, we cannot adequately approximate the pc-scores by numerical integration procedures. The PACE approach of Yao et al. (2005) uses the conditional expectation, $\mathrm{E}(\beta_{tk}|X_t)$, given a joint normal distribution of the random vector $(\beta_{tk}, X_t)'$. This procedure can be applied to our problem, when we use the assumption that the spherical scores, $\tilde{\xi}_{tk}$, and the discrete spherical curve values, $\tilde{Y}_t^* = \frac{Y_t - \mu_t}{||Y_t - \mu_t||_E}$ with $Y_t = (Y_{t,1}, \ldots, Y_{tN_t})'$ and $\mu_t = (\mu(u_{t1}), \ldots, \mu(u_{tN_t}))$, come from a joint Gaussian distribution of $(\tilde{\beta}_{tk}, \tilde{X}_t)$. We estimate the conditional principal component scores, $\tilde{\xi}^c_{tk}$, given the non uniformly distributed discrete observed curve data, $\tilde{Y}_t^*$, by

$$\tilde{\xi}^c_{tk} \quad = \quad E[\tilde{\beta}_{tk}|\tilde{Y}_t^*] = \tilde{\lambda}_k \phi'_{tk} \tilde{\mathbf{\Sigma}}^{-1}_{\tilde{Y}_t^*}(\tilde{Y}_t^*), \tag{9}$$

where $\tilde{\mathbf{\Sigma}}_{Y_t} = [\tilde{\rho}(u_{ti}, u_{tj})]_{i,j=1,\ldots,N_t} + \tilde{\sigma}^2 I_{N_t}$ is a $N_t \times N_t$ symmetric matrix and $\phi_{tk} = (\phi(u_{t1}), \ldots, \phi(u_{tN_t}))'$ a N_t dimensional vector.

3.3. Fitted Curves

Different from traditional methods our estimation procedure does not focus on estimation of the hourly spot prices directly, but on the estimation of daily mean price curves (or supply curves, respectively). The left panel of figure 6 shows the estimated price curve of Thursday the 9^{th} February in 2006. The circle points are assigned to the marginal cost regime and contribute to the estimation procedure. The two prices corresponding to the two triangle points are assigned to the opportunity regime and do not contribute to the estimation procedure. The right panel of figure 6 shows the whole week from Monday the 6^{th} to Sunday the 12^{th} February in 2006. Here, the prices are plotted in the traditional in correspondence to their trading periods.

We want to emphasize that our separation of the data into a marginal cost regime and an opportunity cost regime is more fundamental than standard regime switch models that usually switch between two or more (often comparable) time series models. (See Jong (2006) for an overview of classical regime switch models in the context of electricity spot market data.) Here, we base the regime switch on a change in the bidding behavior of the traders. Ex post this is easily done with a hard threshold price; see discussion to figure 2. Ex ante this is not a trivial thing to do, since every trader may be forced to switch into the opportunity cost regime on basis of private information such as delivery obligations and unexpected changes in power plant capacities. On basis of public available data it will be hardly possible to predict individual regime switches, but it might be possible to predict situations in which (nearly) all market participants will have to switch into the opportunity regime. This will be part of future research.

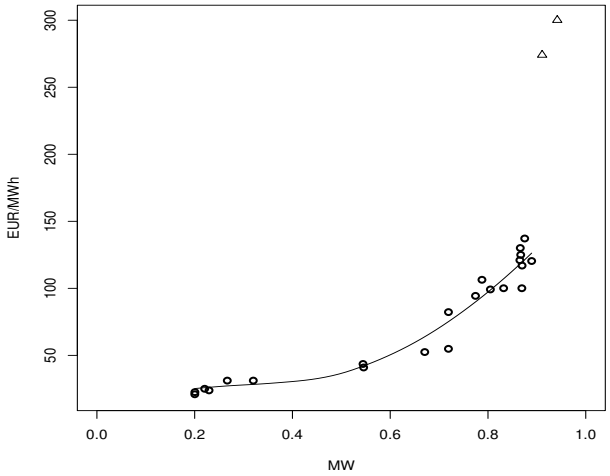

Figure 6: Estimated price curve, $X_t(u)$, of Th. 9/Feb./06; with superimposed 24 prices, $Y_t = \{Y_{t1}, \ldots, Y_{t24}\}$, of that day. The two prices corresponding to the two triangle points are assigned to the opportunity regime.

