Forecasting next-day electricity demand and price using nonparametric functional methods

Juan M.Vilar[*], Ricardo Cao[†] and Germán Aneiros[‡]
Departamento de Matemáticas
Facultad de Informática
Universidade da Coruña
Campus de Elviña, s/n
15071 A Coruña
Spain

Abstract

One-day-ahead forecasting of electricity demand and price is an important issue in competitive electric power markets. These problems have been studied in previous works using, for instance, ARIMA models, dynamic regression and neural networks. This paper provides two new methods to address these two prediction setups. They are based on using nonparametric regression techniques with functional explanatory data and a semi-functional partial lineal model. Results of these methods for the electricity market of mainland Spain, in years 2008-2009, are reported. The new forecasting functional methods are compared with a naïve method and with seasonal ARIMA forecasts.

Keywords: Demand and price, electricity markets, functional data, time series forecasting.

1 Introduction

Nowadays, in many countries all over the world, the production and sale of electricity is traded under competitive rules in free markets. The agents involved in this market: system operators,

[*]Corresponding author. E-mail: eijvilar@udc.es

[†]E-mail: rcao@udc.es

[‡]E-mail: ganeiros@udc.es

regulatory agencies, producers and consumers have a great interest in the study of electricity load and price. Since electricity cannot be stored, the demand must be satisfied instantaneously and producers need to anticipate to future demands to avoid overproduction. Good forecasting of electricity demand is then very important for the agents in the market. In the past, demand was predicted in centralized markets [1] but competition has opened a new field of study. On the other hand, if producers and consumers have reliable predictions of electricity price, they can develop their bidding strategies and establish a pool bidding technique to achieve a maximum benefit. Consequently, prediction of electricity demand and price are significant problems in this sector. In demand and price prediction two types of forecasts are considered: short term forecasts, this is one day ahead hourly forecasts, and mid term forecasts, i.e. several days ahead predictions for daily data.

In most of electricity markets the series of demand and prices present the following features. They have non-constant mean and variance. Superimposed levels of seasonality, in particular, daily and weekly seasonality, are present in these series. Price and demand series exhibit short-term dynamics and special days (calendar effect on weekend and holidays). These series incorporate nonlinear effects of meteorological variables, especially temperature. Outliers are also present. Usually, the price series is more volatile than the demand series. The price series also presents more outliers, mainly in periods of high demand. The previous features imply that finding an adequate fit for these series is a complex problem.

This paper focuses on next day forecasting of electricity demand and price. Therefore, for each day of the week, 24 forecasts (of demand or price) need to be computed. Plenty of algorithms have been developed for short term demand prediction. Two different statistical strategies have been used to this aim. The first one consists in building a single model to derive the load profile, see [2] and [3]. The second approach provides 24 different models to compute one day ahead hourly predictions, using, for each hour, a different time series, see [4] and [5]. In some papers the temperature is incorporated as an explanatory variable in the model [6]. Artificial neural networks have been also used in this context [7] and [8]. Usually short term time dependence is assumed to be generated by an ARMA process, see [9]. This idea is exploited by Cancelo *et al.* [9] to present a complex model used by *Red Eléctrica de España* (REE), the Spanish system operator.

Price forecasting techniques in power systems are relatively recent procedures. Most papers are focused on computing one-day-ahead forecasting. Dynamic regression and transfer functions [10] have been used in this context. ARIMA models for price forecasting in the Spanish market have been also studied, see [11] – [13]. Troncoso *et al.* [14] compared two weighted k-nearest neighbours methods and dynamic regression in this setup. Kim *et al.* [15] used wavelet transform analysis in the England & Wales pool. Jump diffusion/mean reversion models [16] and fuzzy regression [17] have been also applied to model electricity prices. Artificial neural networks have been used to forecast electricity prices too. Some relevant works in this field include [18] – [22]. Conejo *et al.* [23] compared several methods, including ARIMA models, neural networks and wavelet approximation, for price forecasting. In a recent paper [24] the main methodologies used in electricity price forecasting are reviewed.

This paper proposes functional nonparametric and semi-functional partial linear models to forecast electricity demand and price. Nonparametric regression estimation under dependence is a useful tool for forecasting time series. Some relevant works in this field include [25] – [27]. Other papers more specifically focused on prediction using nonparametric techniques are [28] – [30]. The literature on methods for time series prediction in the context of functional data is much more limited. The books [31] and [32] are comprehensive references for linear and nonparametric functional data analysis, respectively. A functional wavelet-kernel approach for time series prediction has been also proposed, see [33]. Aneiros and Vieu [34] dealt with the problem of nonparametric time series prediction using a semi-functional partial linear model. Nadaraya-Watson and local linear methods for functional data have been also used in the problem of time series prediction [35].

The approach in this paper uses functional data methods to take into account the daily seasonality of the electricity demand and price series. Nonparametric regression estimation methods are used to forecast these series. The idea is to cut the observed time series into a sample of functional trajectories and to incorporate in the model just a single past functional observation (last day trajectory) rather than multiple past time series values. A functional version of the partial linear model has been considered in order to incorporate vector covariates in the forecasting procedure. As a consequence, this approach allows both to consider additional explanatory covariates and to use continuous past paths to predict future values. This method is applied in short-term forecasting of electricity demand and price in the market of mainland Spain. This is an area of almost 500 000 square kilometres, with more than 45 million people.

The remaining of this paper is organized as follows. In Section 2, a mathematical description of the functional nonparametric model and the semi-functional partial linear model is given. The accuracy measures and the tunning parameters used in this paper are presented in Section 3. Section 4 shows numerical results concerning one-day ahead forecasting of electricity demand in mainland Spain. A comparative study of the proposed models and some classical methods is also included in this section. Section 5 presents a similar study concerning the problem of electricity price forecasting. Finally, Section 6 provides some relevant conclusions.

2 Functional models

Functional data analysis is a branch of statistics that analyzes data providing information about curves, surfaces or any other mathematical object varying over a continuum. The continuum is often time, but may also be spatial location, wavelength, etc. These curves are defined by some functional form. In particular, functional data analyses often make use of the information in the slopes and curvatures of curves, as reflected in their derivatives. Plots of first and second derivatives as functions of time, or plots of second derivative values as functions of first derivative values, may reveal important aspects of the processes generating the data. As a consequence, curve estimation methods designed to yield good derivative estimates can play a critical role in functional data analysis.

Models for functional data and methods for their analysis may resemble those for conventional multivariate data, including linear and nonlinear regression models, principal components analysis, and many others. But the possibility of using derivative information greatly extends the power of these methods, and also leads to purely functional models such as those defined by differential equations.

