Contribute 4
Modelling Electricity Prices as Functional Data on Random Domains

Kneip Alois, Dominik Liebl

Abstract We model electricity spot prices as noisy discretization points of a latent functional time series. It is assumed that there are densely sampled discretization points, but only within a random sup-part of the original domain. Under mild regularity assumptions, it is possible to recover the functions on the whole original domain. We present in-probability convergence rates for the local linear estimators of the mean and covariance function as well as for the functional principal components. In our application, we show that the recovered functions are powerful functional predictors for hourly electricity spot prices.

4.1 Introduction

The merit order model assumes that the spot prices at electricity exchanges are based on the marginal generation costs of the last power plant that is required to cover the demand. The resulting (monotonically increasing) merit order curve reflects the increasing generation costs of a specific portfolio of power plants. Often, nuclear and lignite plants cover the minimal demand for electricity. Higher demand is mostly served by hard coal and gas fired power plants. A statistical treatment of these micro-economic insights using functional data analysis was already proposed in [4]. There it is suggested to estimate empirical merit order curves, or more generally, price functions from the $n = 24$ data pairs (Y_{it}, X_{it}) of hourly electricity spot prices Y_{it} and hourly electricity demand values X_{it} with $i \in \{1, \dots, n\}$ and $t \in \{1, \dots, T\}$. A scatter plot of the data is shown in Figure 4.1. There the price functions are pre-smoothed for the purpose of visualization, though, we work with the discretization points plotted as gray filled circles.

Kneip Alois
University of Bonn, Germany, e-mail: akneip@uni-bonn.de

Dominik Liebl
Université libre de Bruxelles, Belgium, e-mail: dominik.liebl@ulb.ac.be

Modeling hourly electricity spot prices using this functional data point of view turns out to be very useful and yields a superior forecast performance; see [4]. This paper can be seen as an extension of [4] inasmuch as we use the concept of price functions, but augment it by the additional covariable of daily mean air-temperatures. Furthermore, in contrast to [4] we provide a deeper consideration of the rarely studied phenomenon of functional data observed on random domains.

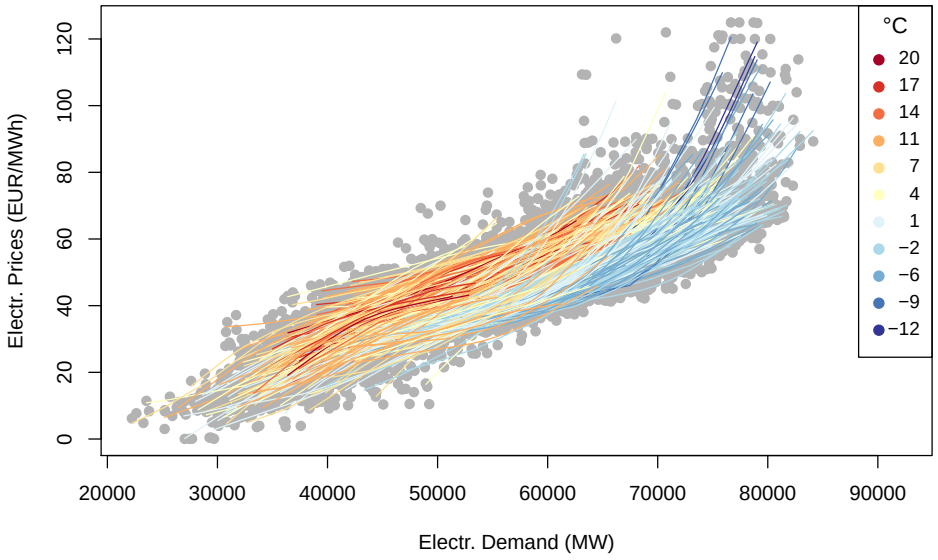

Figure 4.1: Scatter plot of electricity spot prices Y_{it} (circles points) and daily price functions (solid lines) of the whole data set from January 1, 2011 to August 31, 2013.

