ARTICLE IN PRESS

ENEECO-02836; No of Pages 18

Energy Economics xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Energy Economics

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

An empirical comparison of alternative schemes for combining electricity spot price forecasts

Jakub Nowotarski [a], Eran Raviv [b], Stefan Trück [c], Rafał Weron [a,*]

[a] Institute of Organization and Management, Wrocław University of Technology, Wrocław, Poland
[b] Department of Econometrics, Erasmus University, Rotterdam, The Netherlands
[c] Faculty of Business and Economics, Macquarie University, Sydney, Australia

ARTICLE INFO

Article history:
Received 20 August 2013
Received in revised form 14 May 2014
Accepted 19 July 2014
Available online xxxx

JEL Codes:
C22 Time-Series Models
C52 Model Evaluation, Validation, and Selection
C53 Forecasting and Prediction Methods
L94 Electric Utilities
Q47 Energy Forecasting

Electricity price forecasting
Forecast combination
ARX model
Day-ahead market
Constrained least squares regression

ABSTRACT

In this comprehensive empirical study we critically evaluate the use of forecast averaging in the context of electricity prices. We apply seven averaging and one selection scheme and perform a backtesting analysis on day-ahead electricity prices in three major European and US markets. Our findings support the additional benefit of combining forecasts of individual methods for deriving more accurate predictions, however, the performance is not uniform across the considered markets and periods. In particular, equally weighted pooling of forecasts emerges as a simple, yet powerful technique compared with other schemes that rely on estimated combination weights, but only when there is no individual predictor that consistently outperforms its competitors. Constrained least squares regression (CLS) offers a balance between robustness against such well performing individual methods and relatively accurate forecasts, on average better than those of the individual predictors. Finally, some popular forecast averaging schemes – like ordinary least squares regression (OLS) and Bayesian Model Averaging (BMA) – turn out to be unsuitable for predicting day-ahead electricity prices.

1. Introduction

Since the early 1990s, structural reforms and deregulation have lead to significant changes in worldwide electricity markets. Like other commodities, electricity is now traded under competitive rules using spot and derivative contracts (Bunn, 2004; Shahidehpour et al., 2002). However, one particular feature of most, especially European, electricity markets is that what is called the spot market is actually a day-ahead market (Weron, 2006). This is a result of system operators requiring advance notice in order to verify that the schedule is feasible and lies within transmission constraints. In a day-ahead market agents submit their bids and offers for delivery of electricity during each load period (typically an hour or a half-hour) of the next day before a certain market closing time. Thus, when dealing with the modeling and forecasting of intraday prices it is important to recall that prices for all deliveries on the next day are typically determined at the same time using the same available information (Conejo et al., 2005; Huisman et al., 2007; Misiorek et al., 2006; Peña, 2012). The system price is then calculated as the equilibrium point for the aggregated supply and demand curves for each of the load periods. It should be noted, that although we use here the terms spot and day-ahead interchangeably, the former does not necessarily refer to the day-ahead market. In particular, in the US the spot market is another name for the so-called balancing or real-time market, which is a technical market used to price intraday deviations in supply and demand, while the day-ahead market is called the forward market. Also some markets in Europe (e.g. in the UK) nowadays admit continuous trading for individual load periods up to a few hours before delivery. With the shifting of volume from the day-ahead to balancing markets, also in Europe the term spot is more and more often being used to refer to the real-time market (Weron, 2014).

In contrast to other tradable commodities, electricity is practically non-storable. As a result the time series of electricity spot prices exhibit specific characteristics. The seasonal character of the prices is a direct consequence of demand fluctuations that mostly arise from deterministic conditions (such as business hours at the weekly level and the number of daylight hours at the yearly level) or climate conditions (like temperature and precipitation levels). In addition to seasonality and mean reversion, electricity prices exhibit an extremely high price

* Corresponding author.
E-mail addresses: jakub.nowotarski@gmail.com (J. Nowotarski), raviv@ese.eur.nl (E. Raviv), stefan.trueck@mq.edu.au (S. Trück), rafal.weron@pwr.wroc.pl (R. Weron).

http://dx.doi.org/10.1016/j.eneco.2014.07.014
0140-9883/

Please cite this article as: Nowotarski, J., et al., An empirical comparison of alternative schemes for combining electricity spot price forecasts, Energy Econ. (2014), http://dx.doi.org/10.1016/j.eneco.2014.07.014

volatility as well as infrequent, but large price spikes. These features have forced producers and wholesale consumers to hedge not only against volume risk but also against price movements. This in turn has significantly enhanced research efforts towards modeling and forecasting spot electricity prices.

A wide range of econometric or statistical models have been suggested in the literature including regression models, jump-diffusions, GARCH-type models and regime-switching models (see e.g. Garcia-Martos and Conejo, 2013; Hong, 2014; Huisman, 2009; Weron, 2014). In the context of predicting spot price movements, each model specification yields a different forecast. Facing the variety of alternative models available in the literature, one could discard all of the models but one on the basis of their goodness of fit and forecasting performance. Alternatively, one can allocate weights to the various forecasts produced by individual models in order to obtain a combined forecast for the day-ahead electricity price. The latter strategy may be more favorable in the context of changing model and predictor relevance through time and can potentially achieve a better forecasting performance by virtue of smoothed model selection. The best model is not known in advance so allocating weights to the individual models is used as a hedge against the possibility of an inaccurate model choice. That said, combining forecasts is not a one-size-fits-all technique to increase forecast accuracy. There are cases in which combining forecasts leads to worse predictions (Hubrich, 2005). Intuitively, there is no reason to opt for combining forecasts if a superior model is easily recognized beforehand.