The books [36] and [37] are nice general references for functional data analysis using a linear view. Due to the nature of the data, nonparametric smoothing methods are also useful tools for functional data analysis. The book [32] is devoted to nonparametric methods in the functional data context. This is the approach followed in this paper. In the following subsections we will present the functional nonparametric model and the semi-functional partial linear model used for demand and price forecasting.

2.1 Functional nonparametric model

The interest time series (electricity demand or price) will be considered as discrete time realizations of a continuous time stochastic process, $\{\chi(t)\}_{t\in R}$, observed for $t \in [a, b)$. We first concentrate on predicting $\chi(b + r)$, for some $r \geq 0$. Let us assume that $\chi(t)$ is a seasonal process, with seasonal length τ and $b = a + (n + 1)\tau$. In other words, we assume that the interval $[a, b)$ consists of $n + 1$ seasonal periods of length τ of the stochastic process $\{\chi(t)\}_{t\in R}$. In our application the seasonal length, τ, will be one day and $n + 1$ will be the number of days in the data set.

For simplicity we will assume the following Markov property for the process $\{\chi(t)\}_{t\in R}$:

$$\chi(b + r)|_{\{\chi(t),t\in[a,b)\}} \overset{\mathrm{d}}{=} \chi(b + r)|_{\{\chi(t),t\in[b-\tau,b)\}}. \tag{1}$$

Equation (1) states that the probability distribution of a future value of the process given the whole past only depends on the process values at the last observed seasonal interval.

We define the functional data $\{\chi_i\}_{i=1}^{n+1}$ in terms of the continuous time stochastic process:

$$\chi_i(t) = \chi(a + (i - 1)\tau + t), \text{ with } t \in C = [0, \tau).$$

We may look at the problem of predicting the future value $\chi(b + r)$ (for some $r \in C$) by computing nonparametric estimations, $\widehat{m}(\chi)$, of the autoregression function in the functional nonparametric (FNP) model

$$\chi_{i+1}(r) = m\left(\chi_i\right) + \varepsilon_{i+1}, \quad i = 1, \ldots, n. \tag{2}$$

Equation (2) assumes that the values of the series at instant r in a seasonal interval is an unknown nonparametric function of the series at the whole previous seasonal interval plus some error term. Thus, $\widehat{m}\left(\chi_{n+1}\right)$ gives a functional forecast for $\chi\left(b+r\right)$.

In our context this approach consists on estimating the autoregression functional, m, using full-day demand or price values and apply this estimated functional to the last observed day.

Whereas the Euclidean norm is a standard distance measure in finite dimensional spaces, the notion of seminorm or semimetric arises in this infinite-dimensional functional setup. Let us denote by $\mathcal{H} = \{f : C \rightarrow R\}$ the space where the functional data live and by $d(\bullet, \bullet)$ a semimetric associated with $\mathcal{H}$. Thus $(\mathcal{H}, d)$ is a semimetric space (see the book [32] for details).

A Nadaraya-Watson type estimator for m in (2) is defined as

$$\widehat{m}_h^{FNP}\left(\chi\right) = \sum_{i=1}^{n} w_h(\chi, \chi_i)\chi_{i+1}(r) \tag{3}$$

where the bandwidth $h > 0$ is a smoothing parameter,

$$w_h(\chi, \chi_i) = \frac{K\left(d(\chi, \chi_i)/h\right)}{\sum_{j=1}^{n} K\left(d(\chi, \chi_j)/h\right)}$$

and the kernel function $K : [0, \infty) \rightarrow [0, \infty)$ is typically a probability density function chosen by the user.

The choice of the kernel function is of secondary importance. However, both the bandwidth and the semimetric are relevant aspects for the good asymptotic and practical behavior of (3).

A key role of the semimetric is that related to the so called "curse of dimensionality". From a practical point of view the "curse of dimensionality" can be explained as the sparseness of data in the observation region as the dimension of the data space grows. This problem is specially dramatic in the infinite-dimensional context of functional data. Ferraty and Vieu [32] have proven that it is possible to construct a semimetric in such a way that the rate of convergence of the nonparametric estimator in the functional setting is similar to that of the finite-dimensional one. It is important to remark that we use a semimetric rather than a metric. Indeed, the "curse of dimensionality" would appear if a metric were used instead of a semimetric.

In functional data it is usual to consider semimetrics based on seminorms. Thus, [32] recommend, for smooth functional data, to take as seminorm the L_2 norm of some q-th derivative of the function. For the case of rough data curves, these authors suggest to construct a seminorm based on the first q functional principal components of the data curves.

2.2 Semi-functional partial linear model

Very often there exist exogenous scalar variables that may be useful to improve the forecast. In many of these situations, previous experience also suggests that an additive linear effect of these variables on the values to forecast might occur. In such setups, it seems natural to generalize model (2) by incorporating a linear component. This gives the semi-functional partial linear (SFPL) model:

$$\chi_{i+1}(r) = \mathbf{x}_{i+1}^T \boldsymbol{\beta} + m\left(\chi_i\right) + \varepsilon_{i+1}, \quad i = 1, \ldots, n \tag{4}$$

where $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})^T \in R^p$ is a vector of exogenous scalar covariates and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T \in R^p$ is a vector of unknown parameters to be estimated.

Now, based on the SFLP model, we may look at the problem of predicting $\chi\left(b+r\right)$ (for some $r \in C$) by computing estimations $\widehat{\boldsymbol{\beta}}$ and $\widehat{m}\left(\chi\right)$ of $\boldsymbol{\beta}$ and $m\left(\chi\right)$ in (4), respectively. Thus, $\mathbf{x}_{n+2}^T\widehat{\boldsymbol{\beta}}+\widehat{m}\left(\chi_{n+1}\right)$ gives the forecast for $\chi\left(b+r\right)$.

An estimator for $\boldsymbol{\beta}$ based on kernel and ordinary least squares ideas was proposed in [38] in the setting of independent data. More specifically, recall the weights $w_h(\chi, \chi_i)$ defined in the previous subsection and denote $\widetilde{\mathbf{X}}_h = (\mathbf{I} - \mathbf{W}_h)\mathbf{X}$ and $\widetilde{\boldsymbol{\chi}}_h = (\mathbf{I} - \mathbf{W}_h)\boldsymbol{\chi}$, with $\mathbf{W}_h = (w_h(\chi_i, \chi_j))_{1\leq i,j\leq n}$, $\mathbf{X} = (x_{ij})_{\substack{1\leq i\leq n\\ 1\leq j\leq p}}$ and $\boldsymbol{\chi} = (\chi_2(r), \ldots, \chi_{n+1}(r))^T$, the estimator for $\boldsymbol{\beta}$ is defined by

$$\widehat{\boldsymbol{\beta}}_h = (\widetilde{\mathbf{X}}_h^T\widetilde{\mathbf{X}}_h)^{-1}\widetilde{\mathbf{X}}_h^T\widetilde{\boldsymbol{\chi}}_h. \tag{5}$$

It should be noted that $\widehat{\boldsymbol{\beta}}_h$ is the ordinary least squares estimator obtained when the vector of response variables, $\widetilde{\boldsymbol{\chi}}_h$, is linearly linked with the matrix of covariates $\widetilde{\mathbf{X}}_h$. It is worth mentioning that kernel estimation is used to obtain both $\widetilde{\boldsymbol{\chi}}_h$ and $\widetilde{\mathbf{X}}_h$. Actually, both terms are computed as some nonparametric residuals.

