European Journal of Operational Research 264 (2018) 149–164

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Production, Manufacturing and Logistics

Short-run electricity load forecasting with combinations of stationary wavelet transforms

Marie Bessec[a,*], Julien Fouquau[b]

[a] Université Paris Dauphine, PSL Research University, LEDa-CGEMP, Place du Maréchal de Lattre de Tassigny, 75016 Paris, France
[b] ESCP Europe, Labex Refi, 79 avenue de la République, 75011 Paris, France

ARTICLE INFO

Article history:
Received 9 July 2016
Accepted 20 May 2017
Available online 26 May 2017

Keywords:
Forecasting
Electricity demand
Seasonality
Wavelet transform
Combinations

ABSTRACT

Short-term forecasting of electricity load is an essential issue for the management of power systems and for energy trading. Specific modeling approaches are needed given the strong seasonality and volatility in load data. In this paper, we investigate the benefit of combining stationary wavelet transforms to produce one day-ahead forecasts of half-hourly electric load in France. First, we assess the advantage of decomposing the aggregate load into several subseries with a wavelet transform. Each component is predicted separately and aggregated to get the final forecast. One innovation of this paper is to propose several approaches to deal with the boundary problem which is particularly detrimental in electricity load forecasting. Second, we examine the benefit of combining forecasts over individual models. An extensive out-of-sample evaluation shows that a careful treatment of the border effect is required in the multiresolution analysis. Combinations including the wavelet predictions provide the most accurate forecasts. This result is valid with several assumptions about the forecast error in temperature and for different types of hours (peak, normal, off-peak), different days of the week and various forecasting periods.

1. Introduction

Accurate short-term forecasting of electricity load is crucial for the management of power systems and for energy trading. An underestimation or overestimation of the energy demand requires an additional production at high operational costs or leads to a waste of resources. Traders and brokers present in the market and purchasing electricity for resale also need accurate load forecasts to elaborate profitable bidding strategies. Forecast precision is particularly important in the French wholesale market. The third largest European electricity market is currently going through a drastic transformation to open up to more competition and, simultaneously, to a higher share of renewable energy sources. Our focus on the French market is the day-ahead forecast of electricity load.

Electricity load displays several well-known characteristics (see Weron, 2006). Load series are highly volatile with occasional large spikes. As underlined by Taylor (2010), they also exhibit multiple levels of seasonality: a daily cycle (overnight and afternoon troughs versus morning and evening peaks), a weekly cycle (strong increase on working days), a yearly cycle (much higher demand during the colder months of each year) as well as calendar effects (public holidays). Modeling electricity load is also complicated by the nonlinear link to meteorological variables (Bessec & Fouquau, 2008; Lee & Chiu, 2011). Electricity consumption is highly dependent on temperature, especially in France due to the large share of electric heating. There is a U-shaped relationship between electric power demand and temperature. In winter, higher temperatures diminish the need for heating, whereas they increase electricity consumption in summer due to the additional use of cooling devices.

Several approaches have been adopted in the literature to model electricity load (see Hahn, Meyer-Nieberg, and Pickl, 2009; Hong and Fan, 2016 for a survey). In the recent literature using time series techniques, many papers employ seasonal ARMA models and seasonal exponential smoothing. Dummy variables are often included for weekdays, weekends and special days with abnormally low demand (e.g. Cancelo, Espasa, & Grafe, 2008; Kim, 2013). At this level, electricity load is modeled either as a single time series (e.g. Amaral, Castro Souza, & Stevenson, 2008; Taylor, 2012) or in a multi-equation context (e.g. Clements, Hurn, & Li, 2016; Do, Lin, & Molnar, 2016) and different treatments of seasonality are applied in each case. The incorporation of temperature in the model as a determinant of electricity demand also helps to capture the intrayear cycle. To account for the nonlinear impact of temperature, several transformations of degree-day variables are applied

* Corresponding author.
E-mail addresses: marie.bessec@dauphine.fr (M. Bessec), jfouquau@escpeurope.eu (J. Fouquau).

http://dx.doi.org/10.1016/j.ejor.2017.05.037
0377-2217/

(see, among others, Cancelo, Espasa, & Grafe, 2008; Dordonnat, Koopman, & Ooms, 2012). Zivanovic (2001) employs polynomial regressions, while Moral-Carcedo and Vicéns-Otero (2005), Amaral, Castro Souza, and Stevenson (2008) and Lee and Chiu (2011) estimate smooth transition regression models with the temperature as a transition variable.

In the particular case of French data, Dordonnat, Koopman, Ooms, Dessertaine, and Collet (2008) adopt a multiple equation linear regression model, with one equation for each hour and time-varying parameters in order to capture changing trends and seasonal patterns. The specification contains several weather variables and Fourier terms for the yearly seasonal effect in the electricity load. With a multivariate approach, Dordonnat, Koopman, and Ooms (2012) consider the daily vector of the 24 hourly loads of the day and use a dynamic factor state-space model with time-varying coefficients for trend, seasonal and regression effects. Cho, Goude, Brossat, and Yao (2013) model and forecast the overall trend and seasonality at a weekly frequency, by fitting a generalized additive model with temporal and weather variables, and forecast the short run component via curve linear regressions. Finally, Goude, Nedellec, and Kong (2014) use a semi-parametric approach based on generalized additive models theory to forecast electricity load over about 2000 substations in the French distribution network.

