Forecasting hourly supply curves in the Italian Day-Ahead electricity market with a double-seasonal SARMAHX model

Guillermo Mestre[a,*], José Portela[a,b], Antonio Muñoz San Roque[a], Estrella Alonso[c]

[a] *Universidad Pontificia Comillas. Escuela Técnica Superior de Ingeniería ICAI. Instituto de Investigacińn Tecnológica. Madrid, Spain*
[b] *Universidad Pontificia Comillas. Facultad de Ciencias Económicas y Empresariales ICADE. Madrid, Spain*
[c] *Universidad Pontificia Comillas. Escuela Técnica Superior de Ingeniería ICAI. Madrid, Spain*

Abstract

This paper is devoted to the short-term forecasting of the hourly aggregated supply curves in Day-Ahead electricity markets. The time series of supply curves can be considered as a functional time series, which is the realization of a stochastic process where each observation is a continuous function defined on a finite interval. In order to forecast these time series, models that operate with continuous functions are required. The standard approach for estimating these supply curves relies on dimensionality reduction techniques, hence losing some information in the process. This paper proposes a functional forecasting model that uses the continuous supply curves as inputs and does not require turning the curves to a limited number of components, thus avoiding the corresponding information loss. The proposed model is based on a double-seasonal functional SARMAHX model which extends the classical ARMA models and it is able to capture the daily and weekly seasonality of the series of supply curves. In addition, exogenous variables can be included to account for the external factors that influence the offering behavior of the agents. The effectiveness of the proposed model is illustrated by forecasting the hourly aggregated supply curves of the competitors of the main Italian generation company and is compared to other reference models found in the literature.

Keywords: Supply curve forecasting, functional SARMAHX model, Functional data analysis, functional time series.

1. Introduction

The electric power industry in different countries has experienced a deregulation process in the last decades which has given rise to liberalized markets. They allow companies to trade energy in organized auctions. Generally, day-ahead electricity markets are based on sealed-bid auctions where companies submit their selling offers and buying bids to the Market Operator who then determines the market-clearing price and the set of accepted bids and offers for each time period [1].

In a simple-bid market, each offer (or bid) is defined by a price p and a quantity q, which refers to the amount of energy q the agent is willing to sell (or buy) at that price p. By sorting the selling (buying) offers in increasing (decreasing) prices, the aggregated supply (demand) function for the agent is built. Once all the agents have submitted their supply curves, the sum of the supply functions results in the system supply function $S_h(p)$ and the sum of the demand functions of each firm results in the system demand function $D_h(p)$. Market-clearing price p_h^* is computed for each hour as the intersection of the system aggregated supply and demand curves, hence $D_h(p_h^*) - S_h(p_h^*) = 0$. Then, two possible outcomes generally take effect depending on the auction type. In pay-as-bid auctions, prices paid to winning suppliers are based on their actual bids. On the other hand, in marginal pricing auctions all suppliers are paid the market-clearing price.

For a given generation company i participating in a market, it is of utmost importance to plan ahead and manage the available resources as efficiently as possible. The forecasts of different significant market indicators and variables can provide useful estimations for the agents. For example, load forecasting allows for an efficient management of resources, optimal scheduling and production planning for minimizing generation costs [2]. Estimates of the price allow to plan ahead and cover their operation costs and hedge against price movements [3]. As a consequence, load [4], price [5, 6] or wind forecasting [7] are widely studied and continuously being improved.

Furthermore, the market agent is usually interested in optimizing its bidding strategy in the market [8]. This can be done using Residual Demand Curves (RDCs)[9, 10, 11, 12]. RDCs can be defined for each hour as the function that models the offering and bidding behavior of all the competitors and can be calculated as

$$q_i = R_h(p) = D_h(p) - S_h^{-i}(p) \qquad (1)$$

*Corresponding author

Email addresses: guillermo.mestre@comillas.edu (Guillermo Mestre), jose.portela@iit.comillas.edu (José Portela), antonio.munoz@iit.comillas.edu (Antonio Muñoz San Roque), estrella.alonso@comillas.edu (Estrella Alonso)

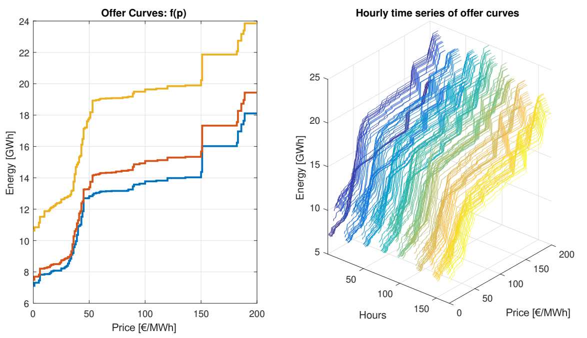

Figure 1: Supply curves in electricity markets. On the left, three offer curves defined in the price range 0 to 200 €/MWh are shown . On the right, the hourly sequence of offer curves for one week is plotted. For each hour, a function defined in the price range 0 to 200 €/MWh is observed.

where $S_h^{-i}(p)$ is the supply function of the competitors of firm i, which is given by the system supply function $S_h(p)$ minus the firm's supply function $S_h^i(p)$ i.e. $S_h^{-i}(p) = S_h(p) - S_h^i(p)$. For a given price value p, R_h gives the maximum energy quantity q_i that the generation company i can sell in the market at hour h.

Assuming that the demand is inelastic, the residual demand at Eq. (1) becomes

$$q_i = D_h - S^{-i}(p). \tag{2}$$

Therefore, the problem of forecasting the residual demand could be split into the forecast of the hourly inelastic demand and the forecast of the supply curves. Demand forecasting is widely studied in the literature [13, 14, 15], hence this work will focus on forecasting the supply functions.

A time series of supply curves can be considered as a time sequence of curves. Figure 1 shows a representation of an hourly supply curves time series. The high dimensionality and complexity of the supply curves make classical multivariate forecasting techniques obsolete, highlighting the need for new methodologies that are able to capture the dynamic of such processes.

For example, [16] uses Principal Component Analysis to reduce the dimensionality of the curves. Then, the reduced set of variables is forecasted using time series models and finally, the estimated curves are obtained by reconstructing the forecasted scores. Recently, [17] proposed a price forecasting approach based on estimating supply and demand curves by means of lasso-based estimation methods, which was extended in [18] to obtain probabilistic forecasts of electricity prices by modeling the market bids for the supply and demand side. In [19], the authors use functional autoregressive models to obtain estimations of sale and purchase curves in order to forecast future values of electricity prices.

In this paper, a functional forecasting approach is proposed. Functional data analysis has been a growing field in statistics, which studies data that are observed in the form of functions (see [20] and [21] for an overview). In particular, supply functions can be viewed as a Functional Time Series $\{S_t\}_{t\in\mathbb{N}}$, which is a time-ordered sequence of random functions $\{S_t(p); t = 1, 2, ..., T\}$. Modeling supply curves as functional data can take advantage of the mathematical background to mitigate the dimensionality issues that often arise with high dimensional data.

In fact, functional forecasting techniques have been successfully applied in electicity markets. [22] proposed a non-parametric functional model to obtain prediction intervals for French electricity consumption. Although the prediction bands obtained by their model are able to capture the shapes of the load series, the authors admit that the model could be improved with the inclusion of exogenous variables such as the temperature. [23] obtains forecasts of the German electicity market demand using a forecasting model based on functional data analysis of generalized quantile curves, whereas [24] applied non-parametric techniques to forecast the residual demand curves of the Spanish electricity market. In the latter, the authors include exogenous variables (such as the wind production and the electricity demand) into their model, highlighting the need to incorporate these variables to capture the dynamic of the series. Finally, in order to forecast electricity price profiles, [25, 26] applied a semi-parametric model that incorporates scalar exogenous variables that are relevant for price forecasting (such as the wind production and the electricity demand) obtaining pointwise prediction bands. [27] developed a functional factor model that, unlike the other models, relies on the decomposition of the daily price profiles into a functional basis. Finally, [28] proposed a functional seasonal ARMAX model to model the daily price profiles of the Spanish and German electricity market. However, the previous models analyzed daily time series, which only exhibit weekly seasonality. The supply curves form an hourly functional time series, which will exhibit daily and weekly seasonalities. Hence, the previous functional approaches to the supply curves problem are insufficient. The work [19] illustrates how functional modeling can be used to forecast this kind of data: the authors propose functional autoregressive models to obtain forecasts of supply and demand curves. However, this approach does not account for the effect that exogenous covariates could have on the shape of the curves, as well as assuming an autoregressive structure on the data without a proper test.

Based on the model developed in [28], which extends the ARMAX model to the functional framework, this paper proposes several significant theoretical and methodological improvements that allows the proposed SARMAHX model to be successful for forecasting hourly supply curves. More specifically the contributions presented are:

- Introducing up to two seasonal terms in the formulation of the SARMAHX model, which are needed in order to take into account seasonal dependencies of hourly functional time series (daily and weekly).

- Including scalar covariates in the SARMAHX model to account for scalar inputs. Exogenous variables, such as the wind and solar production, are necessary to capture the complex dynamics of the supply curves, hence a competitive forecasting model for supply curves should be able to include them.

- Introducing a new functional operator that can improve the performance of the SARMAHX model. If the modeled supply curves do not exhibit cross-correlations between curves, this new functional operator can ease the optimization of the parameters, improving the estimation of the model.

- Including an identification and diagnosis procedure for the SARMAHX model based on the functional autocorrelation of the series and the residuals of the model. This methodology is of utmost importance when fitting the SARMAHX model, allowing the practitioner to model the dynamic behavior of the supply curves with a graphical representation of the autocovariance structure of the data.

- Successful application of the proposed model to a real case study in the Italian Day-Ahead electricity market.

The paper is structured as follows: Section 2 introduces the proposed double-seasonal SARMAHX model with exogenous variables. Section 3 is devoted to the identification and diagnosis methodology developed for the SARMAHX model by means of the lagged autocorrelation of the functional time series and the residuals of the model. Section 4 evaluates the performance of the proposed model comparing the methodology developed in this paper against other well known forecasting models for supply curve forecasting. Finally, Section 5 provides the concluding remarks.

2. Double seasonal Hilbertian SARMAHX model

This section introduces the functional SARMAHX model. It is defined following the standard time series modeling approach proposed in [29] but extended to functional time series using Hilbert operators [28]. The proposed SARMAHX model generalizes the scalar ARMAX model by extending the scalar model to functional time series using Hilbert operators as the parameters of the model. While the parameters of the classical ARMAX model are scalar values, the parameters of the SARMAHX model are functional operators that model the relation between input and output curves. As such, the modelization and estimation of these functional operators is also detailed in this section.