Finally, nonparametric estimation is used again to construct the estimator for $m\left(\chi\right)$ in (4)

$$\widehat{m}_h^{SFPL}\left(\chi\right) = \sum_{i=1}^{n} w_h(\chi, \chi_i)\left(\chi_{i+1}(r) - \mathbf{x}_{i+1}^T\widehat{\boldsymbol{\beta}}_h\right). \tag{6}$$

Rates of convergence of the estimators (5) and (6) in a setting of independent data were obtained in [38]. Recently, [34] gave conditions under which those rates hold for time series. It is interesting to note that $\widehat{\boldsymbol{\beta}}_h$ is a $\sqrt{n}$-consistent estimator of $\boldsymbol{\beta}$. It is also worth mentioning that the rate of convergence of $\widehat{m}_h^{SFPL}\left(\chi\right)$ is the same as that of $\widehat{m}_h^{FNP}\left(\chi\right)$.

Other estimators for m in (2) (and therefore for $\boldsymbol{\beta}$ and $m\left(\chi\right)$ in (4)) could be obtained by means of wavelet-kernel approaches [33] or local linear functional procedures [35].

3 Prediction accuracy and tuning parameters

Our goal is forecasting next-day both electricity demand and electricity prices using several functional models and nonparametric techniques. We will obtain one prediction for each hour in the day. To present a complete study, we will predict the seven days in each of four selected weeks corresponding to the four seasons. Thus, the day and season effects should be small in our results.

The accuracy of each model and forecast considered will be measured using daily and weekly errors. The daily errors (DE) are defined by

$$DE_j = \frac{1}{24}\sum_{r=1}^{24} e_{j,r}, \quad j = 1, \ldots, 7$$

where $e_{j,r}$ denotes the hourly relative percentage error

$$e_{j,r} = 100 \times \frac{\left|V_j(r) - \widehat{V}_j(r)\right|}{V_j(r)}$$

and $V_j(r)$ and $\widehat{V}_j(r)$ are the actual and the forecasted value (demand or price) at the r-th hour in the j-th day of the week to be predicted. The weekly error (WE) is then

$$WE = \frac{1}{7}\sum_{j=1}^{7} DE_j.$$

In practice, several "parameters" need to be selected for the functional nonparametric estimate and the semi-functional partial linear estimate. This is done in the following way:

1. The kernel K has low impact on the estimates. We use the Epanechnikov kernel defined by $K(u) = \frac{3}{4}(1 - u^2)I_{(0,1)}$.

2. For the bandwidth, h, we consider the k-nearest-neighbours method. The number of neighbours, k, is selected by means of the local cross-validation method. See [39] and [40] for details.

3. The semimetric, d, chosen is based on a seminorm. Since our data (curves) are rough (see figures below), we consider a seminorm based on functional principal component analysis. Cross-validation ideas are used to select the tuning parameter, q, the number of principal components.

4 Case study: electricity demand

We are interested in one-day electricity demand forecasting (specifically, j-hours ahead electricity demand forecasting, for $j = 1, \ldots, 24$). We focus on the seven days in each of four weeks corresponding to the four seasons: Summer week 2008 (August 17 – August 23), Autumn week 2008 (November 16 – November 22), Winter week 2009 (February 22 – February 28) and Spring week 2009 (May 17 – May 23). Historical data used for forecasting include the hourly demand (MW) of the 49 previous days to the day which demands are to be predicted. Note that $n = 48$ in the functional models (2) and (4). These data correspond to Spain, and are available at http://www.omel.es, the official website of Operador del Mercado Ibérico de Energía.

Fig. 1 displays both the historical demand and the demand to be predicted, corresponding to the Autumn week mentioned above. Similar pictures for the other three weeks are omitted. The presence of trend in the data is clearly seen in Fig. 1. Thus, the data were differentiated.

[FIGURE 1 AROUND HERE]

We denote by $\chi_i : [0, 24) \to R$ the differentiated electricity demand along the day i ($i = 1, \ldots, n$). Fig. 2 shows the historical curves used to forecast the demand along the first day in the Autumn week. The roughness of the curves exhibited in Fig. 2 remains for the other three weeks.

[FIGURE 2 AROUND HERE]

In this case study a three-dimensional covariate, $\mathbf{x}_i = (x_{i1}, x_{i2}, x_{i3})^T$, was used for the SFPL model (4). The covariates in this vector are $x_{i1} = I_{\text{Saturday}}$, $x_{i2} = I_{\text{Sunday}}$ and $x_{i3} = I_{\text{Monday}}$, i.e. the indicators of a Saturday, a Sunday and a Monday. The reason for this choice is the fact that Saturdays and Sundays are the days of the week with smaller demand. Since the autoregressive degree of the model is 1, these dummy covariates are useful to characterize four different scenarios:

1. Saturdays: relatively small demand with respect to the previous day.
2. Sundays: where the demand is relatively similar to that of the previous day and both tend to be lower than for weekdays.
3. Monday: which is a day with a high demand, when compared to the previous day.
4. Tuesday-Friday: where the high demand of this day is similar to the demand of the previous one.

Of course, different types of covariates could also be incorporated, as, for instance, temperature.

When the prediction horizon is larger than one, point forecasts are carried out in two different ways. The first one is the direct (D) method and consists in the approach mentioned in the previous section. The second alternative is the recursive (R) method. It computes a one-ahead forecast and includes it in the sample to perform again a one-lag prediction, as many times as needed.

We will use the notation FNP-D and FNP-R for the forecasts based on the FNP model (2) obtained from the direct and the recursive methods, respectivaly. SFPL-D and SFPL-R will denote the forecasts based on the SFPL model (4) obtained from the direct and the recursive methods, respectively.

Forecasts obtained by means of the functional models (2) and (4) will be compared with those based on seasonal ARIMA models. More specifically, ARIMA$(p,1,q)\times(P,1,Q)_{24}$ models, with $p,q \in \{1,\ldots,12\}$ and $P,Q \in \{1,\ldots,7\}$, are considered. Note that these values for the orders of the ARIMA model allow for hourly, daily and weekly dependence. In practice the values for p and q ranged between 6 and 12 while P and Q took the values 6 and 7 most of the times.