While the hourly values of electricity demand X_{it} are classical discretization points of the functional data, the daily temperature values Z_t determine the general shape of the functions. For example, price functions of cold days tend to be convex, while those of hot days can be concave; see Figure 4.1. Correspondingly, we regard hourly electricity spot prices Y_{it} as noisy discretization points of a conditional, stationary time series of square integrable random functions, which can be written as

$$\begin{aligned} Y_{it} &= \mu(X_{it}|Z_t) + f_t(X_{it}|Z_t) + \epsilon_{it} \quad \text{with} \qquad (4.1) \\ f_t(X_{it}|Z_t) &= \sum_{k=1}^{\infty} \beta_{tk}(Z_t)\psi_k(X_{it}|Z_t), \end{aligned}$$

where μ is a deterministic mean function, f_t is the tth element of the centered

functional time series, ϵ_{it} is an iid mean zero error with finite variance $V(\epsilon_{it}) = \sigma_\epsilon^2$, $\psi_k(\bullet|Z_t)$ are deterministic conditional eigenfunctions, and $\beta_{tk}(Z_t)$ are random conditional principal component (pc) scores with $E(\beta_{tk}(Z_t)) = 0$ and $Var(\beta_{tk}(Z_t)) = \lambda_k(Z_t)$ with $\lambda_k(Z_t)$ being the kth conditional eigenvalue of the covariance operator

$$(\Gamma g|z)(x) = \int_{[a(z),b(z)]} \gamma(x, y|z)g(y)dy \tag{4.2}$$

for all $g(\bullet) \in L^2[a(z), b(z)]$ with $x, y \in [a(z), b(z)] \subset \mathbb{R}$, and $z \in \mathbb{R}$. The nonparametric functions $a(z)$ and $b(z)$ are smooth boundary functions, which determine the set of physically feasible electricity demand values $D = \{(x, z) \in \mathbb{R}^2 | a(z) \leq x \leq b(z)\}$. A visualization of the set D is given as light-gray colored region in the left panel of Figure 4.2.