A functional time series Y is defined as a sequence of functional observations $\{Y_t(v); t = 1, 2, ..., T; v \in V\}$ where each observation at time t is a continuous function taking values on the interval V. Following [21], these functions are considered as elements belonging to the L^2 Hilbert space of real square integrable functions defined on an interval V, this is, $\int_V (Y_t(u))^2 du < \infty$.

The SARMAHX$(P_0, Q_0)\times(P_1, Q_1)_{s_1}\times(P_2, Q_2)_{s_2}$ model is a functional Seasonal Autoregressive Moving Average Hilbertian model with two seasonalities (although it could be generalized to any number of seasonalities) which includes both functional and scalar explanatory variables. The full expression for the model is defined as follows:

$$\prod_{j=0}^{2}\left(I - \sum_{i=1}^{P_j} \Psi_{j,i} B^{i\cdot s_j}\right)(Y_t) = \prod_{k=0}^{2}\left(I - \sum_{l=1}^{Q_k} \Theta_{k,l} B^{l\cdot s_k}\right)(\varepsilon_t) + \sum_{z=1}^{Z_f} \Gamma_z^f(X_t^z) + \sum_{z=1}^{Z_c} \Gamma_z^c(x_t^z) \tag{3}$$

where:

- $\{Y_t\,(v)\,; t = 1, 2, ..., T; v \in V\}$ a stationary and zero-mean functional time series
- $\{X_t^z(v_z); z = 1, 2, \cdots, Z_f; t = 1, 2, \cdots, T; v_z \in V_z\}$ a set of Z_f functional covariates
- $\{x_t^z; z = 1, 2, \cdots, Z_c; t = 1, 2, \cdots, T\}$ a set of Z_c scalar covariates
- ε_t is a functional white noise process.
- I is the identity operator.
- Parameters P_0, P_1 and P_2 are the regular and the two seasonal autoregressive orders respectively.
- Parameters Q_0, Q_1 and Q_2 are the regular and the two seasonal moving average orders respectively.
- Parameters s_1 and s_2 are the seasonal time span for repeated patterns. Parameter s_0 is equal to 0.
- $\Psi_{0,i}$, $\Psi_{1,i}$ and $\Psi_{2,i}$ are the regular and seasonal autoregressive operators.
- $\Theta_{0,l}$, $\Theta_{1,l}$ and $\Theta_{2,l}$ are the regular and seasonal moving average operators.
- Γ_z^f are the operators related to the Z_f explanatory variables.
- Γ_z^c are the operators related to the Z_c explanatory variables.
- B^n is the lag operator which is defined as $B^n Y_t = Y_{t-n}$ where $n \in \mathbb{N}$.

In [28], integral operators in the L^2 space were used for the ARMA terms, which have the form:

$$\Psi_i(f)(v) = \int \psi_i(u, v) f(u) du \qquad f \in L^2 \tag{4}$$

The operators for the remaining terms are defined alike except for the scalar covariates' operator, which is defined as $\Gamma_z^c(x)(v) = \beta_z(v)\,x$.

When this model is used for forecasting, all the error terms are unobserved. Thus, expanding the functional operators to its integral form and simplifying the lag operators, the empirical forecast equation can be expressed as follows:

$$\begin{aligned}\widehat{Y}_t(v') = & \sum_{i=0}^{p}\sum_{j=0}^{P_1}\sum_{k=0}^{P_2}\iiint A(v,v',u,u')Y_{t-\tilde{t}}(u)du du' dv \\ & -\sum_{i=0}^{q}\sum_{j=0}^{Q_1}\sum_{k=0}^{Q_2}\iiint B(v,v',u,u')\widehat{\varepsilon}_{t-\tilde{t}}(u)du du' dv \\ & +\sum_{z=1}^{Z_f}\int \rho_z(v_z,v')\,X_t^z(v_z)\,dv_z + \sum_{z=1}^{Z_c}\beta_z(v')\,x_t^z,\end{aligned} \tag{5}$$

where $u, v, v' \in V$, $v_z \in V_z$, $\tilde{t}$ is defined as

$$\tilde{t} = i + j \cdot s_1 + k \cdot s_2,$$

$\widehat{\varepsilon}_t$ is the estimation of past innovations, which is given by:

$$\widehat{\varepsilon}_t(u) = Y_t(u) - \widehat{Y}_t(u) \tag{6}$$

and

$$A(v,v',u,u') = \psi_{0,i}(v,v')\,\psi_{1,j}(u',v)\,\psi_{2,k}(u,u') \tag{7}$$

$$B(v,v',u,u') = \theta_{0,i}(v,v')\,\theta_{1,j}(u',v)\,\theta_{2,k}(u,u') \tag{8}$$

In order to ease the comprehension of the SARMAHX model, the following example is detailed. The SARMAHX $(1,1)\times(1,0)_{24}$ model would be defined following equation (3):

$$\begin{aligned}Y_t = & \Psi_{0,1}(Y_{t-1}) + \Psi_{1,1}(Y_{t-24}) - \Psi_{0,1}\Psi_{1,1}(Y_{t-25}) \\ & - \Theta_{0,1}(\varepsilon_{t-1}) + \varepsilon_t,\end{aligned} \tag{9}$$

where the terms are, respectively, the regular autoregressive, the seasonal autoregressive, the autoregressive seasonal interaction and the moving average term. It should be noted that the term $\Psi_{0,1}\Psi_{1,1}(Y_{t-25})$ denotes the composition, i.e. $\Psi_{0,1}(\Psi_{1,1}(Y_{t-25}))$. Then, substituting each operator by its integral expression, the forecasting equation becomes:

$$\begin{aligned}\widehat{Y}_t(v') = & \int \psi_{0,1}(u,v')Y_{t-1}(u)\,du + \int \psi_{1,1}(u,v')Y_{t-24}(u)\,du \\ & - \iint \psi_{0,1}(v,v')\psi_{1,1}(u,v)Y_{t-25}(u)\,du\,dv \\ & - \int \theta_{0,1}(u,v')\widehat{\varepsilon}_{t-1}(u)\,du\end{aligned} \tag{10}$$

In this paper, concurrent operators are proposed instead of integral operators. They are defined as:

$$\Psi_i(f)(v) = \psi_i(v)f(v) \qquad f \in L^2 \tag{11}$$

The representation capabilities of these operators are more limited than integral operators, as they are not able to model cross correlations between curves (i.e. the first part of the input curve affecting the last part of the output curve). However, the model is simplified, which can ease the optimization of the parameters. For example, Eq. (10) using concurrent operators would have the following expression:

$$\begin{aligned}\widehat{Y}_t(u) = & \psi_{0,1}(u)Y_{t-1}(u) + \psi_{1,1}(u)Y_{t-24}(u) \\ & - \psi_{0,1}(u)\psi_{1,1}(u)Y_{t-25}(u) - \theta_{0,1}(u)\widehat{\varepsilon}_{t-1}(u)\end{aligned} \tag{12}$$

As a conclusion, the resulting model admits a wide variety of configurations: autoregressive and moving average terms up to two seasonalities as well as the inclusion of scalar and functional explanatory variables. Therefore, this model is suitable for most functional time series found in the electricity markets.

2.1. Estimation procedure

In order to fit the SARMAHX model, each integral operator must be estimated from the observed data. As shown in Eq. (4), estimating an integral operator Ψ implies estimating the associated kernel surface $\psi(u,v)$. The standard approach for the estimation of these operators is to project the kernel into a subspace spanned by a functional basis, as seen in [30, 31]. Then, the coordinates of the kernel in that subspace are optimized so that the forecasting error is minimized. Thus, the selection of the functional basis is critical when applying these estimation methods.

In order to avoid choosing one fixed functional basis, the SARMAHX model follows a novel approach to functional parameter estimation: each kernel function is modeled as a finite sum of bivariate sigmoid functions [28]. Sigmoid functions are universal function approximates [32] which are commonly used in neural networks due to their properties to model non-linear relations. As such, each bivariate kernel $\psi(u,v)$ can be modeled as:

$$\psi(u,v) = \alpha_0 + \sum_{g=1}^{G_\psi}\alpha_g \tanh\left(w_{g0} + w_{g1}u + w_{g2}v\right), \tag{13}$$

where $w_{g0}, w_{g1}, w_{g2}, \alpha_g$ and α_0 are the parameters that define each sigmoid. The variables u and v take real values in the intervals in which the functional variables are defined (i.e. V or V_z). In the case of a concurrent operator (11), this expression is simplified, modeling each kernel as a finite sum of univariate sigmoids:

$$\psi(v) = \alpha_0 + \sum_{g=1}^{G_\psi}\alpha_g \tanh\left(w_{g0} + w_{g1}v\right). \tag{14}$$

This approach can be viewed as a Multi-Layer Perceptron (MLP) neural network with a particular configuration: an input layer with two input variables and a

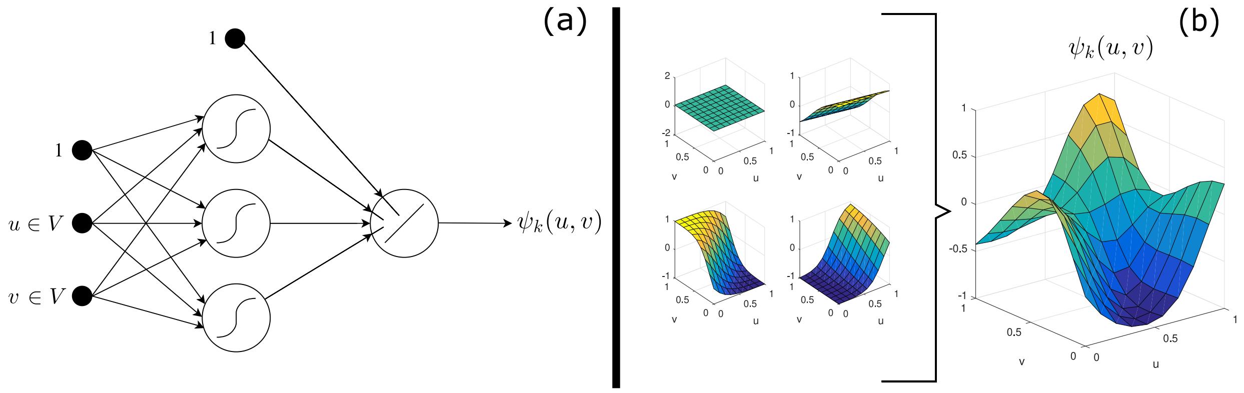

Figure 2: (Part a) Architecture diagram of the neural network used for optimizing the functional parameters of the SARMAHX model. (Part b) Kernel function estimated with 3 bivariate sigmoid functions and a constant surface.

bias; one hidden layer with a number G_ψ of nonlinear hidden units with hyperbolic tangent as the activation function and w_{g0}, w_{g1}, w_{g2} as the weights for each input. Finally, one output layer with one linear output unit having α_g as the weights for the activation of the hidden units. Fig. 2a shows the architecture diagram of the aforementioned network with $G_\psi = 3$ hidden layers. The property of sigmoidal surfaces being universal approximators is illustrated in Fig. 2b, where a rather complex surface is modeled as the sum of 3 sigmoidal functions and a constant surface that models the level of the final surface. It has been observed that using 5 or 6 sigmoid functions when fitting the SARMAHX model is enough in the vast majority of practical applications, due to the flexibility of sigmoid functions.