In addition, a naïve procedure is also used for comparison. It consists in forecasting the demand for a given week by means of the demand in the previous week.

Table 1 gives both the DE and WE corresponding to each model. In general, the results in this table show the good behavior of the SFPL-R forecasts. The recursive versions of the new functional procedures give slightly better results than the direct ones.

[TABLE 1 AROUND HERE]

We now compare hourly errors for the naïve, ARIMA and SFPL-R forecasts only. Fig. 3 shows a visual comparison of the hourly errors of only these three forecasts. The good behaviour of SFPL-R is also evident from this plot. It is worth mentioning the relatively bad results of ARIMA hourly forecasts in the Spring week. The forecasting error is even larger for the naïve method in the Winter week.

[FIGURE 3 AROUND HERE]

Fig. 4 displays the SFPL-R forecasts for the four selected weeks. The good behaviour of SFLP-R is clearly seen from these plots.

[FIGURE 4 AROUND HERE]

5 Case study: electricity price

A similar study is conducted in this section for electricity price forecasting. Prices were available for the same period as demands, and they were obtained from the same source. The four weeks selected to evaluate the performance of the methods are also the same as in the previous section.

Fig. 5 shows the historical prices and the prices to be predicted corresponding to the Autumn week. Similar pictures for the other three weeks are omitted. The presence of outliers in prices is clearly seen from this figure. This is a usual feature for price data [23].

[FIGURE 5 AROUND HERE]

To reduce the problem of outlier effect, pruned curves has also been considered before applying the models and forecasts for the electricity price data. More specifically, the historical prices were pruned in the following way. First, a deterministic trend (with time as covariate) was fitted to the historical prices corresponding to the same hour. A Nadaraya-Watson estimator with cross-validation bandwidth was used to compute this trend. Second, the residuals from that fit were computed and truncated to plus/minus twice their estimated standard deviation. Finally, transformed historical prices were constructed by adding the new residuals to the estimated trend. The pruned functional data were obtained from these transformed historical prices. It should be noted that the percentage of pruned prices was around 5%. Fig. 6 displays the historical pruned curves, used to forecast the price, along the first day in the Autumn week. The roughness of the pruned curves remains for the other three weeks analyzed and, obviously, for the original curves.

[FIGURE 6 AROUND HERE]

Tables 2 and 3 give both the DE and the WE corresponding to each analyzed model. We use the notation FNP-D-P, FNP-R-P, SFPL-D-P and SFPL-R-P for the forecasts obtained from pruned curves, either using the FNP or SFPL functional autoregressive models, with the direct or the recursive method. In general, working with pruned data is better than using the original prices without correcting the outliers. It is also clear that the SFPL model is a very competitive one, although not superior to ARIMA for price forecasting. In this case, the order of these ARIMA models were around 2 and 3 for p and q and 6 and 7 for P and Q. Unlike for demand forecasting, the recursive method and the direct one give comparable results.

[TABLES 2 and 3 AROUND HERE]

We now compare hourly errors for the naïve, ARIMA and SFPL-R-P forecasts only. Fig. 7 shows a comparison of the hourly errors of these three forecasts. The good performance of ARIMA forecasts and the competitive behaviour of SFPL-R-P are clearly seen in this plot. It is worth mentioning the relative stability of the SFPL-R-P forecasting error along hours and days for the four selected weeks.

[FIGURE 7 AROUND HERE]

Fig. 8 displays the SFPL-R-P forecasts for the four selected weeks. The good behaviour of SFLP-R-P is clearly seen from this figure.

[FIGURE 8 AROUND HERE]

6 Conclusions and perspectives

Functional data nonparametric techniques have been successfully used for electricity demand and price forecast in the Spanish market. The semi-functional partial linear model proposed in this paper gives good error results when compared to other complex already existing approaches, like the seasonal ARIMA models. It is worth mentioning that for $0 \leq P, Q \leq 7$ seasonal ARIMA models allow for dependence up to the previous seven days. However, the functional and semi-functional models used in this paper only include dependence on the previous day. The performance of the SFPL model is better for demand than for price forecasting.

Despite its complexity, the SFPL model has been used in a rather simple form. For instance, the outlier pruning has been carried out using a straightforward technique. We focused on point forecasts but prediction bands can also be constructed using this functional data approach. The model is very flexible to incorporate the effect of new informative covariates that can enter it either in parametric (e.g. linear) or nonparametric form. It can be easily applied to high frequency data (e.g. 10-minute data) whenever they become available. All these features make this approach appealing and with plenty of potential for improving.

Acknowledgments

The authors are partly supported by Grants PGIDIT07PXIB105259PR and 07SIN012105PR from Xunta de Galicia (Spain), and by Grant number MTM2008-00166 from Ministerio de Ciencia e Innovación (Spain).

References

[1] G. Gross and F.D. Galiana, "Short-term load forecasting", *Proc. IEEE*, vol. 75, pp. 1558-1573, Dec. 1987.

[2] M. Smith, "Modelling and short-term forecasting of New South Wales electricity system load", *Journal of Business and Economic Statistics*, vol. 22, pp. 17-28, 2000.

[3] J.W. Taylor, L.M. de Menezes and P.E McSharry, "A Comparison of univariate methods for forecasting electricity demand up to a day ahead", *International Journal of Forecasting*, vol. 22, pp. 1-16, 2006.

[4] R. Ramanathan, R. Engle, C.W.J. Granger, F. Vahid-Araghi and C. Brace, "Short-run forecasts of electricity loads and peaks",*International Journal of Forecasting*, vol. 13, pp. 161-174, 1997.

[5] R. Cottet and M. Smith, "Bayesian modeling and forecasting of intraday electricity load", *Journal of the American Statistical Association*, vol. 98, pp. 839-849, 2003.

[6] A. Pardo, V. Meneu and E. Valor, "Temperature and seasonality influences on Spanish electricity load", *Energy Economics*, vol. 24, pp. 55-70, 2002.

[7] H.S. Hippert, D.W. Bunn, and R.C. Souza, "Large neural networks for electricity load forecasting: are they overfitted?", *International Journal of Forecasting*, vol. 21, pp. 425-434, 2005.

[8] A.P. Alves da Silva, V.H. Ferreira and R. Velasquez, "Input space to neural network based load forecasting", *International Journal of Forecasting*, vol. 24, pp. 616-629, 2008.

[9] J.R. Cancelo, A. Espasa and R. Grafe, "Forecasting the electricity load from one day to one week ahead for the Spanish system operator"',*International Journal of Forecasting*, vol. 24, pp. 588-602, 2008.

