CrossMark

ORIGINAL PAPER

Forecasting of daily electricity prices with factor models: utilizing intra-day and inter-zone relationships

Katarzyna Maciejowska · Rafał Weron

Abstract This paper investigates whether using hourly and/or zonal prices can improve the accuracy of short- and medium-term forecasts of average daily electricity prices. We consider a 6 years period (2008–2013) of hourly day-ahead prices from 19 zones of the Pennsylvania–New Jersey–Maryland (PJM) interconnection and the PJM Dominion Hub in Virginia, U.S. The predictive performance of four multivariate models calibrated to hourly and/or zonal day-ahead prices is evaluated and compared with that of a univariate model, which uses only average daily data for the Dominion Hub. The multivariate competitors include a restricted vector autoregressive model and three factor models with the common and idiosyncratic components estimated using principal components in a semiparametric setup. The results indicate that there are statistically significant forecast improvements from incorporating the additional information, essentially for all considered forecast horizons ranging from 1 day to 2 months, but only when the correlation structure of prices across locations and/or hours is modeled using factor models.

K. Maciejowska · R. Weron (✉)
Institute of Organization and Management, Wrocław University of Technology,
Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: rafal.weron@pwr.edu.pl

K. Maciejowska
CERGE-EI, Politickych veznu 7, 11121 Prague 1, Czech Republic
e-mail: katarzyna.maciejowska@pwr.edu.pl

🍦 Springer

1 Introduction

Electricity price forecasting has attracted a lot of attention in the recent years. Short-(up to a few days ahead) and medium-term (up to a few months ahead) price forecasts are of particular interest to power portfolio managers. A generator, a utility company or a large industrial consumer able to forecast the volatile wholesale prices with a reasonable accuracy can adjust its bidding strategy and its own production or consumption schedule to reduce risk or to maximize profits in day-ahead trading. However, electricity is a very special commodity. Demand and to some extent supply is weather and business cycle dependent, yet at the same time electricity is non-storable (at least not economically). The need for keeping a constant balance between production and consumption results in—unobserved in any other financial or commodity market—volatile, complex and hard to predict price dynamics (Eydeland and Wolyniec 2003; Kaminski 2013).

Another consequence of the non-storability is that there is no *spot price* in the strict sense of the word. Electricity delivered at 10 am is a different product than that delivered at 11 am. Hence, continuous trading of 'electricity' is not possible, only of 'electricity delivered during a given load period' (say, from 10 to 11 am). In most liberalized power systems a day-ahead market plays the dominant role and sets the reference price for exchange-traded and OTC derivative contracts. In such an auction market agents submit their bids and offers for delivery of electricity during each hour (or a shorter load period) of the next day before a certain market closing time. The European convention is to refer to this price as the *spot price*. However, in the U.S. the term *spot price* is typically reserved for the intra-day *real-time market*, known as the *balancing* or *intra-day market* in Europe (Weron 2006). The average of the 24 hourly (48 half-hourly) day-ahead prices is called the *daily price*, the *daily spot price* or the *baseload price*. It is also the focus of this paper.

A variety of methods and ideas have been tried for electricity spot price forecasting, with varying degrees of success (see Weron 2014 for a comprehensive review). Thus far, the literature on forecasting daily prices has concentrated on models that use only information at the aggregated, i.e. daily, level (see e.g. Bernhardt et al. 2008; Bierbrauer et al. 2007; Chan and Gray 2006; Koopman et al. 2007; Schlueter 2010). On the other hand, the very rich literature on forecasting hourly or half-hourly prices has used disaggregated data, i.e. hourly or half-hourly, but generally has not utilized the complex dependence structure of the multivariate price series. At least until very recently.

In one of the first papers that touched upon this topic, Chen et al. (2008) converted hourly electricity prices with multiple seasonalities into several time series with only weekly seasonality by manifold learning (an extension of PCA) and predicted them using three techniques. Their approach compared favorably to that of ARIMA, ARX and naive methods in 1 day, 1 week and 1 month ahead forecasting of hourly NYISO (U.S.) prices. Vilar et al. (2012) used a nonparametric regression technique with functional explanatory data and a semi-functional partial linear (SFPL) model to forecast hourly day-ahead prices in the Spanish market and found it superior to ARIMA and naive approaches. Garcia-Martos et al. (2012) proposed to extract common factors from hourly prices and use them for 1 day-ahead forecasting within a dynamic factor

model (DFM) framework. They also reported on some preliminary results showing the usefulness of factor models for mid- and long-term predictions. We will return to this issue later in the text.