4. Conclusion

In this paper we support a new angle of vision in modeling hourly electricity spot market data. We argue that the intra-day seasonality cannot be estimated by traditional time series models, that are based on the assumption of a hourly updating information set (as already done by Huisman et al. (2007)). This assumption is often not valid because most electricity exchanges use a singe price auction where the hourly price vectors, $\{Y_{th}\}_{h=1,\ldots,24}$ for day t, are determined simultaneously the day before at $t-1$. We use a functional time series model and estimate daily mean supply functions by functional principal component analysis. Here, the intra-daily raw data is not the traditional consecutive price vector, $\{Y_{th}\}_{h=1,\ldots,24}$, but the re-ordered price vector, $\{Y_{ti}\}_{i=1,\ldots,24}$, corresponding to the covariate vector of adjusted demand values, $u_{t1} < \cdots < u_{ti} < \cdots < u_{t24}$.

This introduces two problems, first, the data looses its equidistant design, actually the adjusted demand values, u_{ti}, even are not uniformly distributed within the domain $\mathcal{U} \subset \mathbb{R}$. Second, (functional) principal component analysis needs iid data or at least stationary data, but our data set is non stationary. The first problem is solved by an adaption of the principal component analysis for spaces data (see Staniswalis and Lee (1998) and Yao et al. (2005)). The second problem is solved by the introduction of a new decomposition of the functional times series into a stationary spherical component and a non stationary scaling component. The latter is one of our main contributions that might be very useful for many other functional times series estimation problems. Furthermore,

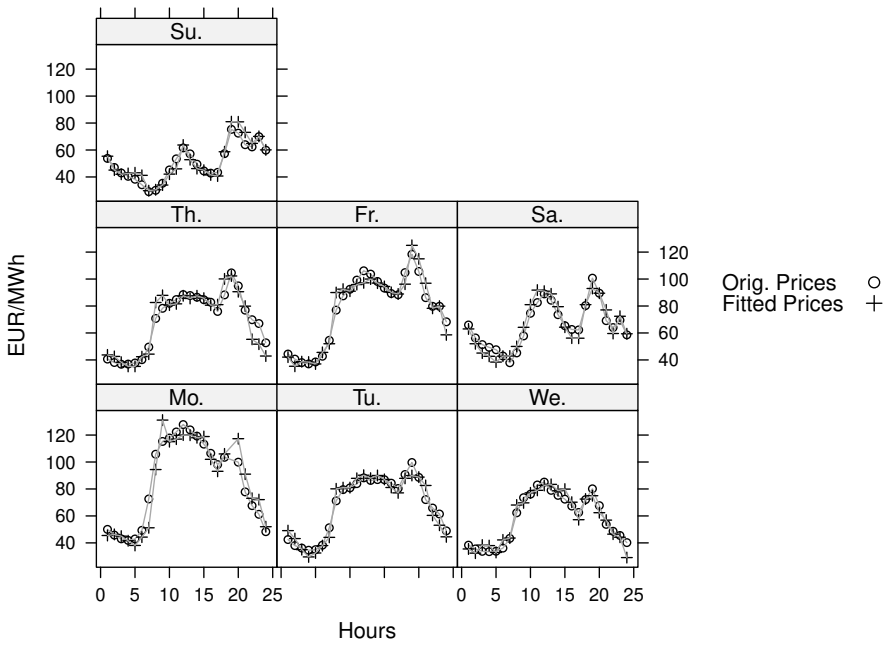

Figure 7: Original and fitted prices of the week: Mo. 6/Feb./06 – Su. 12/Feb./06. Here, the prices are re-ordered into the traditional hourly perspective.

our approach handles the information set correctly as a daily updating information set. The result that we need only three components in order to explain 98.53% of the total variation supports the importance to account for the correct consideration of the data generating process. Other studies that work with similar approaches but use the traditional hourly price vectors, $\{Y_{th}\}_{h=1,\ldots,24}$, need higher numbers of principal components for comparable fractions of explained variance (see e.g. Wolak (1997) and Härdle and Trück (2010)).