[10] F.J. Nogales, J. Contreras and A.J. Conejo, "Forecasting next-day electricity prices by time series models", *IEEE Transactions on Power Systems*, vol. 17, no. 2, pp. 342-348, 2002.

[11] J. Contreras, R. Espínola, F.J. Nogales, and A.J. Conejo, "ARIMA models to predict next-day electricity prices", *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1014-1020, 2003.

[12] A. Conejo, M. Plazas, R. Espínola and B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models", *IEEE Trans Power Syst*, vol. 20, pp. 1035-1042, 2005.

[13] C. García-Martos, J. Rodríguez and M.J. Sánchez, "Mixed models for short-run forecasting of electricity prices: application for the Spanish market", *IEEE Transactions on Power Systems*, vol. 2, no. 2, pp. 544-552, 2007.

[14] A. Troncoso, J. Riquelme, A. Gómez and J.L. Martínez, "A comparison of two techiques for next-day electricity price forecasting", *Lecture Notes In Computer Science*, 2453, Proceedings of the 13th International Conference on Database and Expert Systems Applications, 2002.

[15] C.I. Kim, I.K. Yu, and Y.H. Song, "Prediction of system marginal price of electricity using wavelet transform analysis", *Energy Conversion and Management*, vol. 43, pp. 1839-1851, 2002.

[16] P. Skantze, M. Ilic, and J. Chapman, "Stochastic modeling of electric power prices in a multi-market environment", in *Proc. Power Engineering Winter Meet.*, Singapore, Jan. 2000, pp. 1109-1114.

[17] T. Nakashima, M. Dhalival, and T. Niimura, "Electricity market data representation by fuzzy regresion models", presented at the *Power Eng. Summer Meet.,Seattle, WA,* July 2000.

[18] B. Ramsay and A.J. Wang, "A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays", *Neurocomputing*, vol. 23, pp. 47-57, 1998.

[19] B.R. Szkuta, L.A. Sanabria, and T.S. Dillon, "Electricity price short-term forecasting using artificial neural networks", *IEEE Trans. Power Syst.*, vol. 14, pp. 851-857, 1999.

[20] F. Gao, X. Guan, X.R. Cao, and A. Papalexopoulos, "Forecasting power market clearing price and quantity using a neural network method", in *Proc. Power Engineering Summer Meet.*, Seattle, WA, pp. 2183-2188, July 2000.

[21] Y.Y. Hong and C.Y. Hisiao, "Locational marginal price forecasting in deregulated electricity markets using artificial intelligence", *IEEE Proceedings-Generation, Transmission and Distribution*, vol. 149, no. 5, pp. 621-626, 2002.

[22] L. Zhang, P.B. Luh and K. Kasiviswanathan, "Energy cleary price prediction and confidence interval estimation with cascaded neural networks", *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 99-105, 2003.

[23] A.J. Conejo, J. Contreras, R. Espínola and M.A. Plazas, "Forecasting electricity prices for a day-ahead pool-based electric energy market", *International Journal of Forecasting*, vol. 21, no. 3, pp. 435-462, 2005.

[24] S.K. Aggarwal, L.M. Saini and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation", *Electrical Power and Energy Systems*, vol. 31, pp. 13-22, 2009.

[25] W. Härdle and P. Vieu, "Kernel regression smoothing of time series", *Journal of the Time Series Analysis*, vol. 13, pp. 209-232, 1992.

[26] J. D. Hart, "Some automated methods of smoothing time-dependent data", *Journal of Nonparametric Statistics*, vol. 6, pp. 115-142, 1996.

[27] W. Härdle, H. Lütkepohl and R. Chen, "A review of nonparametric time series analysis", *International Statistical Review,* vol. 65, pp. 49-72, 1997.

[28] M. Carbon and M. Delecroix, "Nonparametric vs parametric forecasting in time series: a computational point of view". *Applied Stochastic Models and Data Analysis*, vol. 9, pp. 215-229, 1993.

[29] E. Matzner-Lober, A. Gannoun and J.G. De Gooijer, "Nonparametric forecasting: a comparison of three kernel based methods", *Communications in Statistics: Theory and Methods*, vol. 27, pp. 1593-1617, 1998.

[30] J.M. Vilar-Fernández and R. Cao, "Nonparametric forecasting in time series - A comparative study", *Communications in Statistics: Simulation and Computation*, vol. 36, pp. 311-334, 2007.

[31] D. Bosq, *Linear Processes in Function Spaces: Theory and Applications.* Lecture Notes in Statistics, 149, Springer- Verlag, 2000.

[32] F. Ferraty and P. Vieu, *Nonparametric Functional Data Analysis.* Springer Series in Statistics, New York: Springer-Verlag, 2006.

[33] A. Antoniadis, E. Paparoditis and T. Sapatinas, "A functional wavelet-kernel approach for time series prediction",*Journal of the Royal Statistical Society, Series B*, vol. 68, pp. 837-857, 2006.

[34] G. Aneiros-Pérez and P. Vieu, "Nonparametric time series prediction: A semi-functional partial linear modeling", *Journal of Multivariate Analysis*, vol. 99, pp. 834-857, 2008.

[35] G. Aneiros-Pérez, R. Cao and J.M. Vilar-Fernández, "Functional methods for time series prediction: a nonparametric approach", *Journal of Forecasting*, to be published.

[36] J.O. Ramsay and B.W. Silverman, *Applied functional data analysis: methods and case studies*, Springer series in statistics, New York; London: Springer, 2002.

[37] J.O. Ramsay and B.W. Silverman, *Functional data analysis.* New York: Springer-Verlag, 2005.

[38] G. Aneiros-Pérez and P. Vieu, "Semi-functional partiall linear regression", *Statistics and Probability Letters*, vol. 76, pp. 1102-1110, 2006.

[39] K. Benhenni, F. Ferraty, M. Rachdi, P. Vieu, "Local smoothing regression with functional data", *Computational Statistics,* vol. 22, pp. 353-369, 2007.

[40] G. Aneiros-Pérez, P. Vieu, "Automatic estimation procedure in partial linear model with functional data", *Statistical Papers*, to be published.