In this paper, we investigate the performance of three alternative approaches to produce electricity load forecasts in France at a daily horizon. The first approach consists in a direct modeling and forecasting of hourly loads with various time series specifications. The models include the temperature, which is known to be a major determinant of electricity consumption in France. At this level, we confront several methods employed separately in the literature to capture the nonlinear link between the two variables: polynomial regression models, piecewise linear regressions with degree-day variables, smooth transition regression models and artificial neural network models. In contrast to the regression-based approach, neural networks (NN) do not require the specification of a particular parametric relationship between electricity load and its covariates. Moreover, they can approximate almost any nonlinear function. One contribution of this paper is to offer a unified framework to compare these four techniques.

Second, we propose to decompose the aggregate load into subseries with a wavelet transform: a low-resolution approximation which captures the long-run trend and a collection of details that are expected to catch the different seasonal cycles in the original data. Each component is predicted separately and the predictions are aggregated to obtain the forecast load. By doing so, we obtain subseries which can be forecasted more easily and therefore a more accurate final aggregate (see e.g. Antoniadis, Brossat, Cugliari, & Poggi, 2016; Ghayekhlooa, Menhaja, & Ghofranic, 2015). In the wavelet-based forecast, we need to address a boundary effect due to the missing values beyond the edge of the signal, as shown by Yu, Goldenberg, and Bi (2001) and Maslova, Ticlavilca, and McKee (2016) among others. In the context of electric load forecasting, Rana and Koprinska (2016) apply neural networks to forecast the missing observations beyond the border in a high frequency analysis of Australian and Spanish data. In this paper, we consider several approaches to handle this boundary problem, including the NN solution employed by Rana and Koprinska (2016). To provide the first missing point on the signal, we use the previous models which are designed to produce day-ahead forecasts and various linear or autoregressive extensions are proposed for the next missing observations.

Third, we examine the benefit of combining forecasts over individual models. The properties of the single forecasting models might vary over time and load periods, so that their combination improves forecast accuracy. This is particularly relevant for electricity loads characterized by a strong volatility and spikes. There is a long tradition in using combinations in time series forecasting (Bates & Granger, 1969; Newbold & Granger, 1974) and there has been a renewed interest in this approach in the last decade. In the energy field, encouraging results are achieved for electricity price forecasting by Bordignon, Bunn, Lisi, and Nan (2013), Raviv, Bouwman, and van Dijk (2015) and Bessec, Fouquau, and Méritet (2016). However, as recently documented by Nowotarski, Liu, Weron, and Hong (2016), only a few papers evaluate the potential of this approach for short-term load forecasting and provide supportive results. In this paper, we assess the forecasting accuracy of pooled forecasts obtained for various weighting schemes in a wide variety of models: univariate regressions and models incorporating meteorological variables, linear, piecewise linear, threshold specifications, neural networks, wavelet-based and direct load forecasts. Several pools of models are considered, including or not including wavelet predictions.

We perform an out-of-sample evaluation on French data. The forecasting accuracy of the models or combination of models is assessed over rolling windows. Particular attention is devoted to the impact of the border distortion in the multiresolution analysis. We calculate several forecasting criteria for five load periods representative of peak and off-peak hours. We perform two tests to compare the forecasting accuracy of the various approaches: the pairwise test of equal conditional predictive ability of Giacomini and White (2006) and a multiple comparison-based test with the Model Confidence Set (MCS) approach developed by Hansen, Lunde, and Nason (2011). The first test has the advantage of capturing the effects of estimation uncertainty on the relative forecast performance and is still implementable when two models are nested. The second one is a multilateral test and therefore does not require the choice of a benchmark model. As a robustness test, we replicate the evaluation for five forecasting periods, from 2010 to 2015, in order to check the stability of the results. This experimental design contrasts with the literature, which generally focuses on a single forecasting period. Yet, this robustness check seems necessary given the huge transformation of the electricity market in the recent period. Moreover, we assess the sensitivity of the results to several assumptions about the forecast error in temperature.

The out-of-sample evaluation on French data highlights the impact of the border effect in the wavelet analysis. Using basic extensions of the signal dramatically deteriorates the performance of the disaggregated method. By contrast, the wavelet approach outperforms the other individual forecasts when we properly deal with the edge effect by using the extensions proposed in the paper. More specifically, prolonging the signal with the one-day-ahead forecast and linear extrapolation outperforms the basic treatments provided in the softwares. However, we show that the ranking of the various approaches varies across load periods and forecasting samples. For this reason, individual forecasts are largely beaten by pooled forecasts. In particular, combinations including the wavelet predictions provide the most accurate forecasts. Time-varying weighting schemes relying on the past performance of the models provide the best results. These findings are valid for several load periods and forecasting samples covering the period 2010–2015. Our favorite combination mostly ranks first in terms of mean absolute error. Hence, we provide supportive evidence for the use of combinations in addition to the wavelet approach in electricity load forecasting.

The remainder of the paper is organized as follows. Section 2 describes the data. The third section presents the estimation and forecast design. The fourth section provides an assessment of the forecasting performance of the various specifications.