More recently, Elattar (2013) proposed to combine kernel principal component analysis (KPCA; to extract features of the inputs and obtain kernel principal components) with a Bayesian local informative vector machine (IVM; to make the predictions) and found it superior in short-term price forecasting to 12 other methods, including ARIMA and neural network techniques, for the Spanish market in 2002. Wu et al. (2013) proposed a recursive dynamic factor analysis (RDFA) algorithm, where the principal components were recursively tracked using an efficient subspace tracking algorithm while their scores were tracked and predicted recursively using a Kalman filter. The RDFA was shown to outperform functional PCA, AR with time varying mean and support vector regression in predicting hourly day-ahead prices in the Australian and New England (U.S.) markets.

In this paper, however, we take a different modeling perspective that has its origins in macroeconomics. Namely, we do not focus on forecasting hourly (i.e. disaggregated) electricity prices, although such price forecasts can be obtained within the multivariate models we study. Instead, we address the question, how to build efficient models for predicting daily (or aggregated) prices. In particular, whether incorporating the intra-day (from hourly day-ahead prices) and inter-zone relationships of electricity prices in the Pennsylvania–New Jersey–Maryland (PJM) Interconnection can improve the accuracy of daily spot price forecasts for a major hub in this market—the PJM Dominion Hub.

We should note that while new to the energy economics literature, the idea of using disaggregated data for forecasting of aggregated variables is not that novel. It has been exploited intensively in the last decade to predict inflation (Bermingham and D'Agostino 2014), the Gross Domestic Product (Perevalov and Maier 2010) or the production index (Stock and Watson 2002). Interestingly, the general conditions under which the use of disaggregated data improves forecasting performance have been recently formulated by Hendry and Hubrich (2011). This result has far reaching consequences well beyond macroeconometrics.

Our work also complements three recent electricity price forecasting papers. In an article, having its roots in the fundamentals of price formation in auction markets, Liebl (2013) proposed to model and predict electricity spot prices by first finding the functional relation between prices and demand in terms of daily price-demand functions, then parametrizing the series of daily price-demand functions using a functional factor model. He demonstrated the power of this approach by comparing 1–20 days ahead forecasts of the model with those of two simple univariate time series models for daily prices (AR and MRS) and two alternative functional data models for hourly prices (DSFM and SFPL). In effect—like us—Liebl compared aggregated daily price forecasts. However, his motivation for working with daily price forecasts was different. In another recent paper, Raviv et al. (2013) exploited the information embedded in the cross correlation of Nord Pool hourly price series to yield more accurate one step-ahead average daily price forecasts for Scandinavia. Finally, Maciejowska and Weron (2013) used a panel of half-hourly data from the UK power market to predict the average daily day-ahead prices from one to 60 days ahead, both directly (via VAR

type models) and indirectly (via factor models). This paper extends these studies by (i) considering not only intra-day (24 h per day) but also inter-zone (19 zones and one hub) relationships and by (ii) utilizing factor models with idiosyncratic components. The motivation for testing the impact of inter-zone relationships comes also from the papers of Dempster et al. (2008) and Higgs (2009), who suggest that joint modeling of prices in connected markets may improve the forecast accuracy.

It should be also noted here that the modeling approach we use is semiparametric in nature, as defined by Powell (1994). In order to decompose a set of variables, presented in a form of a panel, into common and idiosyncratic components, no assumptions about a particular type of distribution of neither factors nor residuals are required. The assumptions, which are necessary to identify the two components, restrict only the correlation structures and moments of the underlying processes. Our approach can be further extended and made more explicitly semiparametric by smoothing the factor loadings (e.g. using B-splines as in the DSFM model of Park et al. 2009). However, since our focus is on forecasting aggregated daily prices (not disaggregated hourly prices as in Härdle and Trück 2010) we do not require smooth factor loadings and, hence, do not use the DSFM approach here.

The remainder of the paper is structured as follows. In Sect. 2, we briefly describe the PJM market and present electricity price data used in this study. In Sect. 3, we describe the benchmark univariate model and alternative multivariate models, which use the information contained in hourly and/or zonal prices. Next, in Sect. 4, we evaluate the forecasting performance of the five tested models. Finally, in Sect. 5, we wrap up the results and make suggestions for future work in this area.

