Energy Economics 57 (2016) 228–235

Contents lists available at ScienceDirect

Energy Economics

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

On the importance of the long-term seasonal component in day-ahead electricity price forecasting

Jakub Nowotarski, Rafał Weron*

Department of Operations Research, Wrocław University of Technology, Wrocław, Poland

ARTICLE INFO

Article history:
Received 16 March 2016
Received in revised form 21 May 2016
Accepted 25 May 2016
Available online 2 June 2016

JEL classification:
C14
C22
C51
C53
Q47

Keywords:
Electricity spot price
Forecasting
Day-ahead market
Long-term seasonal component

ABSTRACT

In day-ahead *electricity price forecasting* (EPF) the daily and weekly seasonalities are always taken into account, but the long-term seasonal component (LTSC) is believed to add unnecessary complexity to the already parameter-rich models and is generally ignored. Conducting an extensive empirical study involving state-of-the-art time series models we show that (i) decomposing a series of electricity prices into a LTSC and a stochastic component, (ii) modeling them independently and (iii) combining their forecasts can bring – contrary to a common belief – an accuracy gain compared to an approach in which a given time series model is calibrated to the prices themselves.

1. Introduction

Without doubt *electricity price forecasting* (EPF) is of prime importance to the functioning of today's energy business. Alongside load forecasting, *short-term* (also called *spot* or *day-ahead*; for a discussion see Weron, 2014) EPF has become the core process of an energy company's planning activities at the operational level. Although it is very hard to quantify the benefits of improving load and/or price forecasts, Hong (2015) provides interesting back-of-the-envelope calculations. Based on U.S. data from the last decade, he concludes that for a typical medium-size utility with a 5 GW peak load, savings from a 1% reduction in the Mean Absolute Percentage Error (MAPE) are as much as $1.5 million per year from short-term load forecasting and $3 million per year from short-term load and price forecasting! Hong's simplified calculations coincide quite well with more technical studies of Hobbs et al. (1999), who conclude that on average a reduction of 1% in MAPE for short-term load forecasts decreases variable generation costs by 0.1%–0.3% when MAPE is in the typical range of 3%–5%, and of Zareipour et al. (2010), who find that when MAPE is the 5% to 15% range commonly observed for short-term price forecasts, a 1% improvement in forecast accuracy would result in about 0.1%–0.35% cost reductions. In both studies, the level of actual savings depends to a large extent on generator characteristics.

As has been noted in a number of studies, a key point in electricity spot price modeling and forecasting is the appropriate treatment of seasonality (Janczura et al., 2013; Keles et al., 2016; Lisi and Nan, 2014; Maciejowska, 2014; Nowotarski et al., 2013). For mid-term horizons – ranging from a few days to a few months ahead and typically considered in derivatives pricing and risk management applications – the daily profile is usually regarded as irrelevant (Ignatieva and Trück, 2016; Janczura, 2014). In fact, most mid-term EPF models work with average daily prices and focus on the *annual* or *long-term seasonal component* (LTSC; also called the *trend-seasonal component*). However, in short-term EPF the daily and weekly seasonalities are always taken into account, but the LTSC is believed to add unnecessary complexity to the already parameter-rich models and is generally ignored (for recent reviews see Garcia-Martos and Conejo, 2013; Weron, 2014). But is this the right approach? Should

* Corresponding author.
E-mail addresses: jakub.nowotarski@pwr.edu.pl (J. Nowotarski), rafal.weron@pwr.edu.pl (R. Weron).

http://dx.doi.org/10.1016/j.eneco.2016.05.009

the LTSC be included in day-ahead EPF models, contrary to a common belief that it is redundant in the short-term?

It is exactly the aim of this paper to address these two important questions, that have not been investigated in the EPF literature to date. We perform an extensive empirical study which involves:

- two 2-year long, hourly resolution test periods from two distinct power markets (GEFCom2014 and Nord Pool),
- two autoregressive model structures — one originally proposed by Misiorek et al. (2006) and later used in a number of EPF studies (Gaillard et al., 2016; Kristiansen, 2012; Maciejowska et al., 2016; Nowotarski et al., 2014; Nowotarski and Weron, 2016; Serinaldi, 2011; Weron, 2006; Weron and Misiorek, 2008; Ziel, 2016) and one which evolved from it during the successful participation of TEAM POLAND in the Global Energy Forecasting Competition 2014 (GEFCom2014; see Hong et al., 2016; Maciejowska and Nowotarski, 2016),
- two novel *Seasonal Component AutoRegressive* (SCAR) models that combine a 24 hour-ahead extrapolation of an estimated LTSC with the forecasts of autoregressive models,
- two well-performing LTSC model classes — wavelet smoothing and the Hodrick and Prescott (1997) filter, as advocated by Janczura et al. (2013), Lisi and Nan (2014), Nowotarski et al. (2013) and Weron and Zator (2015),
- model validation in terms of the robust *weekly-weighted mean absolute error* (WMAE; see Weron, 2014) and the Diebold and Mariano (1995) test,

and draw statistically significant conclusions with far reaching consequences for day-ahead EPF.