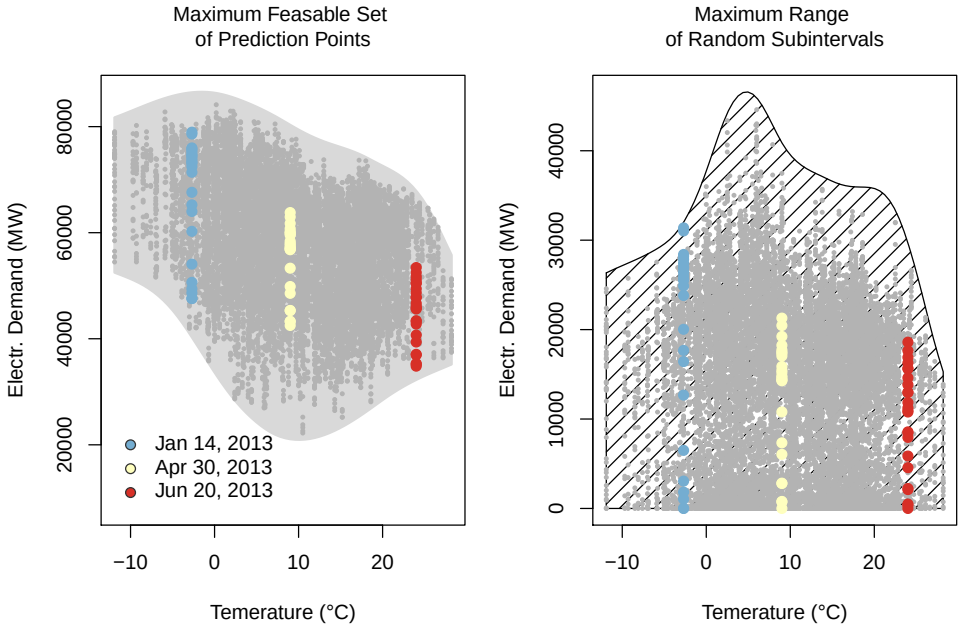

Figure 4.2: LEFT PANEL: Visualization of the set D of physically feasible electricity demand values (light-gray colored region). The prediction points (X_{it}, Z_t) of the whole data set from Jan 1, 2011 to Aug 31, 2013 are plotted as dark-gray filled circles. RIGHT PANEL: Visualization of the temperature dependent maximal range, which can be attained by the random intervals $[A(z_t), B(z_t)]$. BOTH PANELS: The colored prediction points visualize the ranges of the random intervals.

For a given day t with temperature $Z_t = z_t$, the observed values of electricity demand X_{it} are clustered within a random subinterval, say, $[A_t(z_t), B_t(z_t)] \subset$

$[a(z_t), b(z_t)]$. However, outside of the subinterval there is no observation at all; see Figure 4.3.

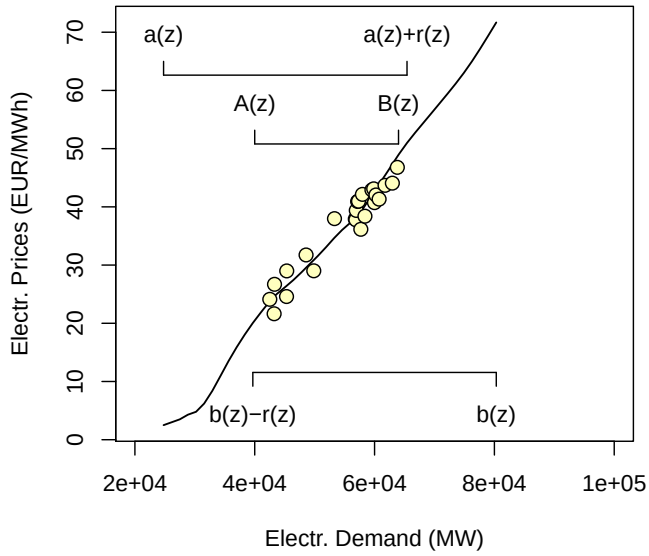

Figure 4.3: Fitted price function of the day April 30, 2013.

In order to estimate the model components in Eq. (4.1) we use functional principal component analysis (FPCA). In particular, we estimate the covariance function $\gamma(x, y|z)$ using a nonparametric local linear estimator applied the pooled raw covariance points $(Y_{it} - \mu(X_{it}, Z_t))(Y_{jt} - \mu(X_{jt}, Z_t))$. This approach was originally proposed by [5], in order to be able to use FPCA in the case of sparse functional data with only a few discretization points per function.

Though, our case differs form the case of sparse functional data as we allow the data to be densely sampled, but only within the random subinterval $[A(z_z), B(z_t)]$; see Figure 4.3. The most substantial difference, however, arises from the assumption that the random subintervals $[A(z_t), B(z_t)]$ are proper subsets of the interval $[a(z_t), b(z_t)]$. I.e., the ranges $R(z_t) = B(z_t) - A(z_t)$ are strictly smaller than the original ranges $b(z_t) - a(z_t)$. In fact, it is reasonable to assume that the random ranges $R(z_t)$ are bounded from above by the nonparametric deterministic bound $r(z_t)$. The latter is visualized by the line-shaded region in the right panel of Figure 4.2; there the points $X_{1t} - A(z_t), \ldots, X_{nt} - A(z_t)$ are bounded from above by $r(z_t)$.

The bound $r(z)$ has sever consequences as it makes the covariance operator in (4.2) infeasible to estimate by restricting the set of observable raw covariance points. The latter issue is visualized by the missing data points in the

upper-left and lower-right corners of the scatter plot in Figure 4.4. The raw covariance points of a single day are restricted to the relatively small plain $[A(z_t), B(z_t)]^2$ visualized by the yellow-filled circle points. Fortunately, pooling the raw-covariance points of all days that have approximately same temperature values $z \approx z_t$, allows us to estimate the covariance function $\gamma(x, y|z)$ over an enlarged plain.