Appendix A. Explicit Formulas

Mean function $\mu(u)$:

$$\begin{aligned}
\hat{\mu}(u) &= Sm[u,(u_{ti},Y_{ti}),T,N_t,h_\mu] \\
\hat{\mu}(u) &= \hat{\beta}_0(u) \\
\text{With } \hat{\beta}_0(u) \text{ from: } \min_{\beta_0,\beta_1} & \sum_t^T \sum_i^{N_t} \mathcal{K}_1(\frac{u-u_{ti}}{h_\mu})[Y_{ti}-\beta_0-\beta_1(u-u_{ti})]^2
\end{aligned}$$

$$\begin{aligned}
\tilde{\rho}_n(u,v) &= Sm\left[u,v,(t,\tilde{G}_t),T,h_{\tilde{\rho}},\text{linear}\right] \\
\text{with: } \tilde{G}_t &= \left[\frac{(Y_t(u_{ti})-\hat{\mu}(u_{ti}))\,(Y_t(u_{tj})-\hat{\mu}(u_{tj}))}{||Y_t(u_{ti})-\hat{\mu}(u_{ti})||_E||Y_t(u_{tj})-\hat{\mu}(u_{tj})||_E}\right]_{\substack{i,j=1,\dots,N_t \\ \text{for all } i\neq j}} \\
\tilde{\rho}(u,v) &= \hat{\beta}_0(u,v) \\
\text{With } \hat{\beta}_0(u,v) \text{ from: } \min_{\beta_0,\beta_{11},\beta_{12}} & \sum_t^T \sum_{1\leq i\neq j\leq N_t} \mathcal{K}_2(\frac{u-u_{ti}}{h_G},\frac{v-u_{tj}}{h_G}) \\
& [G_t(u_{ti},u_{tj})-\beta_0-\beta_{11}(u-u_{ti})-\beta_{12}(v-u_{tj})]^2
\end{aligned}$$

Where $\mathcal{K}_1(w) = \begin{cases} (1-|w|^3)^3 & |w|<1 \\ 0 & \text{otherwise} \end{cases}$ (Or any other valid univar. kernelfunction.)

Where: $w = \frac{u-u_{ti}}{h_\mu}$

$\mathcal{K}_2(w,x) = \mathcal{K}_1(w)\mathcal{K}_1(x)$

Where: $w = \frac{u-u_{ti}}{h_G}$ and $x = \frac{v-u_{ti}}{h_G}$

The bandwidths are determined by Generalized Cross Validation (CGV). These routines are already implemented in the `R` package `locfit` (Loader, 2010).

Appendix B. Proofs

Proof of theorem 2.1.
Part *a)*:
From the definition of the covariance operator and the random walk.

$$\begin{aligned}
\Gamma_{X_t^*}(u,v) &= \mathrm{E}(X_t^* \otimes X_t^*(u,v)) = \mathrm{E}(X_t^*(u)\,X_t^*(v)) \\
\Gamma_{X_t^*}(u,v) &= \mathrm{E}((\sum_{i=0}^{t-1} e_{t-i}(u)+Z_0^*(u))(\sum_{i=0}^{t-1} e_{t-i}(v)+Z_0^*(v))).
\end{aligned}$$

With $\Gamma_e(u,v)=\sum_{k=1}^K \lambda_k^e\phi_k^e\otimes\phi_k^e(u,v)$ we can write $e_t(u)=\sum_{k=1}^\infty \beta_k^e\phi_k^e(u)$ for all $t=\{1,2,\dots\}$, where $\beta_k^e\sim N(0,\lambda_k^e)$. And similar for Z_0, with $\Gamma_Z(u,v)=$

$\sum_{k=1}^{K} \lambda_k^Z \phi_k^Z \otimes \phi_k^Z(u,v)$, we can write $Z_0(u) = \sum_{k=1}^{\infty} \beta_k^Z \phi_k^Z(u)$, where $\beta_k^Z \sim N(0, \lambda_k^Z)$. This yields,

$$\Gamma_{X_t^*}(u,v) \quad = \quad t\, \Gamma_e(u,v) + \Gamma_Z(u,v), \tag{B.1}$$

which corresponds to the usual univariate and multivariate random walk characteric of an with $O(t)$ increasing covariance. Given the spectral decompositions of the covariance operators, we have,

$$\Gamma_{X_t^*}(u,v) \quad = \quad t \sum_{k=1}^{\infty} \lambda_k^e \phi_k^e \otimes \phi_k^e(u,v) + \sum_{k=1}^{\infty} \lambda_k^Z \phi_k^Z \otimes \phi_k^Z(u,v).$$