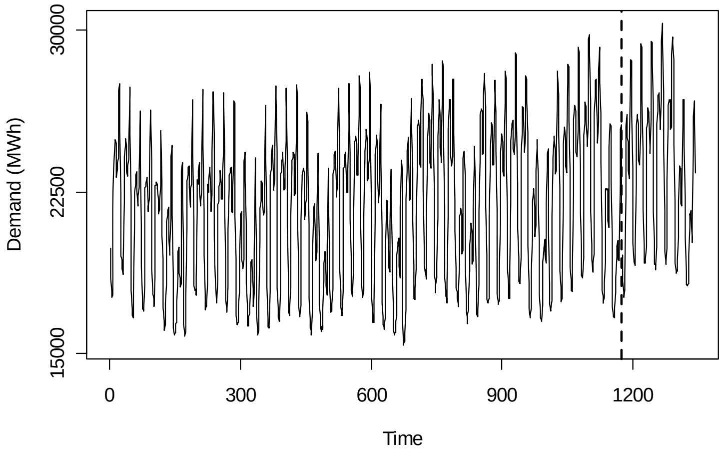

Figure 1: Historical data (left of the vertical line) and data to be predicted (right of the vertical line) corresponding to the electricity demand for the Autumn week.

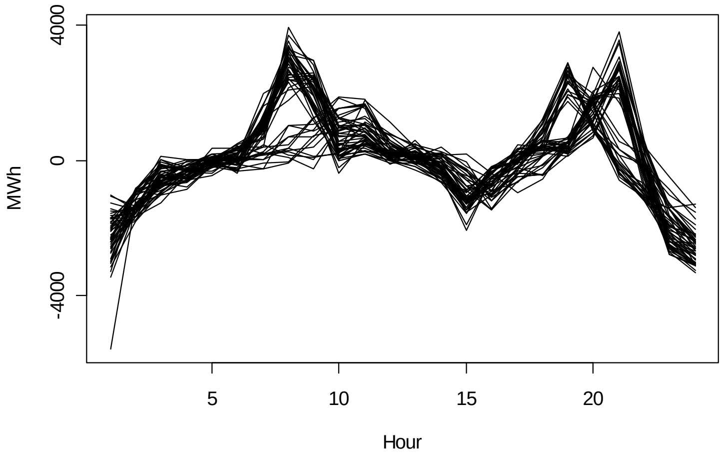

Figure 2: Functional data (daily curves) for the differenciated electricity demand.

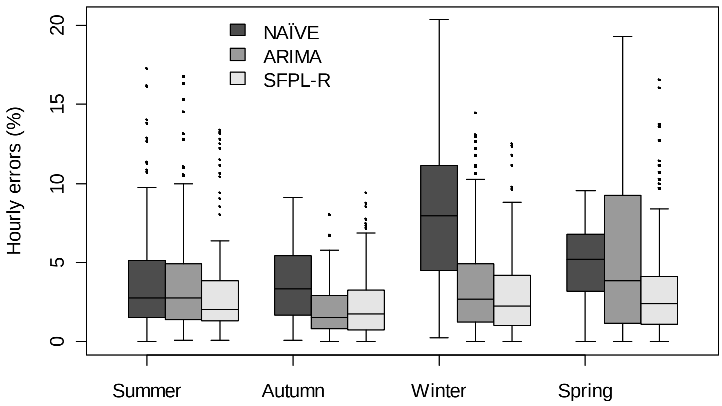

Figure 3: Hourly errors for the electricity demand corresponding to naïve, ARIMA and the recursive version of semi-functional partial linear models.

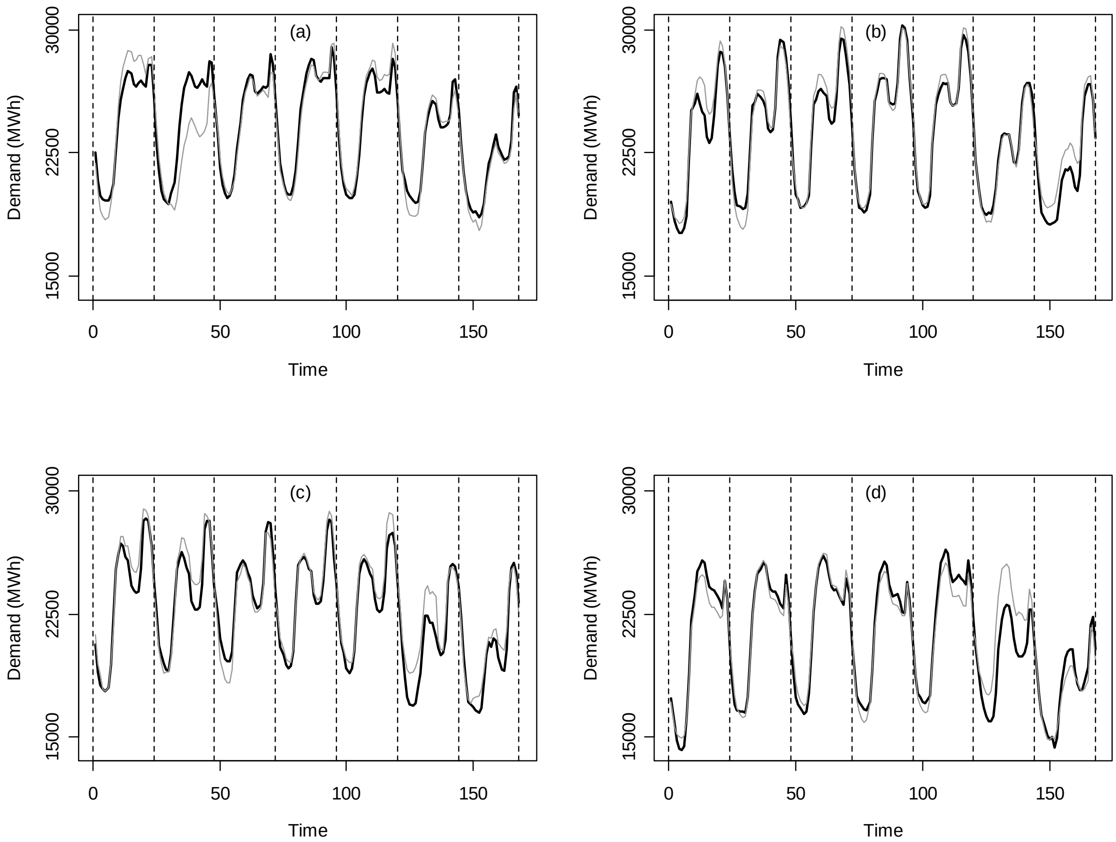

Figure 4: Actual demand (black lines) and predicted values (grey lines) using the recursive version of the semi-functional partial linear model for the analyzed weeks: (a) Summer, (b) Autumn, (c) Winter, and (d) Spring.