The parameters $(w_{g0}, w_{g1}, w_{g2}, \alpha_g, \alpha_0)$ completely define each bivariate sigmoid in Eq. (13). Therefore, the proposed SARMAHX model is estimated when values for all these parameters are estimated. In order to achieve this, a low-memory Quasi Newton method with random initial weights has been implemented to optimize these real-valued parameters so as to minimize a certain cost function. In this paper, the cost function C for estimating the SARMAHX model is defined as the sum of the L^2 square errors,

$$C = \sum_{t=1}^{T} e_t, \tag{15}$$

where

$$e_t = \left\| Y_t - \widehat{Y}_t \right\|_{L^2}^2 = \int \left(Y_t(u) - \widehat{Y}_t(u) \right)^2 du. \tag{16}$$

The implemented Quasi Newton method is a gradient descent algorithm, so the derivatives of the error term (15) with respect to the sigmoid parameters are needed. The derivative of the error function with respect to a general parameter W is given by:

$$\frac{\partial C}{\partial W} = \sum_{t=1}^{T} \int 2 \left(Y_t(u) - \widehat{Y}_t(u) \right) \left(-\frac{\partial \widehat{Y}_t(u)}{\partial W} \right) du, \tag{17}$$

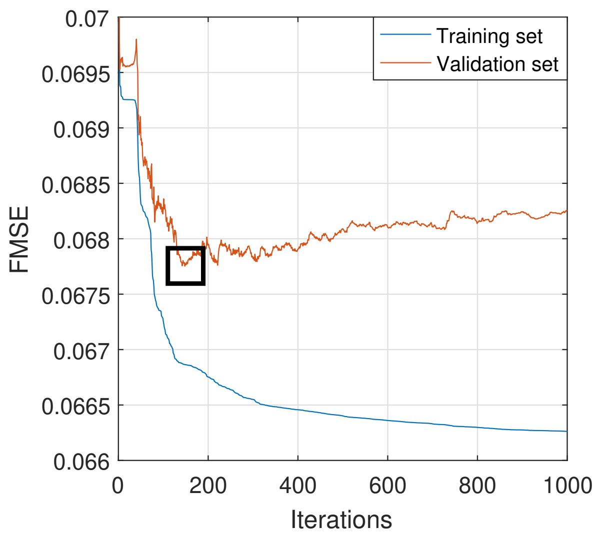

Figure 3: Evolution of both training and validation errors in the optimization process for the SARMAHX model. The training error always decreases in each iteration of the algorithm, but the same is not true for the validation set. The early-stopping method is used to select the parameter configuration that yields a lower validation error (marked in the figure with a square).

where $\frac{\partial \widehat{Y}_t(u)}{\partial W}$ is the derivative of the estimation with respect to the generic parameter W. In order to reduce computational times, analytical expressions for these derivatives have been obtained. The reader is referred to [28] for the formulation of the derivatives of the cost function with respect to the weight parameters.

However, neural network procedures are known to have a tendency towards overfitting. If no regularization techniques are used in the learning process, the resulting network will be so closely fitted to the training data that it causes the model to not generalize well in the presence of new data. In order to prevent this effect, early stopping methodology is used: the observed data is split into a training and a validation set and executing the optimization algorithm with the training set, thus minimizing the cost function over these data. At the same time, the adjusted model in each iteration is being evaluated with the validation set. Fig. 3 illustrates this methodology: the training error of the model decreases in each iteration while the error among the validation set is oscillating. As the best model is the one with better generalization performance among an unknown dataset, the final model will be the one that produces a lower error in the validation data.

The training and evaluation of the SARMAHX model has been implemented in C programming language. This includes the quasi-Newton algorithm and the derivatives which feed the quasi-Newton. It is a common procedure in neural network implementations to normalize or standardize the input and output variables, so both the time series $\{Y_t\}$ and the functional covariates $\{X_t\}$ are centered and normalized as

$$Y_t^*(v) = Y_t(v)/\sigma_Y \quad , \quad X_t^*(v) = X_t(v)/\sigma_X, \tag{18}$$

where σ_Y and σ_X are normalization values for which $Y_t^*(v) \in [-1, 1]$ and $X_t^*(v) \in [-1, 1]$. Similarly, in order to help the algorithm to converge faster and as a way to minimize the risk of falling into a local minimum, the domain of the time series $\{Y_t^*\}$ and the functional covariates $\{X_t^*\}$ is moved to the range $[-1, 1]$. By default, the sigmoid parameters are initialized with small values, causing the sigmoids to be centered around 0. As such, the position of the sigmoids wont need to move far from that initial position.

2.2. Increasing transformation

In the previous sections, the estimation procedure for the SARMAHX has been detailed. However, when forecasting certain types of data, some post-processing method has to be implemented to improve the raw output of the model. One such case is the one that concerns us: forecasting aggregated supply curves in the Day-Ahead electricity market. These curves are always monotonically increasing curves, due to the sheer definition of the supply functions. The formulation of the SARMAHX model does not ensure that the estimations of the model are non-decreasing functions, so in order to obtain faithful estimations of these curves with the proposed model, output curves of the model $\widehat{Y}_t$ are transformed into monotonically increasing curves $\widehat{Y}_t^+$ obtained as the solution of the following optimization problem:

$$\underset{\widehat{Y}_t^+}{\text{minimize}} \quad \frac{1}{2}\left\|\widehat{Y}_t^+ - \widehat{Y}_t\right\|^2 \tag{19a}$$

$$\text{subject to} \quad 0 \le \widehat{Y}_t^+(v_i) \le \widehat{Y}_t^+(v_{i+1}),\ i = 1, \ldots, N-1, \tag{19b}$$

where $\{v_1, \ldots, v_N\}$ denotes the discretization points of the functional observations. The objective function to be minimized (19a) is the distance between the estimated curves and their monotonic transformation, while constraint (19b) ensures that the new curve is non-decreasing. The optimization problem is solved by linear least squares.

This transformation guarantees that the estimated curves are monotonically increasing, while retaining the shape that has been estimated by the SARMAHX model. It should be noted that the problem of estimating curves that can be non-monotonic curves is a known issue that happens to the vast majority of forecasting models, for example if the parameter estimates for one regressor is negative and a new value of this variable is positive. Henceforth, in order to provide a fair comparison between the SARMAHX model and other forecasting models found in the literature, the proposed transformation will be applied to the estimates of all models considered in this paper.

3. Functional autocorrelation and model identification

A common procedure in time series analysis is analyzing the correlation structure of the series in order to select the proper configuration for the model. Order selection is a central issue in forecasting, and while the literature for scalar time series is extensive (see [33, 34, 29] for a comprehensive review), few studies have addressed the problem of selecting the order of a functional time series model. Initial work on this topic has focused on developing statistical tests for the adequacy of a more complex model instead of a simpler one. In [35], the authors develop a statistical test to check if using a functional autoregressive model of order p ($ARH(p)$) produces better results than using a $ARH(p-1)$. Other works, such as [36] and [37] measure the deviation between the residuals of a fitted functional linear model and a functional white noise series to diagnose the fitted model. However, none of these techniques provides the user with a general identification tool that can be used to select the ARMA order and seasonality of a functional linear model.

Classical identification methods for scalar time series, such as the Box-Jenkins methodology [29], use the sample autocorrelation function of the time series in order to identify the underlying correlation structure of the series. This concept can be extended to the functional framework using the functional equivalent, the lagged autocovariance function. Given a functional time series $\{Y_t(v); t = 1, 2, ..., T; v \in V\}$, one can define the sample lagged autocovariance functions of the series for lag h as

$$\hat{C}_h(u, v) = \frac{1}{T}\sum_{i=1}^{T-h}(Y_i(u) - \overline{Y}_T(u))(Y_{i+h}(v) - \overline{Y}_T(v)) \tag{20}$$

where

$$\overline{Y}_T(u) = \frac{1}{T}\sum_{i=1}^{T} Y_i(u) \tag{21}$$

denotes the sample mean function. The lagged autocovariance function $\hat{C}_h(u, v)$ shows the existing relationship

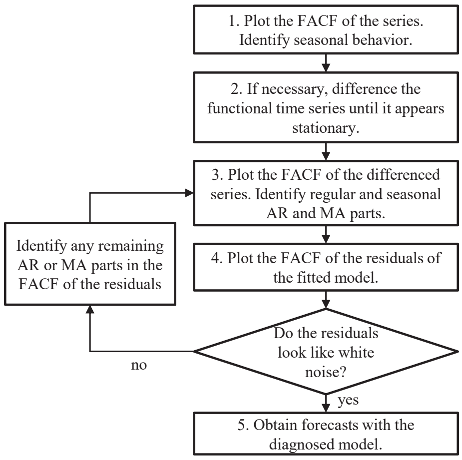

Figure 4: General procedure for identifying and diagnosing a SARMAHX model using the proposed FACF.

between the time series of curves with its lagged functional observations. By considering the L^2 norm of these surfaces, [38] develops a Functional Autocorrelation plot (FACF) defining the functional autocorrelation coefficient at lag h as

$$\hat{\rho}_h = \frac{\|\hat{C}_h\|}{\int \hat{C}_0(u,u)du} \tag{22}$$

The distribution of these coefficients under the hypothesis of white noise was obtained in [39], thus it can be used to identify if there is a significant correlation between observations t and $t-h$.