Note that $\Gamma_{X_t^*}(u,v)$ is an element of an addition of two vector spaces,

$$\begin{aligned} \mathcal{P}_t \otimes \mathcal{P}_t + \mathcal{Q}_t \otimes \mathcal{Q}_t \quad &= \quad \{p | p = \sum_{k=1}^{\infty} \alpha_{p,k}\, t\, \phi_k^e \otimes \phi_k^e, \sum_{k=1}^{\infty} (\alpha_{p,k})^2 < \infty, \alpha_{p,k} \in \mathbb{R} \forall k\} \\ &+ \quad \{q | q = \sum_{k=1}^{\infty} \alpha_{q,k}\, \phi_k^e \otimes \phi_k^e, \sum_{k=1}^{\infty} (\alpha_{q,k})^2 < \infty, \alpha_{q,k} \in \mathbb{R} \forall k\}. \end{aligned}$$

Without loss of generality we can investigate the degenerated case where $\lambda_k^Z = 0$ for all k. Then we have

$$\Gamma_{X_t^*}(u,v) \quad = \quad t \sum_{k=1}^{\infty} \lambda_k^e \phi_k^e \otimes \phi_k^e(u,v) \;\in\; \mathcal{P}_t \otimes \mathcal{P}_t.$$

With the property that vector spaces are closed with respect to scalar multiplication we can directly show that $\mathcal{P}_t \otimes \mathcal{P}_t = \mathcal{P}_{t+i} \otimes \mathcal{P}_{t+i}$,

$$\begin{aligned} \mathcal{P}_t \otimes \mathcal{P}_t \quad &= \quad \text{span}\{t\, \phi_1^e \otimes \phi_1^e, t\, \phi_2^e \otimes \phi_2^e, \dots\} \\ &= \quad \{p | p = \sum_{k=1}^{\infty} \alpha_k\, t\, \phi_k^e \otimes \phi_k^e \text{ and } \alpha_k \in \mathbb{R} \forall k\} \\ &= \quad \{p | p = \sum_{k=1}^{\infty} \gamma_k\, (t+i)\, \phi_k^e \otimes \phi_k^e \text{ and } \gamma_k \in \mathbb{R} \forall k\} \text{ where } \gamma_t = (\alpha_k\, t)/(t+i) \\ &= \quad \text{span}\{(t+i)\, \phi_1^e \otimes \phi_1^e, (t+i)\, \phi_2^e \otimes \phi_2^e, \dots\} \;=\; \mathcal{P}_{t+i} \otimes \mathcal{P}_{t+i} \;=\; \mathcal{P} \otimes \mathcal{P}, \end{aligned}$$

for arbitrary $t, i = \{1, 2, \dots\}$. Therefore each covariance operator, $\Gamma_{X_t^*}(u,v)$, is an element of the same space

$$\mathcal{P} \otimes \mathcal{P} \quad = \quad \{p | p = \sum_{k=1}^{K} \alpha_{p,k}\, \phi_k^e \otimes \phi_k^e, \sum_{k=1}^{K} (\alpha_{p,k})^2 < \infty, \alpha_{p,k} \in \mathbb{R} \forall k\}.$$

This shows part *a)* of theorem 2.1.

Part *b)*:
From equation (B.1) we have,

$$\lim_{\substack{t,\, s \to \infty \\ (t/s) = \text{const.}}} \left(\frac{\Gamma_{X_t^*}}{\Gamma_{X_{t+s}^*}} \right) = \lim_{\substack{t,\, s \to \infty \\ (t/s) = \text{const.}}} \left(\frac{t\,(\,\Gamma_e + (1/t)\,\Gamma_Z)}{(t+s)\,(\,\Gamma_e + (1/(t+s))\,\Gamma_Z)} \right) = \frac{1}{(1 + const.)}.$$

