Electrical Power and Energy Systems 96 (2018) 457–472

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Prediction intervals for electricity demand and price using functional data

Juan Vilar*, Germán Aneiros, Paula Raña

Departamento de Matemáticas, Facultad de Informática, Universidad de A Coruña, Campus de Elviña, s/n, 15071 A Coruña, Spain

A R T I C L E I N F O

Keywords:
Load and price
Electricity markets
Functional data
Functional time series forecasting
Prediction intervals

A B S T R A C T

This paper provides two procedures to obtain prediction intervals for electricity demand and price based on functional data. The proposed procedures are related to one day ahead pointwise forecast. In particular, the first method uses a nonparametric autoregressive model and the second one uses a partial linear semi-parametric model, in which exogenous scalar covariates are incorporated in a linear way. In both cases, the proposed procedures for the construction of the prediction intervals use residual-based bootstrap algorithms, which allows also to obtain estimates of the prediction density. Applications to the Spanish Electricity Market, in year 2012, are reported. This work extends and complements the results of Aneiros et al. (2016), focused on pointwise forecasts of next-day electricity demand and price daily curves.

1. Introduction

Nowadays, in many countries all over the world, the production and sale of electricity is traded under competitive rules in free markets. Thus, the electricity demand and price forecasting is a main target for the agents and companies involved in the electricity markets. In particular, short term forecast, which is the one day ahead hourly forecast, has been extensively studied in the literature. A wide range of methodologies and models for forecasting have been proposed and studied. These methods can be classified into two large groups: the first one based on statistical approaches including time series, dynamic regression, exponential smoothing, regression analysis, etc. and the second one based on artificial intelligence techniques such as neuronal networks, fuzzy neural networks, support vector machines, etc. A nice monograph on electricity demand and price forecasting can be found in [1]. Also [2,3] contain reviews on electricity demand forecasting and [4,5] on electricity price forecasting. In recent years, one can find some studies addressing this problem from a functional perspective, this is, defining the daily curves of electricity demand or price by some functional form (functional data). The books [6,7] are comprehensive references for functional data analysis. See also [8] for a recent monograph on inference for functional data. Statistical models for functional data were also used to predict electricity demand and price. The reader will find in [9] a parametric model to predict electricity consumption curves; [10] introduced a novel functional time series methodology that is applied to historical daily curves of load; [11] proposed a hybrid approach which was applied to French demand curves; [12] proposed a new methodology to obtain probabilistic forecasts of electricity load that is based on functional data analysis of generalized quantile curves; [13] studied the short term forecasting of household-level intra-day electricity load curve (they proposed a nonparametric functional approach based on functional kernel regression estimator with the use of an unsupervised clustering step of the historical segments); [14] used three approaches, two of them of functional type, to forecast the France's daily electricity load consumption; [15] proposed an adaptive functional autoregressive (AFAR) forecast model to predict electricity price curves; finally, the case of residual demand curves in Spain was analysed in [16]. When the interest is to forecast scalar values (not curves) from functional data, the reader can see [17] (nonparametric and semiparametric models), [18] (functional factor model) or [19] (censored response). The case of forecasting curves (as well as scalar values) of demand and price from functional data are studied in [20]. In particular, nonparametric and semi-functional partial linear models are analysed within the Spanish Electricity Market.

There are many methods and studies about short-term point forecast of electricity demand and price. However, methods for obtaining prediction intervals (PI) or prediction densities (PD) have not been studied extensively up to date. The importance of the PIs and PDs relies on the information that they provide, related to the evolution and variability of future demand or price, which allows to plan different strategies of action. Some papers on this topic are the following: [21–23] obtained PIs in the problem of electricity price prediction; [24–26] computed PIs in forecast of demand, and [27,28] studied methods of estimation of PD in this context.

This paper proposes two algorithms to obtain PIs and PDs in the problem of the next-day forecasting of electricity demand and price.

* Corresponding author.
E-mail addresses: juan.vilar@udc.es (J. Vilar), german.aneiros@udc.es (G. Aneiros), paula.rana@udc.es (P. Raña).

http://dx.doi.org/10.1016/j.ijepes.2017.10.010
Received 7 April 2017; Received in revised form 21 September 2017; Accepted 7 October 2017
0142-0615/

These algorithms are based on the use of functional nonparametric and semi-functional partial linear regression models, considering scalar response, to compute point predictions of demand or price (note that those prediction models have been studied in [17,20]). The approach in this paper uses functional data methods to take into account the daily seasonality of the electricity demand and price series. Bootstrap procedures are employed within the algorithms in order to avoid assumptions about the distribution of the data. Also, homoscedasticity and heteroscedasticity cases are considered. It is worth being noted which is the main difference between our proposal and the ones existing in the literature dealing with PIs and/or PDs from functional data: the used regression model and/or the implemented bootstrap procedure.

The proposed methods are applied in short-term forecasting of electricity demand and price in the Spanish Electricity Market. The accuracy of the PIs is evaluated by comparing their nominal coverages with the true coverage (unconditional coverages), and studying the length of PIs and the Winkler (or interval) score. The Christofferseńs test has been also applied. PIs computed from the empirical distribution of the residuals are obtained as benchmark method. A comparative study between the PIs from the proposed bootstrap procedure and the benchmark is performed.