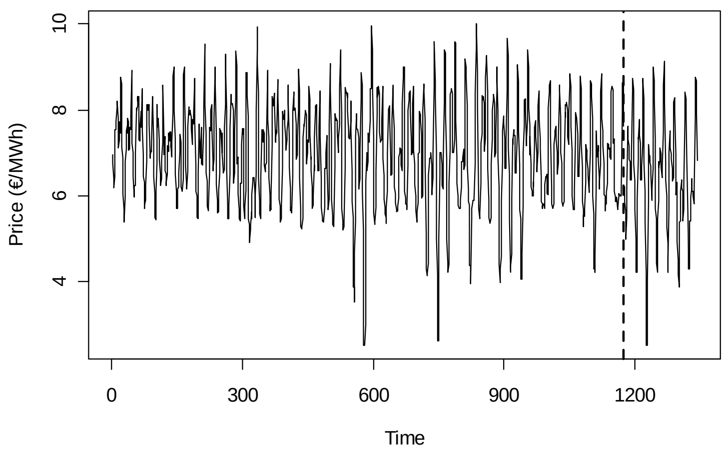

Figure 5: Historical data (left of the vertical line) and data to be predicted (right of the vertical line) corresponding to the electricity prices for the Autumn week.

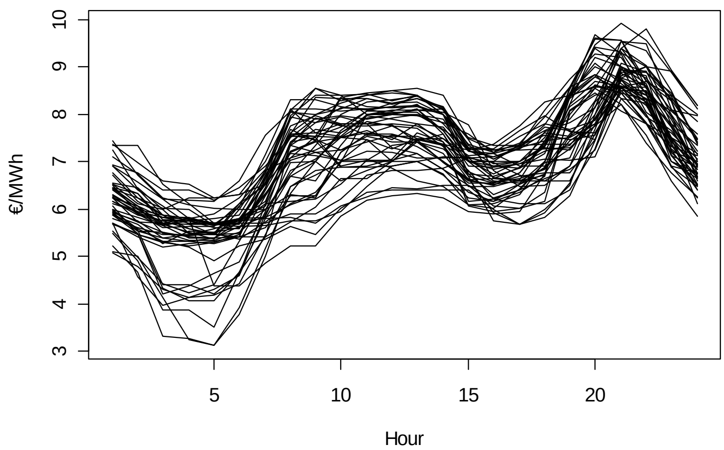

Figure 6: Pruned curves from the electricity price data.

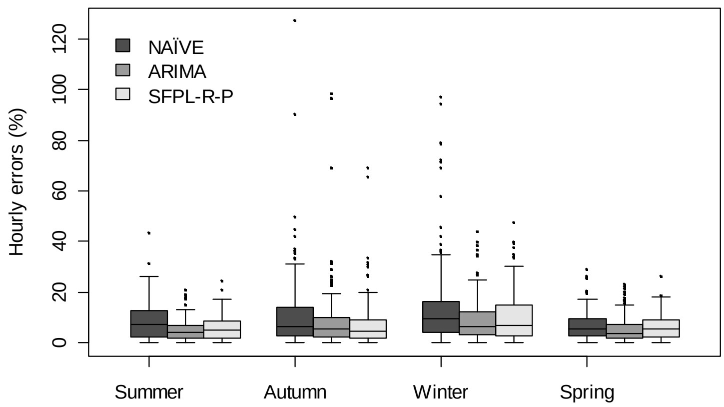

Figure 7: Hourly electricity price errors corresponding to naïve, ARIMA and semi-functional partial linear models (with recursive estimator on the pruned curves).

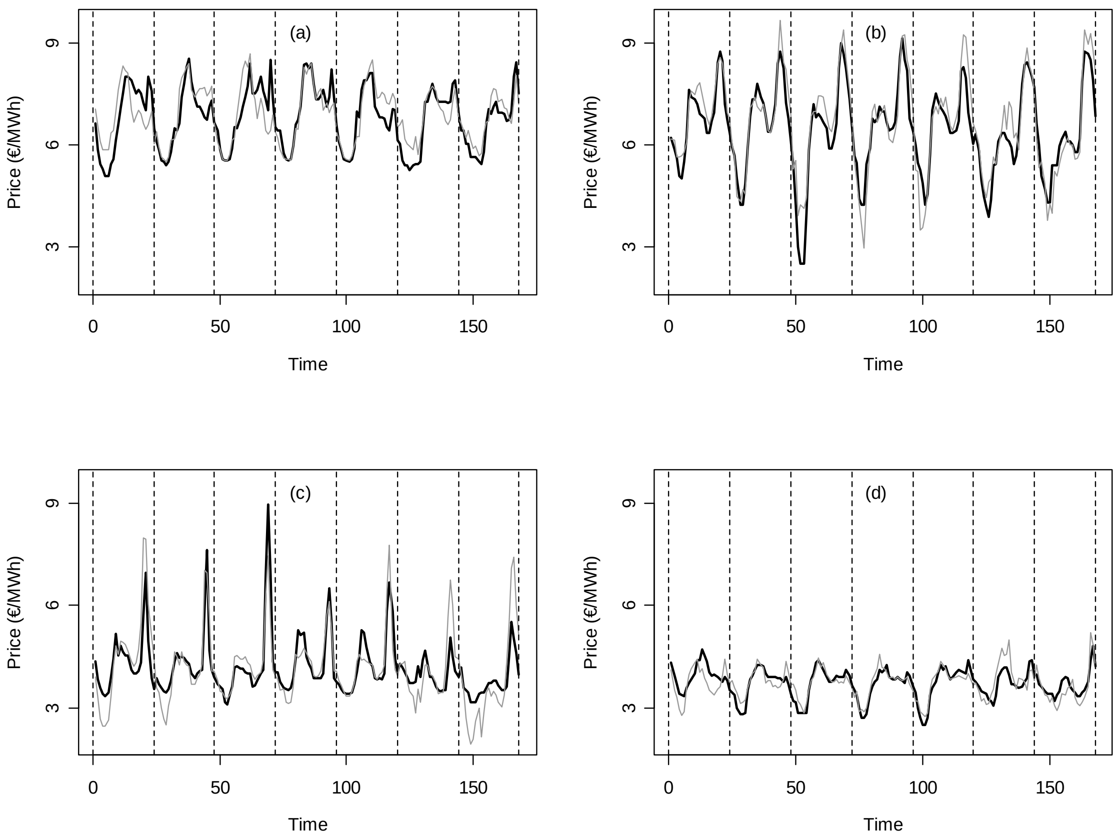

Figure 8: Actual prices (black lines) and predicted values (grey lines) for the semi-functional partial linear model with the recursive method using the pruned data for the analyzed weeks: (a) Summer, (b) Autumn, (c) Winter, and (d) Spring.

Table 1: Relative daily and weekly forecasting errors (%) for the electricity demand.