This shows part *b)* of theorem 2.1. □

Proof of theorem 2.2.
Part *a)*:
Without loss of generality, we investigate the de-meand process, $(\tilde{X}_t^*)$, given by

$$\begin{aligned}
\pi(X_t - \mu_Z) &= \delta + \pi(X_{t-1} - \mu_Z) + \pi e_t = \sum_{i=0}^{t-1} \pi e_t + \pi(Z_0 - \mu_Z) \\
\tilde{X}_t^* &= \delta + \tilde{X}_{t-1}^* + \tilde{e}_t \qquad \text{with } X_t^* = (X_t - \mu_Z) \text{ and } \tilde{(.)} = \pi(.) = \frac{(.)}{||(.)||_2} \\
&= \delta + \sum_{i=0}^{t-1} \tilde{e}_t + \tilde{Z}_0^* \quad \text{with } Z_0^* = (Z_0 - \mu_Z) \text{ and } \tilde{Z}_0^* = \pi Z_0^*
\end{aligned}$$

We proof that the spherical component, $(\tilde{X}_t)$, is a (weak) stationary process. I.e. (i) has got constant mean function for all $t = \{0, 1, 2, \dots\}$, (ii) finite covariance operator, and (iii) autocovariance operators that are independent of t (see any introductory time series book, such as Shumway and Stoffer (2006)). Given the functional random walk process defined by equation (2). Without loss of generality, we set δ equal to the null function such that,

$$\tilde{X}_t^* = \tilde{X}_{t-1}^* + \tilde{e}_t$$

Condition (i):

The mean function is given by,

$$\begin{aligned}
\tilde{X}_t^* &= \tilde{X}_{t-1}^* + \tilde{e}_t \;=\; \sum_{i=0}^{t-1} \tilde{e}_t + \tilde{Z}_0^* \\
\mathrm{E}\frac{X_t^*}{||X_t^*||_2} &= \mathrm{E}\left(\frac{\sum_{i=0}^{t-1} \tilde{e}_t + Z_0^*}{||\sum_{i=0}^{t-1} \tilde{e}_t + Z_0^*||_2}\right) \;=\; \frac{0}{\mathrm{E}\sqrt{C_1}}. \\
\text{With } C_1 &= \int_{\mathcal{U}} (\sum_{i=0}^{t-1}\sum_{k=1}^{\infty} \beta_{t,k}^e \phi_k^e + \sum_{k=1}^{\infty} \beta_{t,k}^Z \phi_k^Z)^2
\end{aligned}$$

Taylor series expansion at $\mathrm{E}(C_1)$:

$$\begin{aligned}
\mathrm{E}\sqrt{C_1} &= \sqrt{\mathrm{E}(C_1)} + r = \sqrt{t\,\Gamma_e + \Gamma_Z} + r \\
\text{Where } 0 \geq r &= \frac{1}{2}(\mathrm{E}(C_1))^{-\frac{3}{2}}(\mathrm{E}(C_1) - C_1) - \frac{3}{8}(\mathrm{E}(C_1))^{-\frac{5}{2}}(\mathrm{E}(C_1) - C_1)^2 - O(C_1^3) \\
\text{sucht that} \quad & 0 < \mathrm{E}\sqrt{C_1} < \infty \\
\text{(We exclude the degenerated case:)} \quad & \Gamma_e = \Gamma_Z = 0. \\
\text{Finally:} \quad & \mathrm{E}(\pi X_t) = \frac{0}{\mathrm{E}\sqrt{C_1}} = 0, \qquad \forall t = \{0, 1, 2, \dots\}
\end{aligned}$$

And therefore suffices the condition (i) of a (weak) stationary process.

Condition (ii):
The covariance operator is given by (similar to proof of theorem 2.1),

$$\begin{aligned}
\Gamma_{\tilde{X}_t^*}(u,v) &= \mathrm{E}\tilde{X}_t^* \otimes \tilde{X}_t^*(u,v) \;=\; \mathrm{E}\left(\frac{X_t^*}{||X_t^*||_2} \otimes \frac{X_t^*}{||X_t^*||_2}(u,v)\right) \\
\Gamma_{\tilde{X}_t^*}(u,v) &= \mathrm{E}\left(\frac{(\sum_{i=0}^{t-1} e_{t-i}(u) + Z_0(u))(\sum_{i=0}^{t-1} e_{t-i}(v) + Z_0(v))}{||\sum_{i=0}^{t-1} e_{t-i} + Z_0||_2^2}\right) \\
&= \frac{t\,\sum_{k=1}^{\infty} \lambda_k^e \phi_k^e \otimes \phi_k^e(u,v) + \sum_{k=1}^{\infty} \lambda_k^Z \phi_k^Z \otimes \phi_k^Z(u,v)}{t\,\sum_{k=1}^{\infty} \lambda_k^e + \sum_{k=1}^{\infty} \lambda_k^Z}
\end{aligned}$$