This paper shows how the FACF can be used as an identification methodology for the proposed SARMAHX model in a practical application. Firstly, any seasonalities in the data can be detected by looking at the FACF plot of the series. As in the univariate case, integrated time series will show a slow decrease in the FACF [40, 41], hence the functional time series must be differentiated to obtain a stationary series. Secondly, the FACF of the stationary series can be used to identify the AR and MA orders of the SARMAHX model following a similar approach as in the scalar case [29]. According to (3), if the model has been succesfully identified, its residuals should look like a functional white noise process, thus not showing any kind of autocorrelation. This can be checked by analyzing the FACF of the residuals of the fitted model: if the FACF of the residuals reveals that there is some significative correlation remaining, the orders of the SARMAHX model should be modified in order to capture that relation. Finally, the residuals of the model can be diagnosed by means of their FACF. If significant autocorrelations are detected, then the null hypothesis of functional white noise is rejected and the identification process continues. Once all the values of the FACF are within the threshold limits, the identification of the model comes to an end and forecasts can be obtained. Fig. 4 summarizes this procedure.

4. Case study: Forecasting supply curves in the Italian Day-Ahead electricity market

This section is devoted to the case study of forecasting supply curves in electricity markets. This is very useful for electricity companies that want to foresee the competitors' behavior and optimize their bidding strategy. The set up of this case study is based on [16], where the author analyzes the forecasting of supply functions of competitors in the Italian electricity market using functional models. Despite the fact the the Italian market consists in 6 inteconnected zones, for simplicity, and following [16], the Italian market is considered as a single market, as if there were no saturation constraints between zones. The detailed research of each zone would be of much interest, but is out of the scope of this paper.

This paper is devoted to illustrate the applicability of the proposed SARMAHX model to obtain short-term forecasts of the aggregated supply curves in the Italian Day-Ahead Market. The Day-Ahead Market (MGP) is the venue for the trading of electricity supply offers and demand bids for each hour of the next day. The MGP opens at 8:00 A.M. of the ninth day before the day of delivery and closes at 12:00 P.M. of the day before the day of delivery. Participants submit their bids where they specify the amount and the maximum and minimum price at which they are willing to buy and sell electricity for each hourly auction. Once the MGP closes, bids and offers are accepted based on the economic merit-order criterion and taking into account transmission capacity limits between zones. Where these constraints are binding, the market effectively splits and distinct zonal clearing prices are determined. While generators are paid the relevant zonal price, the accepted demand bids pertaining to consuming units are evaluated at the single national price (PUN) which is computed as the average of the zonal prices weighted by zonal consumption [42].

The results of the MGP are made known within 12.55 P.M. of the day before the day of delivery. As pointed out in [19], due to confidentiality policies, the Italian Market Operator (GME) does not publish the detailed information relative to demand bids and supply offers for a confidentiality period of seven days, beginning on the day following the close of the market sitting to which they refer. However, for market participants, anonymized confidential results (including the acceptance/rejection of the bids and the aggregated supply curves without information about the geographical zone of each bid) are available at `https://www.IPEX.it`. As this study is concerned with forecasting the aggregated supply curves (instead of the zonal supply curves), all the data from hourly supply curves up to day D is available to forecast the hourly supply curves for day $D+1$ before the closing of MGP auction for day $D+1$.

Consequently, the setup for the real case study is as follows. Supply curves for the competitors of Enel, a major electricity company in Italy, are obtained by aggregating all the zonal competitors' supply curves. For these study, supply curves are limited to the price range $[0, 200]$ €/MWh, as in Figure 1. Therefore, each functional observation is the aggregated supply curve submitted to the market in each hour. The time range of the data is from 01/03/2015 up to 29/02/2016, thus consisting of 8784 curves. The data has been obtained from the Italian Electricity Market Operator (`www.mercatoelettrico.org`) and is divided into two sets: The In-Sample period is considered from 01/03/2015 to 31/08/2015, which will be used for training the models. The remaining data is left for the Out-Of-Sample period.

The offering behavior of the agents is conditioned to the weather and the particular circumstances of each day. Consequently, explanatory variables are used as a way to account for the external factors that might influence the traders decision. The explanatory variables used in this study are the following:

- Total electricity demand of Italy. The demand is of utmost importance to account for the consumption of energy in the country.
- Total wind power production. The south of Italy hosts a great number of wind farms which have a significant impact on the supply curves in windy days.
- Total solar production. The north of Italy is the region with the greatest solar power capacity installed. Therefore, the solar production should be significant.
- Thermal availability. It is the sum of all the energy offered in the market by thermal units.
- Energy exchanged from Italy to to the adjacent European countries: France, Switzerland, Austria, Slovenia, Greece and Malta. These exchanges play a very important role in the energy trading of the country.

It is worth noting that real values, instead of forecasts, of the explanatory variables are used for both training and validation of all the models considered in this study. In real operation the actual values of these variables are unknown, so future scenarios of the exogenous variables have to be used instead. The Italian Market Operator publishes hourly forecasts of the total electricity demand and international exchanges for day $D + 1$ at day D, so the proposed models can use them to obtain estimations of the houry supply curves before the auction for day $D + 1$ closes. Other variables, such as the the thermal availability and solar and wind production should be forecasted in order to use them as inputs for the forecasting models. These models provide useful indicators for any market agent, hence in this study it is assumed that market participants already have forecasts of these variables. Historical information about the aforementioned variables can be obtained from the Transparency Platform of ENTSO-E, the European Network of Transmission System Operators (`https://transparency.entsoe.eu/`). As the output time series are hourly curves and these explanatory variables are hourly values, the model will consider them as scalar covariates and not functional covariates.

The remainder of the section is divided into two parts. Firstly, the fitting of the SARMAHX model is detailed, showing the identification process proposed in this work. Secondly, a comparison with other reference models is presented.

4.1. SARMAHX model identification

The proposed functional SARMAHX model is identified and trained with the In-Sample data of this case study. Both integral and concurrent operators are used. Firstly, the autocorrelation structure of the series of supply curves in the Italian electricity market is identified using the FACF and the proposed identification procedure shown in Fig. 4. The first graph in Fig. 5 shows the seasonal behavior of the series: a high autocorrelation on lags 24 and 168 indicates that the series has strong daily and weekly seasonalities. The decay pattern of the FACF indicates the presence of an auto-regressive part in the series. The serial autocorrelation is still present in the second graph of Fig. 5, that shows the autocorrelation structure of the residuals of a fitted SARMAHX$(0, 0) \times (0, 0)_{24} \times (0, 0)_{168}$ including exogenous variables. These explanatory variables are included to capture the effect of these exogenous terms in the supply curves. As the output time series are hourly curves and the explanatory variables are hourly values, the model will consider them as scalar covariates. The strong serial correlation in the regression residuals of the SARMAHX model indicates that the dynamic of the series has to be modeled in order to obtain accurate forecasts of our data.

A SARMAHX$(1, 0) \times (1, 0)_{24} \times (1, 0)_{168}$ model is fitted to model the auto-regressive part observed in the residuals of the regression model. The third graph in Fig. 5 shows the functional autocorrelation of the residuals of this SARMAHX$(1, 0) \times (1, 0)_{24} \times (1, 0)_{168}$ model. As can be seen, significant correlated lags are found at multiples of 24 and 168 hours. Hence, there are daily and weekly effects that have not been captured by the explanatory variables and the AR terms that should be modeled.

In order to select the optimal autoregressive and moving average orders of the SARMAHX model, the identification methodology presented in Section 3 was used. Regular and seasonal moving average terms have been added into the model to account for the significant correlated lags found in the previous FACF residual plot of the previous model. An iterative process has been carried out until the residuals of the fitted model are functional white noise. The final adjusted model for both concurrent and functional cases is a SARMAHX$(2, 0) \times (1, 2)_{24} \times (1, 2)_{168}$ model. The formulation of both models can be found in the Appendix. Throughout this paper, these models will be denoted by SARMAHX (functional case) and SARMAHX-conc (concurrent case).

The fourth image in in Fig. 5 shows the autocorrelation function of the residuals of the fitted concurrent

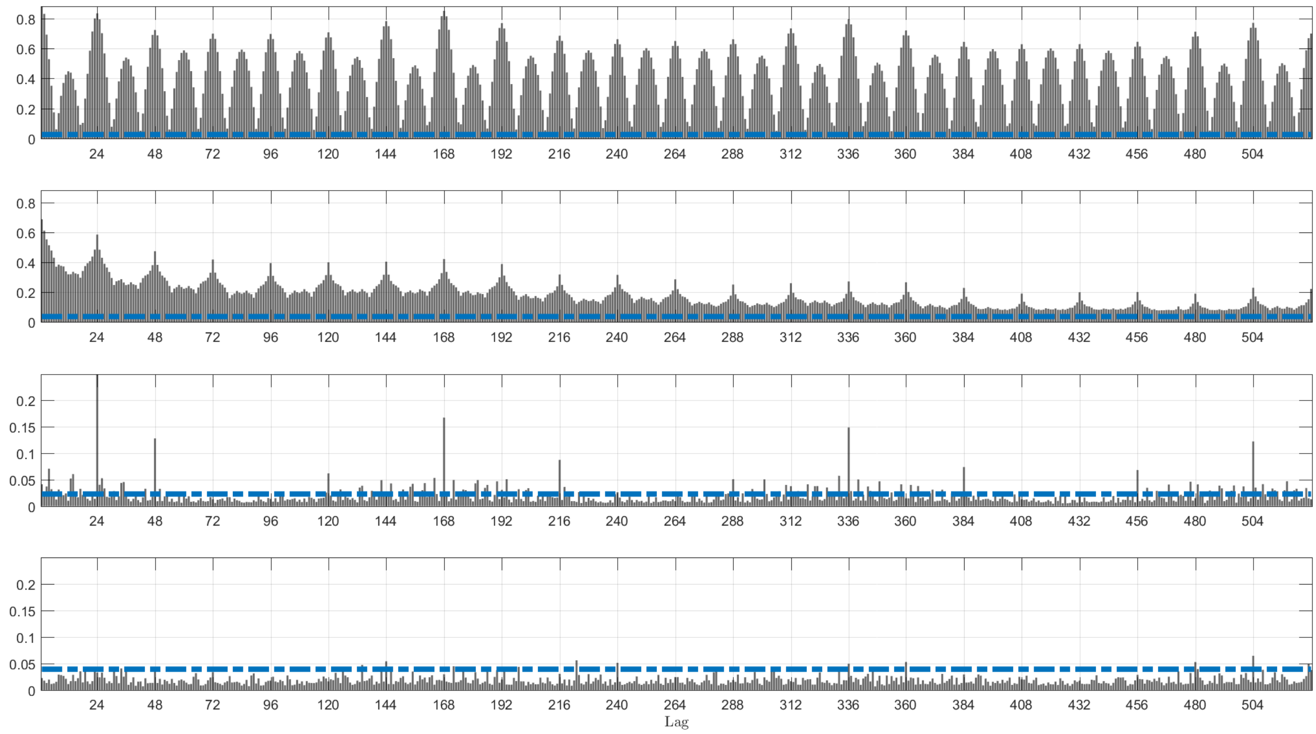

Figure 5: First 528 lags of the functional autocorrelation function (FACF). From top to bottom: (a) FACF of functional time series of Italian supply curves. (b) FACF of the regression residuals of a fitted SARMAHX$(0,0) \times (0,0)_{24} \times (0,0)_{168}$ including the aforementioned exogenous variables. (c) FACF of residuals of the fitted SARMAHX$(1,0) \times (1,0)_{24} \times (1,0)_{168}$ concurrent model including exogenous variables. A high autocorrelation value is found at multiples of 24 and 168 hours. (d) FACF of residuals of the fitted SARMAHX$(2,0) \times (1,2)_{24} \times (1,2)_{168}$ concurrent model including exogenous variables. As there are no significative autocorrelation values for any lags, the residuals can be viewed as a functional white noise process. In all images, the horizontal line represents the 95% upper limit of the distribution of the statistic under the white noise hypothesis.