The remaining of the paper is organized as follows. Section 2 presents the main features of electricity demand and price data corresponding to mainland Spain during the years 2011 and 2012, which will be used in our study. In addition, maximum daily temperature and daily wind power production will be used as exogenous scalar covariates. The forecasting methods involved in the study, functional nonparametric and semi-functional partial linear model, are presented in Section 3. The algorithms proposed to obtain PIs and PDs are described in Section 4. Such algorithms are applied in Section 5 to compute next-day PIs and PDs of the hourly values of electricity demand and price in each day of year 2012. Finally, Section 6 provides some relevant conclusions.

2. The data

The aim of this paper is to obtain PIs and PDs of the hourly electricity demand and price, one day-ahead, corresponding to the Spanish Electricity Market. Functional nonparametric and semi-parametric models, together with bootstrap techniques, will be considered. In order to present a complete study, we will compute PIs and PDs of the electricity demand and price for the 24 h of the day along all the days in 2012. The PIs and PDs for a fixed day are obtained using information from the previous 365 days and thus, our database contains information related to years 2011 and 2012. This database has been already used in [20] within the prediction study of the electric curves. Each daily functional datum comes from the 24 hourly observations of demand or price in each day (available at http://www.omel.es, the official website of OMIE, Operador del Mercado Ib érico de Energía. OMIE is the Market Operator which, together with System Operator REE, Red Eléctrica de España, compose the Regulators of the Electricity Sector in Spain). Smoothing techniques are used to convert the 24 hourly data in a functional observation. Fig. 1 displays the historical demand (MW h) and electricity price (Cents/kW h) for the whole year 2012 in Spain.

Main features of electricity demand can be summarized in the daily and weekly seasonality, the calendar effect on weekend and also the presence of outliers. It is easy to observe differences in the behaviour of the demand on the weekdays (from Monday to Friday), Saturday and Sunday, as it has been remarked in Fig. 1. This distinction is due to the decrease of electrical consumption, affecting the demand, during the weekend. For this reason, different models are used for each one of these three sets of days. Also the demand varies along the year possibly due to the different climate, among other reasons, resulting in some stable periods and others with more variability.

Electricity price shares some of the characteristics with the demand. However, it has some particular properties. The most notable one are the days with zero price. When the production of wind power increases, the price generally decreases, even reaching value zero. There are also some variations in the behaviour of the weekdays and the weekend, but maybe not as significant as in the demand. Daily curves for the electricity price remain in similar values along the year but, for example, the last quarter of the year corresponds to a turbulent period in the market in which the variation is higher and so, the forecasting becomes a quite difficult task.

Both demand and price present some outliers that have to be identify in order to prevent misleading predictions. For that purpose, we use some tools to detect outliers in functional time series, taking into account the dependence in the data, which is the case of the electricity data. Methods presented in [29,30] were applied to our data. [29] performs a test to determine if a curve is an outlier ot not, based on functional depths and bootstrap techniques. The methods in [30] consider robust principal components analysis, looking for outliers in the projections of the first principal component or analysing the residuals from some predictions. The identified outliers were replaced in the sample by weighted moving average of the surrounded days. Specifically, each identified outlier was changed by the weighted moving average of the four closest curves (in a temporal sense) in the same class that the identified outlier, the weights being 0.3 for the two closest and 0.2 for the other two.

There exist exogenous variables that could improve the predictions of the electricity demand and price and could be incorporated in the prediction model included in Section 3.2 (see [20] for details about the relation of the exogenous variables and the electrical data, which is summarize in the next paragraphs). For instance, the temperature influences the electrical demand, as one may observe a large demand of electrical heating when the weather is cold. For that reason, in this study we consider the maximum daily temperature (°C) in Spain. AEMET (Agencia Estatal de Meterología) provided us the maximum daily temperature for each province of the country and then, by population-weighted average, we built the corresponding maximum daily temperature for Spain. Population data were collected from INE (Instituto Nacional de Estadística). In our prediction methods (see Section 3) only exogenous variables with linear effect can be considered but the temperature has a nonlinear effect over the demand. The relation between the maximum daily temperature and the daily mean demand is U-shaped. Therefore, a transformation of the temperature data is needed. We have built two new variables, HDD (Heating Degree Days) and CDD (Cooling Degree Days), that are a measurement of the amount of energy needed to heat/cool a building. Specifically, they are defined as:

$$HDD(t) = \max\{20-T(t),0\} \quad \text{and}$$
$$CDD(t) = \max\{T(t)-24,0\},$$

where $T(t)$ is the maximum daily temperature in day t. See [31,20] for a justification of the values 20 and 24 in the functions HDD and CDD, respectively. Through this transformation, we obtain two variables that summarize all the temperature information with linear effect over the demand.

On the other hand, when the aim is to forecast the hourly electricity price, we will introduce as scalar covariates both the forecasted daily demand (trough the model in Section 3.2, including temperature information) and the wind power production (MW h). Our wind power data come from the System Operator in the Spanish Electricity Market, REE. In their web page, they monitor the demand and its generation structure at each moment. Then, we can know the amount of demand covered by wind power during each ten minutes period of the year. Therefore, we can calculate the corresponding value for each hour or the day. In this case, the relation between electricity price and the wind power production is linear: price decreases as wind power production increases.