Asymptotically,

$$\lim_{t \to \infty}\left(\Gamma_{\tilde{X}_t^*}(u,v)\right) = \sum_{k=1}^{\infty}\left(\frac{\lambda_k^e}{\sum_{k=1}^{\infty}\lambda_k^e}\right)\phi_k^e \otimes \phi_k^e(u,v)$$

By definition of Γ_e : $\quad \sum_{k=1}^{\infty}(\lambda_k^e)^2 < \infty$

And therefore also: $\quad \sum_{k=1}^{\infty}\frac{\lambda_k^e}{\sum_{k=1}^{\infty}\lambda_k^e} < \infty$

And therefore suffices the condition (ii) of a (weak) stationary process. Furthermore, this shows part *b)* of theorem 2.2.

Condition (iii):

$$\begin{aligned}
\mathrm{E}\tilde{X}_t^* \otimes \tilde{X}_{t+h}^*(u,v) &= \mathrm{E}\left(\frac{X_t^*}{||X_t^*||_2} \otimes \frac{X_{t+h}^*}{||X_{t+h}^*||_2}(u,v)\right) \\
&= \mathrm{E}\left(\frac{(\sum_{i=0}^{t-1} e_{t-i}(u) + Z_0(u))(\sum_{i=0}^{t+h-1} e_{t+h-i}(v) + Z_0(v))}{||\sum_{i=0}^{t-1} e_{t-i} + Z_0||_2 \, ||\sum_{i=0}^{t+h-1} e_{t+h-i} + Z_0||_2}\right)
\end{aligned}$$

$$= \frac{\mathrm{E}(\sum_{i=0}^{t-1}\sum_{k=1}^{\infty}\beta_{t-i,k}^e\phi_k^e(u) + \sum_{k=1}^{\infty}\beta_k^Z\phi_k^Z(u))(\sum_{i=0}^{t+h-1}\sum_{k=1}^{\infty}\beta_{t+h-i,k}^e\phi_k^e(v) + \sum_{k=1}^{\infty}\beta_k^Z\phi_k^Z(v))}{\mathrm{E}||(\sum_{i=0}^{t-1}\sum_{k=1}^{\infty}\beta_{t-i,k}^e\phi_k^e(u) + \sum_{k=1}^{\infty}\beta_k^Z\phi_k^Z(u))||_2 \, ||(\sum_{i=0}^{t+h-1}\sum_{k=1}^{\infty}\beta_{t+h-i,k}^e\phi_k^e(v) + \sum_{k=1}^{\infty}\beta_k^Z\phi_k^Z(v))||_2}$$

$$\begin{aligned}
\text{Def. } \alpha(u) &= \sum_{i=0}^{t-1}\sum_{k=1}^{\infty}\beta_{t-i,k}^e\phi_k^e(u) + \sum_{k=1}^{\infty}\beta_k^Z\phi_k^Z(u) \\
\beta(u) &= \sum_{i=0}^{h-1}\sum_{k=1}^{\infty}\beta_{t+h-i,k}^e\phi_k^e(u)
\end{aligned}$$

$$= \frac{t\ \sum_{k=1}^{\infty}\lambda_k^e\phi_k^e \otimes \phi_k^e(u,v) + \sum_{k=1}^{\infty}\lambda_k^Z\phi_k^Z \otimes \phi_k^Z(u,v)}{\mathrm{E}\sqrt{(\int_{\mathcal{U}}\alpha^2)\ \int_{\mathcal{U}}(\beta^2 + 2\,\beta\,\alpha + \alpha^2)}} = \frac{t\,\Gamma_e + \Gamma_Z}{\mathrm{E}\sqrt{(\int_{\mathcal{U}}\alpha^2)\,(\int_{\mathcal{U}}\beta^2 + 2\ \int_{\mathcal{U}}\beta\,\alpha + \int_{\mathcal{U}}\alpha^2)}}$$

$$
\begin{aligned}
\mathrm{E}\tilde{X}_t^* \otimes \tilde{X}_{t+h}^*(u,v) &= \frac{t\,\Gamma_e + \Gamma_Z}{\mathrm{E}\sqrt{C_2}} \\
\text{With } \sqrt{C_2} &= \sqrt{(\int_{\mathcal{U}} \alpha^2)\,(\int_{\mathcal{U}} \beta^2 + 2\int_{\mathcal{U}} \beta\,\alpha + \int_{\mathcal{U}} \alpha^2)}
\end{aligned}
$$