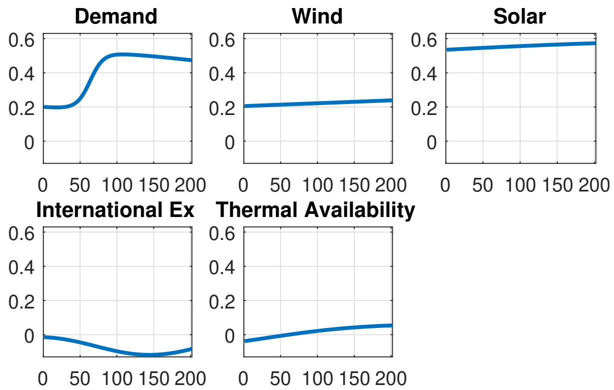

Figure 6: Operators' kernels for the regressors of the SARMAHX model in Supply Curve forecasting study.

SARMAHX model with the 95% upper bound of the statistics under the white noise hypothesis. As most values fall below this upper bound, the hypothesis of the residuals being white noise cannot be rejected, hence validating the model from a linear point of view.

The computation time for a sequential implementation of the SARMAHX estimation algorithm depends heavily on the size of the training dataset: with the configuration $(2,0) \times (1,2)_{24} \times (1,2)_{168}$ and a training dataset with 4416 curves discretized in 201 points; 500 iterations of the optimization algorithm took 3.027 hours when running on an i7-6700 CPU at 3.40 GHz with 32 GB of RAM. However, this process had to be carried only once: once fitted, the evaluation of the SARMAHX model is much faster, taking 6.52 seconds to obtain forecasts for the whole Out-Of-Sample period. In addition, this estimation algorithm could be improved by using parallel computation to speed up the whole process.

It is worth analyzing the resulting shapes of the fitted functional parameters. Fig. 6 shows the concurrent regressor operators. Analyzing these shapes, the effect of each variable on the offering curve can be seen. The demand has positive values for each price, meaning that for higher demand, the energy offered at all prices increases. In fact, the offer increases more at higher prices than at lower prices, meaning that the increase in demand is usually covered with more expensive generation. Wind and solar production have somewhat flat coefficients. Usually,

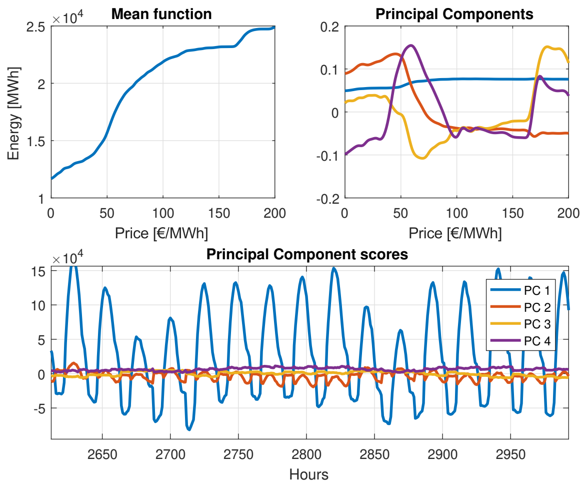

Figure 7: Mean function, Principal Components and scores of the supply curves.

these renewal production is offered at price 0, thus, simply displacing the curve up or down. The international exchange, on the other hand, has negative values. This means that when more energy is imported, the supply curve has less energy being offered.

4.2. Empirical comparison

This section compares the fitted SARMAHX models with some reference models. Two different analysis are presented. On the one hand, as the proposed model is trained to minimize the one-step ahead forecast error, a one-hour ahead forecast is analyzed. However, in the Italian electricity market, the auctions for the 24 hours of the day are cleared at the same time. Therefore, a 24-hour ahead forecast is also considered, validating the use in a real case application.

All the models included in this study were fitted using data from the In-Sample period. Once their parameters have been estimated, estimations are obtained for both the In-Sample and the Out-Of-Sample period, without re-calibrating the parameters of the models. For the 24-hour ahead forecast, a rolling window approach is used: when forecasting the first hour $h = 1$ of day $D+1$, the estimation $\widehat{Y}_{D+h}$ uses the real curves up to day D. However, when forecasting hours $h > 1$, the real curves have not been observed, so past estimations of the model are used as inputs for the model.

Following [43] and [28], forecasting models for functional time series can be classified in 3 main groups, according to the techniques that are used to obtain estimations of the parameters of the model: parametric, non-parametric and dimensionality reduction models. Parametric models, such as the proposed SARMAHX model, assume that the parameters can be expressed as integral operators. Conversely, non-parametric models do not define a fixed structure for the operators of the model, instead relying on kernel estimators or other non-parametric techniques to fit the parameters of the models. Finally, dimensionality reduction models pursue the goal of transforming the functional time series into a reduced set of scalar variables so that multivariate techniques can be applied to the estimation of the final model. In order to provide a fair comparison, one model for each group will be included in this study, to test the forecasting capabilities of the proposed model against the best performing functional models found in the literature. The models to be compared are described ahead:

- *Naïve.* Simple benchmark model that provides a reference to compare the tested models. Two versions are used, depending whether the simulation is one-step or 24-step forecast. In the first case, the forecast is simply the last curve observed in the data, i.e. the curve from the previous hour. In the second case, for Tuesdays, Wednesdays, Thursdays and Fridays, the forecast will be the hourly curve of the previous day, while for Saturdays, Sundays and Mondays, the forecast will be the hourly curve of the previous week.

- *Principal Components approach.* This method is used in [16] for supply curves forecasting. It extracts the first Functional Principal Components of the curves and the corresponding time series of scores. Then, the scores are forecasted by means of Transfer Function (TF) models [44], which include explanatory variables. The final estimation of the curves is done by reconstructing the estimated scores. The In-Sample period is used for training all models. FPCs are extracted for that range and the parameters of the TF are estimated. In the Out-Of-Sample period, the component scores are obtained, by projecting the curves into the basis spanned by the FPCs previously extracted.

  Fig. 7 represents the mean function, the principal components and the scores obtained for the curves time series. Three and four principal components are extracted, which explain the 98% and 99% of the variance of the data respectively. It is worth noting that each forecasted curve from this method is a linear combination of only the 3 or 4 principal component extracted. These models will be denoted as PC_FT3 and PC_FT4, respectively.

- *Functional nonparametric approach.* When forecasting functional time series, one of the most popular models found in the literature is the functional non-parametric model (the reader is referred to [43] for a comprehensive review of nonparametric techniques applied to functional data). The functional non-parametric model [45] assumes that the functional response Y_t can be modeled as a non-linear function of functional covariates X_t.

$$Y_t(v) = m(X_t)(v) + \varepsilon_t(v), \tag{23}$$

Table 1: Average errors for each method in the one-step and 24-step ahead prediction in Supply Curve forecasting study. The lower errors obtained for each forecasting horizon are marked in bold.

Forecasting horizon	Model	In-Sample FMAE [MWh]	In-Sample FRMSE [MWh]	Out-Of-Sample FMAE [MWh]	Out-Of-Sample FRMSE [MWh]
1-Step Ahead	Naïve	909.39	1258.46	818.59	1179.16
	PC_FT3	353.85	470.51	480.54	694.15
	PC_FT4	276.65	378.83	416.05	569.36
	NPARHX	389.30	553.53	641.92	884.50
	SARMAHX	264.62	363.10	313.11	429.66
	SARMAHXconc	**248.58**	**354.97**	**288.66**	**407.81**
24-Step Ahead	Naïve	1075.3	1433.5	1382.3	1812.50
	PC_FT3	672.06	871.33	856.62	1148.1
	PC_FT4	639.46	834.46	811.69	1085.07
	NPARHX	743.47	1010.01	987.41	1311.51
	SARMAHX	590.08	787.34	709.70	963.49
	SARMAHXconc	**557.37**	**764.73**	**655.19**	**885.86**

Table 2: Diebold-Mariano test's p-values for Out-Of-Sample period in the one-step ahead forecasts in Supply Curve forecasting study. p-values that are lower than 0.05 are marked in bold.