		DE							
		Mon	Tue	Wed	Thu	Fri	Sat	Sun	WE
Summer	NAÏVE	2.68	2.29	4.78	5.54	4.38	4.14	3.02	3.83
	ARIMA	2.03	3.30	4.84	5.16	3.76	3.83	2.53	3.64
	FNP-D	2.10	2.37	4.22	4.18	3.37	3.97	2.81	3.29
	FNP-R	2.28	2.43	3.92	3.00	3.12	3.28	2.65	2.95
	SFPL-D	2.42	2.61	3.50	3.44	3.54	3.77	2.42	3.10
	SFPL-R	2.26	2.51	3.89	3.24	3.19	3.47	2.62	3.03
Autumn	NAÏVE	4.91	4.99	3.82	2.66	2.50	2.93	3.60	3.63
	ARIMA	2.13	2.11	2.96	1.95	1.98	1.26	1.21	1.94
	FNP-D	1.70	2.46	5.52	5.14	5.96	3.68	2.78	3.89
	FNP-R	2.00	3.51	5.33	4.83	6.06	4.11	3.36	4.17
	SFPL-D	1.86	2.75	3.38	4.12	3.92	1.80	2.10	2.85
	SFPL-R	1.95	2.85	2.58	2.91	3.28	1.58	1.63	2.40
Winter	NAÏVE	4.84	6.55	8.01	10.60	12.07	11.08	4.39	8.22
	ARIMA	2.37	4.07	5.62	3.44	3.99	4.16	2.01	3.67
	FNP-D	3.35	6.42	8.49	5.03	7.60	6.24	4.11	5.89
	FNP-R	2.89	5.82	6.97	4.07	5.88	5.11	3.40	4.88
	SFPL-D	2.97	4.53	3.42	2.31	4.27	3.32	1.96	3.25
	SFPL-R	2.88	3.66	3.08	2.62	4.25	3.22	1.60	3.04
Spring	NAÏVE	4.82	5.73	4.16	4.47	4.95	5.26	4.94	4.90
	ARIMA	2.07	4.58	8.40	6.02	6.63	6.61	4.00	5.47
	FNP-D	3.18	5.43	10.43	8.34	9.84	9.16	6.59	7.57
	FNP-R	3.01	5.29	7.77	6.61	7.52	8.16	4.84	6.17
	SFPL-D	2.27	2.93	2.95	3.24	4.19	4.31	4.22	3.45
	SFPL-R	2.45	3.73	3.81	3.69	3.87	3.53	2.40	3.35

Table 2: Relative daily and weekly forecasting errors (%) for the electricity price (2008).

		DE							
		Mon	Tue	Wed	Thu	Fri	Sat	Sun	WE
Summer	NAÏVE	10.57	10.23	9.32	6.98	7.76	5.72	10.22	8.69
	ARIMA	6.50	5.28	6.34	4.25	3.82	4.83	5.30	5.19
	FNP-D	5.01	4.98	5.40	4.52	5.70	4.60	5.17	5.05
	FNP-R	4.94	5.50	9.83	5.99	6.38	8.34	7.45	6.92
	FNP-D-P	5.42	5.42	6.55	4.46	5.38	4.94	5.85	5.43
	FNP-R-P	4.89	5.57	8.89	5.17	6.60	8.75	7.42	6.76
	SFPL-D	4.53	4.97	6.67	4.25	5.41	4.65	5.44	5.13
	SFPL-R	4.96	5.61	7.82	3.67	4.48	5.75	6.44	5.53
	SFPL-D-P	5.58	4.54	6.29	4.77	5.83	5.33	5.31	5.38
	SFPL-R-P	5.23	5.73	7.62	3.89	5.13	7.55	6.99	6.02
Autumn	NAÏVE	19.90	25.62	7.80	7.44	8.00	3.93	7.02	11.39
	ARIMA	15.10	16.41	6.58	5.03	6.03	3.47	6.76	8.48
	FNP-D	9.30	11.31	11.80	9.69	10.05	6.43	10.04	9.80
	FNP-R	11.78	14.64	10.39	9.08	6.25	4.11	7.09	9.05
	FNP-D-P	8.85	11.15	12.18	10.13	9.35	6.03	8.93	9.52
	FNP-R-P	12.28	14.51	10.55	8.61	6.83	4.58	7.72	9.30
	SFPL-D	10.34	12.93	5.79	7.74	7.45	4.92	11.22	8.63
	SFPL-R	10.88	11.60	6.91	7.12	6.60	4.32	10.96	8.34
	SFPL-D-P	10.17	10.25	5.84	8.42	8.67	4.93	9.68	8.28
	SFPL-R-P	11.24	11.85	4.92	5.26	6.10	4.38	8.75	7.50

Table 3: Relative daily and weekly forecasting errors (%) for the electricity price (2009).

		DE							
		Mon	Tue	Wed	Thu	Fri	Sat	Sun	WE
Winter	NAÏVE	22.37	20.91	10.41	9.29	6.91	15.12	18.73	14.82
	ARIMA	9.92	9.96	8.86	6.09	3.94	10.22	13.21	8.88
	FNP-D	8.60	14.28	13.43	7.67	8.44	12.99	17.53	11.85
	FNP-R	7.12	10.22	9.73	7.28	8.79	12.64	15.67	10.21
	FNP-D-P	9.94	13.88	12.03	6.49	7.02	12.42	14.64	10.92
	FNP-R-P	6.68	11.00	8.28	6.14	7.53	12.58	15.27	9.64
	SFPL-D	9.45	14.77	10.06	5.62	5.60	14.27	16.22	10.86
	SFPL-R	9.18	12.57	13.84	7.05	6.23	12.14	18.57	11.37
	SFPL-D-P	8.76	12.28	9.14	4.86	5.65	12.31	14.64	9.67
	SFPL-R-P	8.65	13.47	10.19	5.48	5.60	12.46	17.57	10.49
Spring	NAÏVE	9.88	7.61	5.60	6.93	6.53	7.02	5.93	7.07
	ARIMA	4.81	8.24	4.15	4.89	5.28	3.84	5.65	5.26
	FNP-D	8.99	12.67	11.73	8.39	8.08	6.55	10.21	9.52
	FNP-R	8.37	9.55	11.35	7.85	9.31	5.63	11.52	9.08
	FNP-D-P	8.53	11.25	11.10	7.99	7.83	6.06	8.92	8.81
	FNP-R-P	8.54	8.89	11.15	8.62	8.89	5.85	11.2	9.02
	SFPL-D	9.22	9.76	6.02	6.23	8.37	6.85	7.46	7.70
	SFPL-R	7.71	6.95	6.25	7.08	5.48	4.69	5.83	6.28
	SFPL-D-P	9.26	9.66	5.29	6.35	6.89	6.59	9.26	7.61
	SFPL-R-P	7.79	6.70	6.22	6.09	5.89	4.61	6.35	6.23