Despite the increasing body of literature on the use of forecast combinations (also referred to as combining forecasts, forecast averaging or model averaging) for prediction, there is only a small number of applications of these techniques in the area of electricity markets. To our best knowledge, in the context of electricity spot price forecasting existing applications so far only include the work by Bordignon et al. (2013), Maciejowska et al. (2014), Nan (2009), Nowotarski and Weron (2014) and Raviv et al. (2013). The relatively small number of studies on combining forecasts produced by various models is surprising since, on the one hand, research shows that performance of individual models is often unstable and dependent on the considered periods of price behavior (see e.g. Aggarwal et al., 2009; Conejo et al., 2005; Weron and Misiorek, 2008), on the other, forecast averaging has been used in the context of load forecasting for over three decades (see e.g. Bunn, 1985; Bunn and Farmer, 1985; Smith, 1989; Taylor, 2010; Taylor and Majithia, 2000). This motivates us to thoroughly investigate which forecast combination schemes and under what market conditions are able to outperform individual methods in forecasting day-ahead electricity prices.

The contribution of our paper is twofold. First, we apply a great variety of stochastic models and forecast combination techniques to the data. These include, for example, standard autoregressive models, regime-switching models, mean-reversion jump diffusion models and semiparametric autoregressive models. Techniques for forecast combinations include simple equal weighted averaging, forecast combinations based on OLS regression, constrained least squares regression (CLS, PW), Least Absolute Deviation (LAD) regression, as well as model averaging based on a Bayesian approach. The majority of the averaging techniques have not been applied to forecasting electricity spot prices yet.

Second, we provide the so far most extensive study, using four datasets from key electricity markets worldwide. The quite unique behavior of electricity spot prices, including seasonality, periods of extreme volatility and price spikes, may provide good reasons for questioning the good performance of forecast averaging also for this class of assets. To thoroughly investigate this issue we include markets and time periods that are characterized by a different behavior with respect to volatility and the number of price spikes during examined out-of-sample periods. Considered markets include the Nordic power exchange (Nord Pool), the European Energy Exchange in Leipzig (EEX) and the Pennsylvania–New Jersey–Maryland Interconnection (PJM). For these markets, we compare the averaging techniques with the realistic situation where the market participants have to decide ex ante which individual model to use. Hereby, we assume that participants decide to pick one of the models that performed well in the past, and then examine the performance of this model in comparison to the averaging techniques. We evaluate the performance based on different criteria and conduct Diebold–Mariano tests in order to investigate whether combining forecasts can significantly improve the performance.

The remainder of the article is organized as follows. Section 2 provides an overview of the recent literature on forecast averaging and its limited applications in electricity markets. Section 3 describes the four datasets used in this study, while Section 4 reviews the individual models for forecasting electricity spot prices and the applied averaging techniques. Finally, Section 5 presents empirical results and Section 6 concludes.

2. Combining forecasts and electricity markets

The idea of combining forecasts goes back to the late 1960s, with the works of Bates and Granger (1969), Crane and Crotty (1967) and Newbold and Granger (1974). Examining forecast combinations, using various models and weights based on mean squared errors, the authors found a significant improvement in terms of reducing prediction errors. Since then, many authors have suggested the superior performance of forecast combinations over the use of individual models, see e.g. Clemen (1989), Diebold and Pauly (1987), de Menezes et al. (2000), Stock and Watson (2004), Timmermann (2006) and references therein. Forecast averaging has become so popular, with so many different ways to combine forecasts, that Andrawis et al. (2011) suggest to use hierarchical forecast combinations, i.e. combining combined forecasts.

2.1. Electricity loads and transmission congestion

While there is a large body of literature on forecasting day-ahead electricity prices and loads, only few of these studies examine the performance of combining forecasts obtained from individual models. However, already in the 1980s combining has been used in practice—the system load predictor used by the British system operator was a combination of three methods (Bunn and Farmer, 1985). Other examples from that decade include papers by Bunn (1985) and Smith (1989). In the former, the author theoretically discusses concepts and methods for combining several forecasting models in electricity load prediction and suggests that the approach can be seen as a surrogate device, either for finding a simple comprehensive model or identifying one single model that can be considered to be most appropriate. Smith (1989), on the other hand, provides empirical evidence in favor of forecast averaging. He combines several ARIMA time series models for electricity demand and concludes that the combined forecasts are significantly more accurate than any of the forecasts obtained from the individual models. The weights of the forecast combinations change for different days of the week, to overcome cyclic modeling weaknesses of the individual models, and the selection and combination of forecasts from different prediction methods is conducted on the basis of recent forecasting performance only, with no a priori assumptions about demand behavior.

A decade later, Taylor and Majithia (2000) apply switching and smooth transition forecast combination models for electricity demand profiling. The applied models allow for combining weights to vary across half-hourly intervals which is an appealing feature as different forecast models may be more suitable for different periods. A number of criteria are used to control the changing weights, including weather and the shape of demand profiles. Empirical results suggest an improved post-sample forecasting performance of the applied models. In a follow up study, Taylor (2010) applies so-called triple seasonal methods for short-term electricity demand forecasting. Hereby, double

Please cite this article as: Nowotarski, J., et al., An empirical comparison of alternative schemes for combining electricity spot price forecasts, Energy Econ. (2014), http://dx.doi.org/10.1016/j.eneco.2014.07.014