Models	Naïve	PC_FT3	PC_FT4	NPARHX	SARMAHX	SARMAHXconc
Naïve	-					
PC_FT3	**0**	-				
PC_FT4	**0**	**0**	-			
NPARHX	**0**	**0**	**0**	-		
SARMAHX	**0**	**0**	**0**	**0**	-	
SARMAHXconc	**0**	**0**	**0**	**0**	**0**	-

Table 3: Diebold-Mariano test's p-values for Out-Of-Sample period in the 24-step ahead forecasts in Supply Curve forecasting study. p-values that are lower than 0.05 are marked in bold

Models	Naïve	PC_FT3	PC_FT4	NPARHX	SARMAHX	SARMAHXconc
Naïve	-					
PC_FT3	**0**	-				
PC_FT4	**0**	**0**	-			
NPARHX	**0**	**0.001**	**0**	-		
SARMAHX	**0**	**0.0001**	0.1162	-		
SARMAHXconc	**0**	**0**	**0**	**0**	**0**	-

where $m(.)$ is an unknown nonlinear operator and ε_t is a functional white noise process. Nonparametric models do not define a fixed structure for the operator $m(.)$, instead they employ a functional version of the Nadaraya-Watson kernel estimator to estimate the regression operator:

$$\widehat{m}_h(X_t) = \frac{\sum_{i=1}^{T} Y_i K\left(h^{-1} d(X_t, X_i)\right)}{\sum_{i=1}^{T} K\left(h^{-1} d(X_t, X_i)\right)}, \tag{24}$$

where K is a kernel function, $h > 0$ is a bandwidth parameter and d is a semimetric. A nonparametric version of the functional autoregressive model of order 1 can be defined with this model, using past realizations of the process as the functional covariate:

$$Y_t(v) = m(Y_{t-1})(v) + \varepsilon_t(v). \tag{25}$$

This model was used in [19] to obtain forecasts of supply and purchase curves in the Italian Day-Ahead market in order to obtain accurate estimations of the market price. However, as all the models that will be compared against the proposed SARMAHX model use exogenous variables, a variation of this model will be included in the comparison study instead. The semi-functional partial lineal model [24, 25] allows the inclusion of exogenous scalar variables, generalizing the functional nonparametric model (25). The expression of the model is as follows:

$$Y_t(v) = \boldsymbol{x}_t^{\intercal} \boldsymbol{\beta}(v) + m(Y_{t-1})(v) + \varepsilon_t(v) \tag{26}$$

where $\boldsymbol{x}_t^{\intercal} = (x_{t,1}, \ldots, x_{t,p})$ is a vector of p exogenous scalar covariates and $\boldsymbol{\beta}(v) = (\beta_1(v), \ldots, \beta_p(v))$ is a vector of unknown functional parameters to be estimated. The authors obtain estimations for these parameters using kernel smoothing and ordinary least squares ideas. In order to fit this model to the supply curve dataset, the parameters have been selected as follows: the kernel function K used was the Epanechnikov, defined as $K(u) = 3/4(1-u^2)$; the bandwidth parameter h has been selected using the k-nearest neighbours method proposed in [46]; and the semimetric d selected is based on the L^2 norm of the curves. Throughout this paper, this model will be denoted as NPARHX.

All the models are trained with the In-Sample period and using the same exogenous variables. Then, each one produces a 1-step ahead forecast, assuming that the curve of last hour is known, and a 24-step ahead forecast, where the 24 hours of the following day are forecasted being the last observed curve the hour 24 of the current day. Henceforth, 1-step and 24-step ahead forecasts are produced for the whole data range. Functional errors FMAE and FRMSE are calculated, which are defined as

$$\text{FMAE} = T^{-1} \sum_{t=1}^{T} \int |Y_t(u) - \hat{Y}_t(u)| du \tag{27}$$

$$\text{FRMSE} = \sqrt{T^{-1} \sum_{t=1}^{T} \int \left(Y_t(u) - \hat{Y}_t(u) \right)^2 du} \tag{28}$$

Table 1 shows the functional errors for the 1-step and the 24-step ahead forecasts for both the In-Sample and Out-Of-Sample periods. The proposed SARMAHX model shows a clear advantage over the reference models, obtaining a FMAE of 288.66 MWh that is much lower than the FMAE of the PC_FT4 (the best reference model), which achieves a FMAE of 416.05 MWh in the Out-Of-Sample period of the 1-step ahead forecasting study. For the 24-step ahead forecasting case, the results are similar: while the PC_FT4 model obtains a FMAE of 811.69 MWh, it is not able to improve the results of the concurrent SARMAHX model, that achieves a FMAE of 655.19 MWh in the Out-Of-Sample period. The estimation results are analyzed ahead in detail.

Firstly, looking at the 1-step ahead forecasts it can be seen how the functional approach outperforms the other methods in both the In-Sample and Out-Of-Sample periods. Moreover, the concurrent functional approach is the one that provides the best results, with a mean absolute error of 289 MWh in the Out-of-Sample period, whereas the FMAE of the PC_FT4 model is 416 MWh. This difference validates the usefulness of the SARMAHX model as a competitive model to forecast supply curves. Several conclusions can be drawn from these results. Firstly, even though FPCA approach also uses a time series model, the fact of reducing the dimensionality significantly affects the performance. As the Principal Components are kept untouched, the forecasts cannot adapt to changes in the functional time series and the reconstruction of the curves looses precision. On the contrary, the proposed functional model does not rely on any basis expansion of the series and it takes into account the whole curve values from the recent past. Therefore, it can better account for changes in the series. In addition, as the nonparametric approach does not take into account the seasonality structure of the data, it is not able to achieve the same results as the other reference models. Finally, the use of concurrent operators in the SARMAHX model returns better results than integral operators. This means that the cross correlations between supply curves might not be very important, hence, the adjustment of the parameters might be easier for the concurrent model. Fig. 8 shows some forecasting examples for the one step-ahead case, where the capability of the concurrent SARMAHX model of estimating the shape of the supply curves is highlighted. Whereas the proposed model is able to accurately estimate the supply curves, the reference models are not capable of capturing the complex bidding behavior exhibited by the supply curves, providing only a smooth approximation to the curves that does not take into account the step-wise nature of the curves.

Secondly, the 24-step forecasts show similar results to the 1-step ahead case. The concurrent SARMAHX model provides better average errors with respect to the other models. While the PC_FT4 model has a mean absolute error of 811 MWh for the Out-Of-Sample period, the concurrent SARMAHX model has a much lower mean absolute error of 655 MWh, proving itself to be a competitive forecasting model. Globally, 24-step ahead errors are much higher than 1-step ahead errors, i.e. around 4% error for the 24-step and around 2% error for the 1-step. Supply curves are the result of aggregating the offers of competitors in the market. These offers respond to the different strategies of the agents which, in many cases, have a common behavior for the whole day. This is the case if there is some technical constraint on a power plant for that day. Therefore, strategies are planed for the whole day, and consequently, once the curve for the first hour is known, it helps to improve significantly the forecast of the following hours. This can be seen in Fig. 10. Forecasting the first hour is very important to the rest of the prediction, specially for the first hours of the day. Fig. 9 shows some forecasting examples for the 24-step case. As can be seen, the concurrent SARMAHX model provides better estimations of the supply curves than the other models, obtaining a much lower error than the reference models.

Additionally, Tables 2 and 3 show the Diebold-Mariano (DM) [47] test applied to each pair of forecasts. This test is a head-to-head method which compares the forecast error of two different models. Each cell of Tables 2 and 3 contains the p-value that results from comparing the model of the cell's row against the model of the cell's column. A low p-value means that the null hypothesis has to be rejected and therefore the forecasting errors can be considered statistically different. The DM test validates the Out-Sample results, yielding very low p-values between model comparisons. Then, the equal forecast null hypothesis can be rejected. This analysis emphasizes the statistical differ-

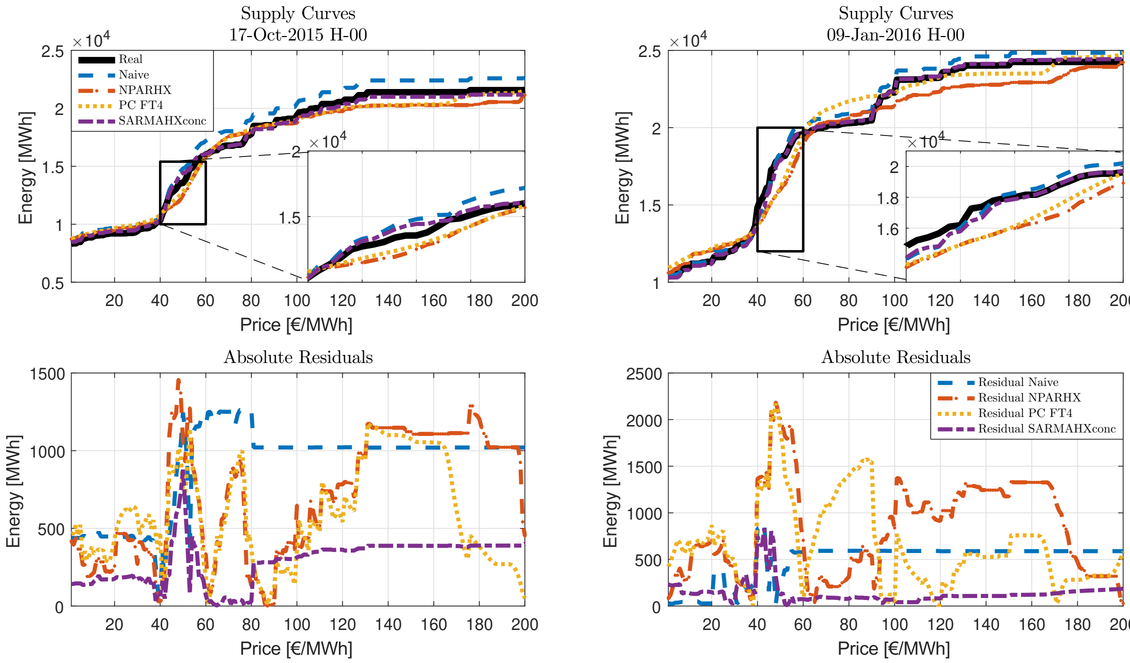

Figure 8: Top: Forecast examples for 1-step ahead estimations in the Out-Of-Sample periods in Supply Curve forecasting study. Curves in the price range $[40, 60]$ €/MWh are shown in detail, highlighting the differences between the different forecasting models. Bottom: Absolute value of the residual curves of the models.