The remainder of the paper is structured as follows. In Section 2 we present the datasets. Then in Section 3 we describe the techniques considered for price forecasting: two baseline autoregressive model structures, two LTSC model classes and two novel SCAR models. In Section 4 we summarize the empirical findings and in Section 5 wrap up the results and conclude.

2. Datasets

The datasets used in this empirical study include two day-ahead time series. The first one comes from the Global Energy Forecasting Competition 2014 (GEFCom2014) — the largest energy forecasting competition to date, both in terms of the diversity of competition topics and wide geographic coverage of the participants (for details see Hong et al., 2016). The dataset includes three time series at hourly resolution: locational marginal prices, day-ahead predictions of zonal loads and day-ahead predictions of system loads and covers the period from January 1, 2011 to December 17, 2013. During the competition the information set was being extended on a weekly basis to prevent 'peeking' into the future. However, now it is available in whole as supplementary material (Appendix A) accompanying Hong et al. (2016). In this paper we only use two subseries — locational marginal prices and day-ahead predictions of zonal loads, see Fig. 1. The origin of the data has never been revealed by the organizers.

The second dataset comes from one of the major European power markets — Nord Pool (NP). It comprises hourly system prices and hourly *consumption prognosis* for four Nordic countries (Denmark, Finland, Norway and Sweden) for the period January 1, 2013–December 26, 2015, see Fig. 2. The time series were constructed using data published by the Nordic power exchange Nord Pool (www.nordpoolspot.com) and preprocessed to account for missing values and changes to/from the daylight saving time (like in Weron (2006), Section 4.3.7). The missing data values were substituted by the arithmetic average of the neighboring values. The 'doubled' values (corresponding to the changes from the daylight saving/summer time) were substituted by the arithmetic average of the two values for the 'doubled' hour.

For both markets, the day-ahead forecasts of the hourly electricity price are determined within a rolling window scheme, using a 360-day calibration window. First, all considered models (their short-term and long-term components) are calibrated to data from the initial calibration period, i.e. January 1 to December 26 (year 2011 for GEFCom2014 and 2013 for Nord Pool) and forecasts for all 24 h of the next day (December 27) are determined. Then the window is rolled forward by one day and forecasts for all 24 h of December 28 are computed. This procedure is repeated until the predictions for the last day in the sample – December 17, 2013 (for GEFCom2014) or December 26, 2015 (for Nord Pool) – are made.

3. Methodology

3.1. The benchmarks

Our choice of the benchmark models is guided by previous literature on electricity price forecasting and experience gained during the successful participation of TEAM POLAND in the GEFCom2014 competition. The modeling is implemented separately across the hours, leading to 24 sets of parameters for each day the forecasting exercise is performed. This approach is inspired by the fact that each hour displays a rather distinct price profile, reflecting the daily variation of demand, costs and operational constraints, and by the extensive research on demand forecasting, which has generally favored the multi-model specification for short-term predictions (see Weron, 2014 for a review).

The first benchmark belongs to the class of similar-day techniques. Most likely it was introduced to the EPF literature by Nogales et al. (2002) and dubbed the *naïve method*. It proceeds as follows: hour h on Monday is similar to the same hour on Monday of the previous week, and the same rule applies for Saturdays and Sundays; hour h on Tuesday is similar to the same hour on Monday, and the same rule applies for Wednesdays, Thursdays and Fridays. As was argued by Conejo et al. (2005) and Nogales et al. (2002), forecasting procedures that are not calibrated carefully fail to pass this 'naïve test' surprisingly often. We will denote this benchmark by *Naïve*.

The second model is a parsimonious autoregressive structure originally proposed by Misiorek et al. (2006) and later used in a number of EPF studies (Gaillard et al., 2016; Kristiansen, 2012; Maciejowska et al., 2016; Nowotarski et al., 2014; Weron, 2006; Weron and Misiorek, 2008; Ziel, 2016). Within this model the natural logarithm of the electricity spot price, $p_t = log(P_t)$, is given by the following formula:

$$p_t = \phi_1 p_{t-24} + \phi_2 p_{t-48} + \phi_7 p_{t-168} + \phi_8 mp_t + \psi_1 z_t + \sum_{i=1}^{3} d_i D_i + \varepsilon_t, \quad (1)$$

where the lagged log-prices p_{t-24}, p_{t-48} and p_{t-168} account for the autoregressive effects of the previous days (the same hour yesterday, two days ago and one week ago), while mp_t creates the link between bidding and price signals from the entire previous day (it is the minimum of the previous day's 24 hourly log-prices). The variable z_t refers to the hourly zonal load of a US utility or Nordic consumption (actually to forecasts made a day before, see Section 2). The three dummy variables – D_1, D_2 and D_3 (for Monday, Saturday and Sunday, respectively) – account for the weekly seasonality. Finally, the ε_ts are assumed to be independent and identically distributed (i.i.d.) normal variables. We will denote this autoregressive benchmark by ARX to reflect the fact that the load (or consumption) forecast is used as the eXogenous variable in Eq. (1).

The third benchmark is an extension of the ARX model, which takes into account the experience gained during the GEFCom2014