Taylor series expansion at $\mathrm{E}(C_2)$:

$$
\begin{aligned}
\mathrm{E}\sqrt{C_2} &= \sqrt{\mathrm{E}((\int_{\mathcal{U}} \alpha^2)(\int_{\mathcal{U}} \beta^2)) + \mathrm{E}((\int_{\mathcal{U}} \alpha^2)^2)} + r \\
\text{Where } 0 \geq r &= \frac{1}{2}(\mathrm{E}(C_2))^{-\frac{3}{2}}(\mathrm{E}(C_2) - C_2) - \frac{3}{8}(\mathrm{E}(C_2))^{-\frac{5}{2}}(\mathrm{E}(C_2) - C_2)^2 - O(C_2^3) \\
\mathrm{E}\sqrt{C_2} &= \sqrt{\mathrm{E}((\int_{\mathcal{U}} \alpha^2)^{-1}(\int_{\mathcal{U}} \alpha^2)^2(\int_{\mathcal{U}} \beta^2)) + \mathrm{E}((\int_{\mathcal{U}} \alpha^2)^2)} + r \\
\mathrm{E}\sqrt{C_2} &= \sqrt{\left(\mathrm{E}(\int_{\mathcal{U}} \alpha^2)^{-1}(\int_{\mathcal{U}} \beta^2) + 1\right)\mathrm{E}((\int_{\mathcal{U}} \alpha^2)^2)} + r \\
\mathrm{E}\sqrt{C_2} &= \sqrt{\left(\frac{h\ \int_{\mathcal{U}} \Gamma_e(u,u)}{t\ \int_{\mathcal{U}} \Gamma_e(u,u) + \int_{\mathcal{U}} \Gamma_Z(u,u)} + 1\right)\mathrm{E}((\int_{\mathcal{U}} \alpha^2)^2)} + r \\
\mathrm{E}\sqrt{C_2} &= \sqrt{\left(\frac{(h/t)\ \int_{\mathcal{U}} \Gamma_e(u,u))}{\int_{\mathcal{U}} \Gamma_e(u,u) + (1/t)\ \int_{\mathcal{U}} \Gamma_Z(u,u)} + 1\right)\mathrm{E}((\int_{\mathcal{U}} \alpha^2)^2)} + r
\end{aligned}
$$

With $h = \{0, 1, 2, \dots\}$ fix but arbitrary:

$$
\begin{aligned}
\mathrm{E}\sqrt{C_2} &= \sqrt{(O(t^{-1}) + 1)\ \mathrm{E}((\int_{\mathcal{U}} \alpha^2)^2)} + r
\end{aligned}
$$

By the displacement law

$$
\begin{aligned}
\mathrm{E}\sqrt{C_2} &= \sqrt{(O(t^{-1}) + 1)\ (\mathrm{E}(\int_{\mathcal{U}} \alpha^2))^2 + \mathrm{V}(\int_{\mathcal{U}} \alpha^2)} + r \\
\mathrm{E}\sqrt{C_2} &= \sqrt{(O(t^{-1}) + 1)\ (\mathrm{E}(\int_{\mathcal{U}} \alpha^2))^2 + O(t)} + r \\
\mathrm{E}\sqrt{C_2} &= \sqrt{(O(t^{-1}) + 1)\ (t\ \int_{\mathcal{U}} \Gamma_e(u,u) + \int_{\mathcal{U}} \Gamma_Z(u,u))^2 + O(t)} + r
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\tilde{X}_t^* \otimes \tilde{X}_{t+h}^*(u,v) &= \frac{t\,\Gamma_e + \Gamma_Z}{\mathrm{E}\sqrt{C_2}} \\
\mathrm{E}\tilde{X}_t^* \otimes \tilde{X}_{t+h}^*(u,v) &= \frac{\Gamma_e + (1/t)\,\Gamma_Z}{\sqrt{(O(t^{-1}) + 1)\ (\int_{\mathcal{U}} \Gamma_e(u,u) + (1/t)\ \int_{\mathcal{U}} \Gamma_Z(u,u))^2 + O(t^{-1})} + (1/t)\ r} \\
\mathrm{E}\tilde{X}_t^* \otimes \tilde{X}_{t+h}^*(u,v) &\xrightarrow[t\to\infty]{} \frac{\Gamma_e}{\int_{\mathcal{U}} \Gamma_e(u,u)}
\end{aligned}
$$

□

References

Bosq, D., 2000. Linear processes in function spaces: Theory and Applications. Springer.