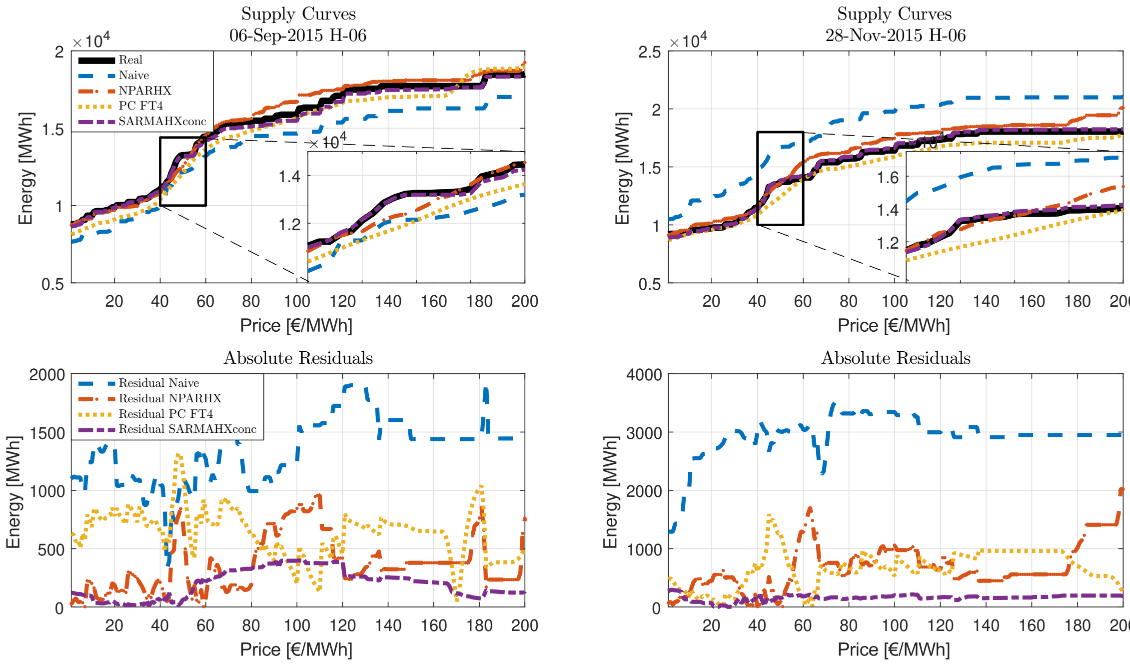

Figure 9: Top: Forecast examples for 24-step ahead estimations in the Out-Of-Sample periods in Supply Curve forecasting study. Curves in the price range $[40, 60]$ €/MWh are shown in detail, highlighting the differences between the different forecasting models. Bottom: Absolute value of the residual curves of the models.

ence between the proposed model and the other models considered in this study.

4.3. In-depth analysis of the results

In the Italian Day-Ahead market, the market-clearing price is usually located in the price range $[30, 70]$ €/MWh [19], hence the volume of the agents' bids will be significantly higher in that portion of the aggregated supply curve. Thus, obtaining accurate forecast of the supply curve in that region is of utmost importance for any company. As can be seen in both Fig. 8 and Fig. 9, the concurrent SARMAHX model is able to capture the shape

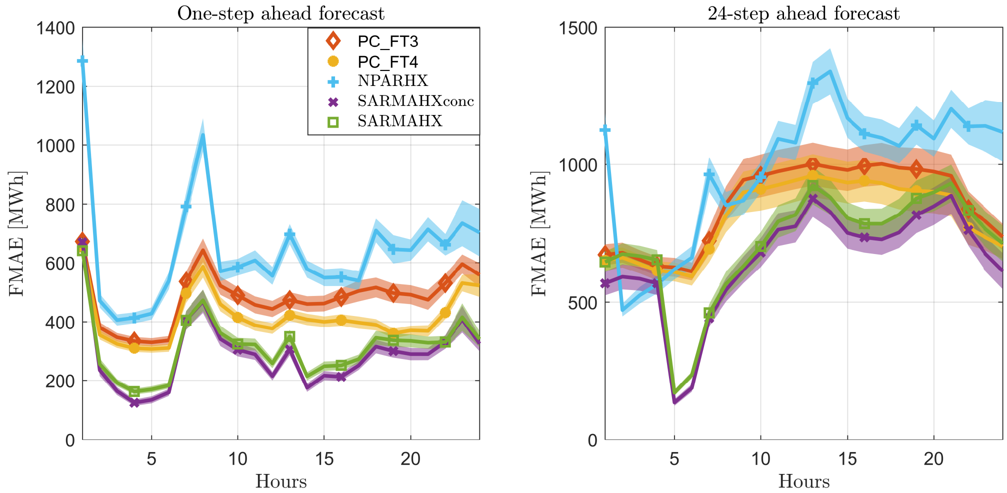

Figure 10: FMAE for each hour in the 1-step and 24-step ahead forecasts in the Out-Of-Sample period in Supply Curve forecasting study. As the FMAE error of the Naive model is significantly higher than the other models, it is not included. The shaded regions are 90% confidence bands for the FMAE.

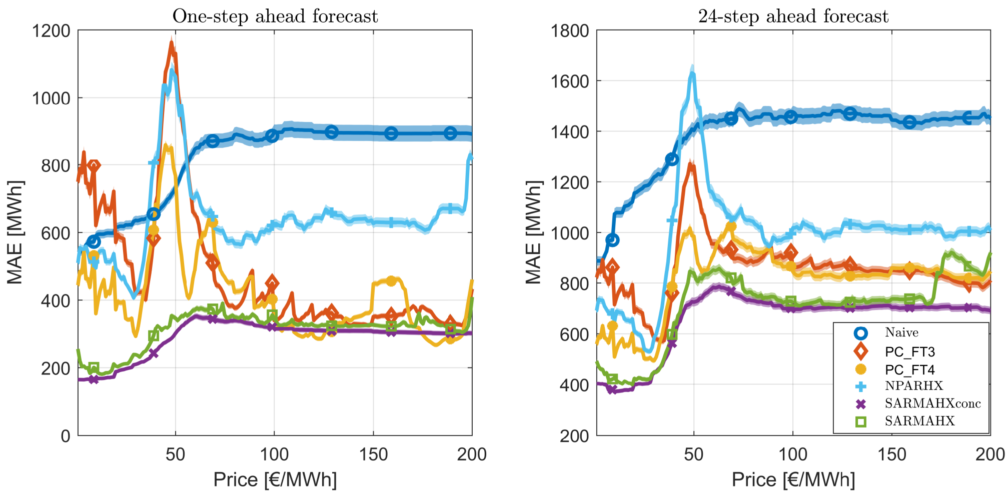

Figure 11: Mean absolute error for each price in the 1-step and 24-step ahead forecasts in the Out-Of-Sample period in Supply Curve forecasting study. The shaded regions are 90% confidence bands for the MAE.

of the supply curve in that price range, providing an accurate description of the agents' bidding behavior in the zone of interest of the supply curve.

The shape of the estimated supply curves is essential when analyzing the strategic offering of the competitors of an electricity company, because it contains valuable information about the offering behavior of the agents. Forecasted supply curves, together with estimations of the demand, can be used to define RDCs (1), which are often used to calibrate market equilibrium models [48] as well as offering optimization models [9]. In [49], the authors analyze the importance of obtaining faithful estimations of the slopes of RDCs, since it indicates an agent's capability to influence market prices. This highlights the need for a forecasting model that not only captures the general level of the supply curves, but also provides an accurate estimation of the shape of the curve.

As can be seen in Fig. 11, MAE errors are presented for each price for the Out-Of-Sample period and for the 1-step and 24-step ahead forecasts, which are obtained as the mean value of the absolute error for each price of the curves. The figures show the performance of each method for different ranges of prices. It can be seen how the the SARMAHX models yield smoother errors across prices. Again, as the proposed models do not depend on the FPCs, it can adapt better to the bidding steps in the curve. The most significant improvement is seen in the aforementioned price range of interest: for example, whereas the mean absolute error of the PC_FT4 model at price 50 €/MWh is 1000 MWh, the concurrent SARMAHX model achieves a mean absolute error of 730 MWh in the 24-step ahead forecast. This difference is more pronounced in the one-step ahead study, where the mean absolute errors of the PC_FT4 model in the price range of interest are close to 800 MWh as opposed to the errors of the SARMAHX model, which are near 350 MWh. In addition, confidence bands for the MAE errors are included to ease the interpretation of the results. As the confidence bands do not overlap, it can be concluded that the errors of the SARMAHX models are significantly lower than the other models considered in this study.

5. Conclusions

This paper has been devoted to the empirical application of functional models to the forecasting of supply curves in electricity markets. Electricity markets provide a suitable framework for the use of this new statistical approach.

This manuscript contributes significantly by expanding the functional SARMAHX model to allow the inclusion of up to two seasonalities and scalar covariates, allowing the modelization of hourly functional time series by considering their daily and weekly seasonalities. This development has lead to more accurate forecasts of the aggregated supply curves of the system, outperforming other well known existing methods for estimating supply curves.

In addition, the identification procedure based on the functional autocorrelation function has proven to be crucial for analyzing the correlation structure of the functional time series and for selecting the correct AR and MA order of the model.

The performance of the model has been validated with a real forecasting case of hourly supply curves in the day-ahead Italian electricity market. It has been shown how the proposed model is a competitive model against other well known existing methods for supply curve forecasting for both 1-step and 24-step ahead forecasting.

The proposed model and identification methodology presented in this paper are not restricted to supply curves alone. These techniques can be extended to other functional time series found in electricity markets or in other fields. Some examples could be residual demand curves forecasting or intraday continuous market price forecasting in power systems or forward curves in financial markets. One of such applications could be the estimation of electricity prices as a result of forecasting both the supply and the demand curves of the system. Finally, additional research could be done for measuring the importance of the explanatory variables included in the model, the inclusion of dummy intervention variables or the usage of an optimization algorithm that trains the model to minimize the 24-step ahead forecasts instead of the one-step ahead error.

6. Appendix

This appendix contains the formulation of the final SARMAHX$(2,0)\times(1,2)_{24}\times(1,2)_{168}$ models that were fitted in Section 4:

$$\begin{aligned}&\big(I-\Psi_{0,1}B^1-\Psi_{0,2}B^2\big)\big(I-\Psi_{1,1}B^{24}\big)\big(I-\Psi_{2,1}B^{168}\big)(Y_t)=\\&\big(I-\Theta_{1,1}B^{24}-\Theta_{1,2}B^{48}\big)\big(I-\Theta_{2,1}B^{168}-\Theta_{2,2}B^{336}\big)(\varepsilon_t)\\&+\Gamma^c_{\mathrm{D}}x^{\mathrm{D}}_t+\Gamma^c_{\mathrm{W}}x^{\mathrm{W}}_t+\Gamma^c_{\mathrm{S}}x^{\mathrm{S}}_t+\Gamma^c_{\mathrm{E}}x^{\mathrm{E}}_t+\Gamma^c_{\mathrm{T}}x^{\mathrm{T}}_t,\end{aligned}\tag{29}$$

where $\Psi_{j,i}$ are autoregressive operators, $\Theta_{k,l}$ are moving average operators, Γ^c_Z are regression integral operators and x^Z_t are the scalar covariates (D stands for demand, W stands for wind production, S stands for solar production, E stands for international exchanges and T stands for thermal availability). In the case of the SARMAHXconc model, the autoregressive and moving average operators of (29) are concurrent operators (11).

Acknowledgments

The authors would like to thank the editor and the anonymous reviewers for their valuable comments which helped in improving the paper.