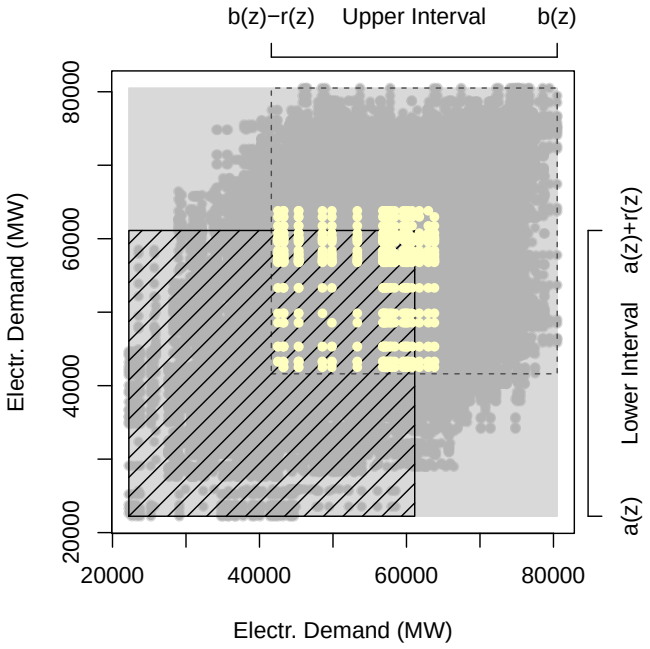

Figure 4.4: Scatter plot of prediction points for estimating the covariance function.

There are many different plains over which it is possible to define a feasible covariance operator. In particular, under some regularity assumptions, the two plains $[b(z_t) - r(z_t), b(z_t)]^2$ and $[a(z_t), a(z_t) + r(z_t)]^2$, as shown in Figure 4.4, allow us to consistently estimate the price functions f_t over the whole original domains $[a(z_t), b(z_t)]$; see Figure 4.3.

FPCA for sparse functional data is considered for the first time in [5]. Based on a "large-T, small-n" asymptotic, the authors derive upper bounds describing uniform-in-probability convergence of the empirical mean function, covariance function, eigenvalues and eigenfunctions. More precise uniform-in-probability convergence rates for the eigenvalues and eigenfunctions are derived in the follow-up paper of [2]. [3] reconsider the problem and derive uniform, almost sure convergence rates. In the latter two papers, the authors take a unifying point of view and derive convergence rates for all possible scenarios between "large-T, small-n" and "large-T, large-n" asymptotics.

We contribute to this literature by deriving the asymptotic bias and variance

of the local linear estimator for the mean and the covariance function. Hereby, we take into account the particular stochastics from the functional time series error term (f_t) of model (4.1), which imposes within-function covariances as well as and between-function covariances.

Under mild assumptions, it is possible to recover the functions on the whole original domain. In our application, we show that these recovered functions are powerful functional predictors in dynamic functional regression models, which have strong forecast performances.

Bibliography

[1] Fan, J., Gijbels, I. (1996). Local Polynomial Modelling and its Applications. Monographs on Statistics and Applied Probability **66**. Chapman & Hall/CRC.

[2] Hall, P., Müller, H. G. and Wang J. L. (2006). Properties of principal component methods for functional and longitudinal data analysis. *The Annals of Statistics*. **34** 1493–1517.

[3] Li, Y. and Hsing, T. (2010). Uniform convergence rates for nonparametric regression and principal component analysis in functional/longitudinal data. *The Annals of Statistics*. **38** 3321–3351.

[4] Liebl, D. (2013). Modeling and forecasting electricity spot prices: a functional data perspective. *The Annals of Applied Statistics*. **7** 1562–1592

[5] Yao, F., Müller, H. G. and Wang J. L. (2005). Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*. **100** 577–590.