Cramton, P., 2004. Competitive bidding behavior in uniform-price auction markets. Proceedings of the Hawaii International Conference on System Sciences.

Ferraty, F., Vieu, P., 2006. Nonparametric Functional Data Analysis, 1st Edition. Springer.

Gervini, D., Sep. 2008. Robust functional estimation using the median and spherical principal components. Biometrika 95 (3), 587–600.

Green, R. J., Newbery, D. M., 1992. Competition in the British Electricity Spot Market. The Journal of Political Economy 100 (5), 929 – 953.

Härdle, W. K., Trück, S., 2010. The dynamics of hourly electricity prices.

Huber, P. J., Ronchetti, E. M., 2009. Robust Statistics. John Wiley & Sons.

Huisman, R., Huurman, C., Mahieu, R., Mar. 2007. Hourly electricity prices in day-ahead markets. Energy Economics 29 (2), 240–248.

Jong, C. D., 2006. The nature of power spikes: A regime-switch approach. Studies in Nonlinear Dynamics & Econometrics 10 (3), Article 3.

Karakatsani, N., Bunn, D., 2008. Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. International Journal of Forecasting 24 (4), 764–785.

Klemperer, P. D., Meyer, M. A., 1989. Supply Function Equilibria in Oligopoly under Uncertainty. Econometrica 57 (6), 1243 – 1277.

Kosater, P., Mosler, K., 2006. Can Markov regime-switching models improve power-price forecasts? Evidence from German daily power prices. Applied Energy.

Loader, C., 2010. locfit: Local Regression, Likelihood and Density Estimation. URL `http://cran.r-project.org/package=locfit`

Locantore, N., Marron, J. S., Simpson, D. G., Tripoli, N., Zhang, J. T., Cohen, K. L., Boente, G., Fraiman, R., Brumback, B., Croux, C., 1999. Robust principal component analysis for functional data. Test 8 (1), 1–73.

Mount, T., Ning, Y., Cai, X., 2006. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. Energy Economics 28 (1), 62–80.

Park, B. U., Mammen, E., Härdle, W., Borak, S., Mar. 2009. Time Series Modelling With Semiparametric Factor Dynamics. Journal of the American Statistical Association 104 (485), 284–298.

Rafal Weron, 2006. Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach. Wiley.

Ramsay, J. O., Silverman, B. W., 2005. Functional data analysis, 2nd Edition. Vol. 8. Springer.

Shumway, R. H., Stoffer, D. S., 2006. Time Series Analysis and its Applications With R examples, 2nd Edition. Springer, New York.

Silverman, B. W., 1984. A Fast and Efficient Cross-Validation Method for Smoothing Parameter Choice in Spline Regression. Journal of the American Statistical Association 79 (387), 584 – 589.

Staniswalis, J. G., Lee, J. J., 1998. Nonparametric Regression Analysis of Longitudinal Data. Journal of the American Statistical Association 93 (444), 1403–1418.

Willems, B., Rumiantseva, I., Weigt, H., Jan. 2009. Cournot versus Supply Functions: What does the data tell us? Energy Economics 31 (1), 38–47.

Wolak, F., 1997. Market Design and Price Behavior in Restructured Electricity Markets: An International Comparison.

Yao, F., Müller, H. G., Clifford, A. J., Dueker, S. R., Follett, J., Lin, Y., Bruce, A. B., Vogel, S. J., 2003. Shrinkage Estimation for Functional Principal Component Scores with Application to the Population Kinetics of Plasma Folate. Biometrics 59 (3), 676 – 685.

Yao, F., Müller, H. G., Wang, J. L., Jun. 2005. Functional Data Analysis for Sparse Longitudinal Data. Journal of the American Statistical Association 100 (470), 577–590.