References

[1] R. Weron, Modeling and Forecasting Electricity Loads and Prices. A Statistical Approach, Wiley, 2006.

[2] D. Bunn, E. D. Farmer, Comparative models for electrical load forecasting, John Wiley and Sons Inc., New York, NY, 1985.

[3] R. Weron, Electricity price forecasting: A review of the state-of-the-art with a look into the future, International Journal of Forecasting 30 (4) (2014) 1030–1081. doi:10.1016/j.ijforecast.2014.08.008.

[4] K. Chen, K. Chen, Q. Wang, Z. He, J. Hu, J. He, Short-Term Load Forecasting With Deep Residual Networks, IEEE Transactions on Smart Grid 10 (4) (2019) 3943–3952. doi:10.1109/TSG.2018.2844307.

[5] J.-L. Zhang, Y.-J. Zhang, D.-Z. Li, Z.-F. Tan, J.-F. Ji, Forecasting day-ahead electricity prices using a new integrated model, International Journal of Electrical Power & Energy Systems 105 (2019) 541–548. doi:10.1016/j.ijepes.2018.08.025.

[6] C. Monteiro, I. J. Ramirez-Rosado, L. A. Fernandez-Jimenez, M. Ribeiro, New probabilistic price forecasting models: Application to the Iberian electricity market, International Journal of Electrical Power & Energy Systems 103 (2018) 483–496. doi:10.1016/j.ijepes.2018.06.005.

[7] Y. Wang, Q. Hu, D. Srinivasan, Z. Wang, Wind Power Curve Modeling and Wind Power Forecasting With Inconsistent Data, IEEE Transactions on Sustainable Energy 10 (1) (2019) 16–25. doi:10.1109/TSTE.2018.2820198.

[8] Z. Liu, J. Yan, Y. Shi, K. Zhu, G. Pu, Multi-agent based experimental analysis on bidding mechanism in electricity auction markets, International Journal of Electrical Power & Energy Systems 43 (1) (2012) 696–702. doi:10.1016/j.ijepes.2012.05.056.

[9] A. Baillo, M. Ventosa, M. Rivier, A. Ramos, Optimal offering strategies for generation companies operating in electricity spot markets, IEEE Transactions on Power Systems 19 (2) (2004) 745–753. doi:10.1109/TPWRS.2003.821429.

[10] F. A. Campos, A. Muñoz, E. F. Sánchez-Úbeda, J. Portela, Strategic Bidding in Secondary Reserve Markets, IEEE Transactions on Power Systems 31 (4) (2016) 2847–2856. doi:10.1109/TPWRS.2015.2453477.

[11] L. Xu, R. Baldick, Transmission-Constrained Residual Demand Derivative in Electricity Markets, IEEE Transactions on Power Systems 22 (4) (2007) 1563–1573. doi:10.1109/TPWRS.2007.907511.

[12] C. L. Pretea, B. F. Hobbs, Market power in power markets: an analysis of residual demand curves in California's day-ahead energy market (1998-2000), The Energy Journal 36 (2). doi:10.5547/01956574.36.2.9.

[13] G. Aneiros, J. Vilar, P. Raña, Short-term forecast of daily curves of electricity demand and price, International Journal of Electrical Power & Energy Systems 80 (2016) 96–108. doi:10.1016/j.ijepes.2016.01.034.

[14] S. J. Huang, K. R. Shih, Short-term load forecasting via ARMA model identification including non-Gaussian process considerations, IEEE Transactions on Power Systems 18 (2) (2003) 673–679. doi:10.1109/TPWRS.2003.811010.

[15] S. Fan, R. J. Hyndman, Short-Term Load Forecasting Based on a Semi-Parametric Additive Model, IEEE Transactions on Power Systems 27 (1) (2012) 134–141. doi:10.1109/TPWRS.2011.2162082.

[16] M. Pelagatti, Supply function prediction in electricity auctions, in: Complex Models and Computational Methods in Statistics, Springer, 2013, pp. 203–213.

[17] F. Ziel, R. Steinert, Electricity price forecasting using sale and purchase curves: The X-Model, Energy Economics 59 (2016) 435–454. doi:10.1016/j.eneco.2016.08.008.

[18] F. Ziel, R. Steinert, Probabilistic mid- and long-term electricity price forecasting, Renewable and Sustainable Energy Reviews 94 (2018) 251–266. doi:10.1016/j.rser.2018.05.038.

[19] I. Shah, F. Lisi, Forecasting of electricity price through a functional prediction of sale and purchase curves, Journal of Forecasting (2019) 1– 18doi:10.1002/for.2624.

[20] J. Ramsay, B. W. Silverman, Functional Data Analysis, Springer Series in Statistics, Springer-Verlag New York, 2005.

[21] L. Horváth, P. Kokoszka, Inference for functional data with applications, Vol. 200, Springer, 2012.

[22] A. Antoniadis, X. Brossat, J. Cugliari, J.-M. Poggi, A prediction interval for a function-valued forecast model: Application to load forecasting, International Journal of Forecasting 32 (3) (2016) 939–947. doi:10.1016/j.ijforecast.2015.09.001.

[23] B. L. Cabrera, F. Schulz, Forecasting Generalized Quantiles of Electricity Demand: A Functional Data Approach, Journal of the American Statistical Association 112 (517) (2017) 127–136. doi:10.1080/01621459.2016.1219259.

[24] G. Aneiros, J. M. Vilar, R. Cao, A. Muñoz, Functional prediction for the residual demand in electricity spot markets, IEEE Transactions on Power Systems 28 (4) (2013) 4201–4208. doi:10.1109/TPWRS.2013.2258690.

[25] J. Vilar, G. Aneiros, P. Raña, Prediction intervals for electricity demand and price using functional data, International Journal of Electrical Power & Energy Systems 96 (2018) 457–472. doi:10.1016/j.ijepes.2017.10.010.

[26] J. M. Vilar, R. Cao, G. Aneiros, Forecasting next-day electricity demand and price using nonparametric functional methods, International Journal of Electrical Power & Energy Systems 39 (1) (2012) 48–55. doi:10.1016/j.ijepes.2012.01.004.

[27] D. Liebl, Modeling and forecasting electricity spot prices: A functional data perspective, The Annals of Applied Statistics 7 (3) (2013) 1562–1592. doi:10.1214/13-AOAS652.

[28] J. Portela, A. Muñoz, E. Alonso, Forecasting Functional Time Series with a New Hilbertian ARMAX Model: Application to Electricity Price Forecasting, IEEE Transactions on Power Systems 33 (1) (2018) 545–556. doi:10.1109/TPWRS.2017.2700287.

[29] G. E. P. Box, G. M. Jenkins, Time series analysis, forecasting and control, Holden-Day, San Francisco, CA, USA., 1970.

[30] D. Bosq, Linear processes in function spaces: theory and applications, Vol. 149, Springer Verlag, 2000.

[31] D. Didericksen, P. Kokoszka, X. Zhang, Empirical properties of forecasts with the functional autoregressive model, Computational Statistics (2012) 285–298doi:10.1007/s00180-011-0256-2.

[32] G. Cybenko, Approximation by superpositions of a sigmoidal function, Mathematics of control, signals and systems 2 (4) (1989) 303–314. doi:10.1007/BF02551274.

[33] R. J. Bhansali, Order selection for linear time series models: a review, in: Developments in time series analysis, Chapman and Hall, London, 1993, pp. 50–66.

[34] P. J. Brockwell, R. A. Davis, Time Series: theory and methods, 2nd Edition, Springer series in statistics, Springer, New York, NY, 2006.

[35] P. Kokoszka, M. Reimherr, Determining the order of the functional autoregressive model, Journal of Time Series Analysis 34 (1) (2013) 116–129. doi:10.1111/j.1467-9892.2012.00816.x.

[36] X. Zhang, White noise testing and model diagnostic checking for functional time series, Journal of Econometrics 194 (1) (2016) 76–95. doi:10.1016/j.jeconom.2016.04.004.

[37] P. Bagchi, V. Characiejus, H. Dette, A Simple Test for White Noise in Functional Time Series, Journal of Time Series Analysis 39 (1) (2018) 54–74. doi:10.1111/jtsa.12264.

[38] G. Mestre, J. Portela, G. Rice, A. Muñoz, E. Alonso, Functional time series identification and diagnosis by means of autocorrelation analysis, Working Paper IIT-18-102A (Aug. 2018).

[39] P. Kokoszka, G. Rice, H. L. Shang, Inference for the autocovariance of a functional time series under conditional heteroscedasticity, Journal of Multivariate Analysis 162 (2017) 32–50. doi:10.1016/j.jmva.2017.08.004.

[40] W. W. Wei, Time Series Analysis : Univariate and Multivariate Methods, 2nd Edition, Addison Wesley, 2005.

[41] R. J. Hyndman, G. Athanasopoulos, Forecasting: Principles and Practice, 2nd Edition, Otexts, 2018.

[42] A. Gianfreda, L. Grossi, Forecasting Italian electricity zonal prices with exogenous variables, Energy Economics 34 (6) (2012) 2228–2239. doi:10.1016/j.eneco.2012.06.024.

[43] F. Ferraty, P. Vieu, Nonparametric Functional Data Analysis: Theory And Practice, Springer, 2006.

[44] A. Pankratz, A Primer on ARIMA Models, Wiley Online Library, 1991.

[45] F. Ferraty, I. Van Keilegom, P. Vieu, Regression when both response and predictor are functions, Journal of Multivariate Analysis 109 (2012) 10–28. doi:10.1016/j.jmva.2012.02.008.

[46] A. Antoniadis, E. Paparoditis, T. Sapatinas, Bandwidth selection for functional time series prediction, Statistics & Probability Letters 79 (6) (2009) 733–740. doi:10.1016/j.spl.2008.10.028.

[47] F. X. Diebold, R. S. Mariano, Comparing predictive accuracy, Journal of Business & Economic Statistics 13 (3) (1995) 253–263. doi:10.1080/07350015.1995.10524599.

[48] C. A. Díaz, J. Villar, F. A. Campos, J. Reneses, Electricity market equilibrium based on conjectural variations, Electric Power Systems Research 80 (12) (2010) 1572–1579. doi:10.1016/j.epsr.2010.07.012.

[49] J. Portela, A. Muñoz, E. F. Sánchez-Úbeda, J. García-González, R. González, Residual demand curves for modeling the effect of complex offering conditions on day-ahead electricity markets, IEEE Transactions on Power Systems 32 (1) (2017) 50–61. doi:10.1109/TPWRS.2016.2552240.