2 The PJM Interconnection and market data

The PJM interconnection is the world's largest competitive wholesale electricity market. Similar to the Scandinavian Nord Pool market, PJM provides an interesting example of market design where organized markets and transmission pricing are integrated. PJM is a regional transmission organization (RTO) that coordinates the movement of wholesale electricity in all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia and the District of Columbia. As of today it serves over 60 million people and has more than 800 market participants (see www.pjm.com). PJM combines the role of a power exchange, a clearing house and a system operator. It operates several markets, although different in detail: two generating capacity markets (daily and long-term), two energy markets (day-ahead and real-time), a financial transmission entitlements market and an ancillary services market.

The data used in this study was downloaded from the GDF Suez website (www.gdfsuezenergyresources.com) and contains hourly day-ahead prices for 19 PJM zones (APS, AEP, AECO, ATSI, BGE, ComEd, Dayton, Delmarva, Dominion, Duke, Duquesne, JCPL, Metro Edison, PennElec, Rockland, PECO, PEPCO, PPL, PSEG) and the Dominion Hub. The latter is a major market hub and comprises a group of approximately 650 nodes in Virginia (U.S.) within Dominion's Virginia Power control area. The Dominion control area is also referred to as PJM South.
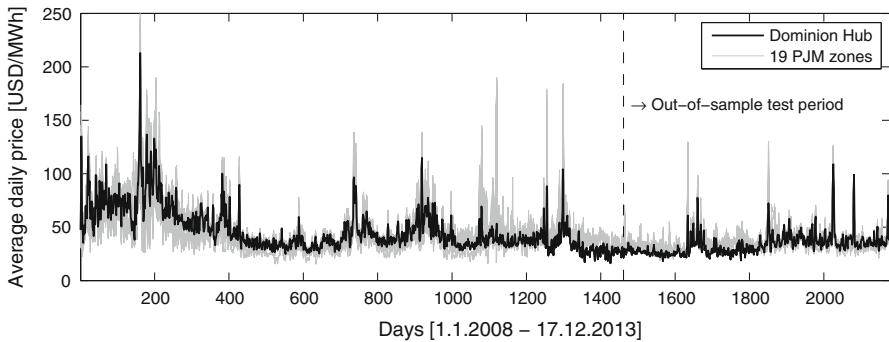
Fig. 1 Average daily (baseload) day-ahead prices for the Dominion Hub and 19 PJM zones from the nearly 6 years period 1.1.2008–17.12.2013. To show the inter-zone price variability all zonal prices are plotted in *gray*

The data spans a nearly 6 years period—from 1.1.2008 to 17.12.2013—and includes hourly prices. Depending on the model structure, we use one of three data panels for calibration. The largest panel (Panel-HL) includes hourly prices for all 20 locations and consists of $24 \times 20 = 480$ variables, the intermediate panel (Panel-H) contains hourly prices for the PJM Dominion Hub (24 variables), whereas the smallest panel (Panel-L) includes daily prices across all 20 locations (20 variables). We use the last 2 years (1.1.2012 to 17.12.2013) or 717 days to evaluate the out-of sample forecasting performance, see Figs. 1 and 2. For each daily forecast, we roll the four-year calibration window forward by 1 day to ensure that all models are estimated on a sample of the same size.

3 The models

In this article, we focus on autoregressive (AR) and vector autoregressive (VAR) models, augmented by deterministic terms. Since a stable (Lütkepohl 2005) AR(q) or VAR(q) process has a moving average representation, it will return to its mean after any shock, even for $q > 1$. The dynamics of the return to the process mean depends on the model parameters and the lag order. To model the seasonal pattern of the process mean, we extend the AR and VAR models with a 3×1 vector of deterministic variables—denoted by D_t—composed of a constant, a dummy representing the day type (working day vs. weekend) and the number of daylight hours (which mimics the annual seasonality). Note that the approach we take is less popular than the classical seasonal decomposition of the price series into the long-term trend-seasonal component (LTSC), the short-term seasonal component (STSC) and the remaining stochastic component, and separate estimation of the three parts (Janczura et al. 2013). Instead, we jointly estimate the deterministic (hence, straightforward to predict) trend-seasonal component and the stochastic part. In our models, D_t includes a simple STSC (working day vs. weekend; not a separate dummy for each day of the week) and a periodic LTSC (the number of daylight hours for a particular day of the year), which is better at describing annual seasonality than a simple sine wave but worse than a wavelet
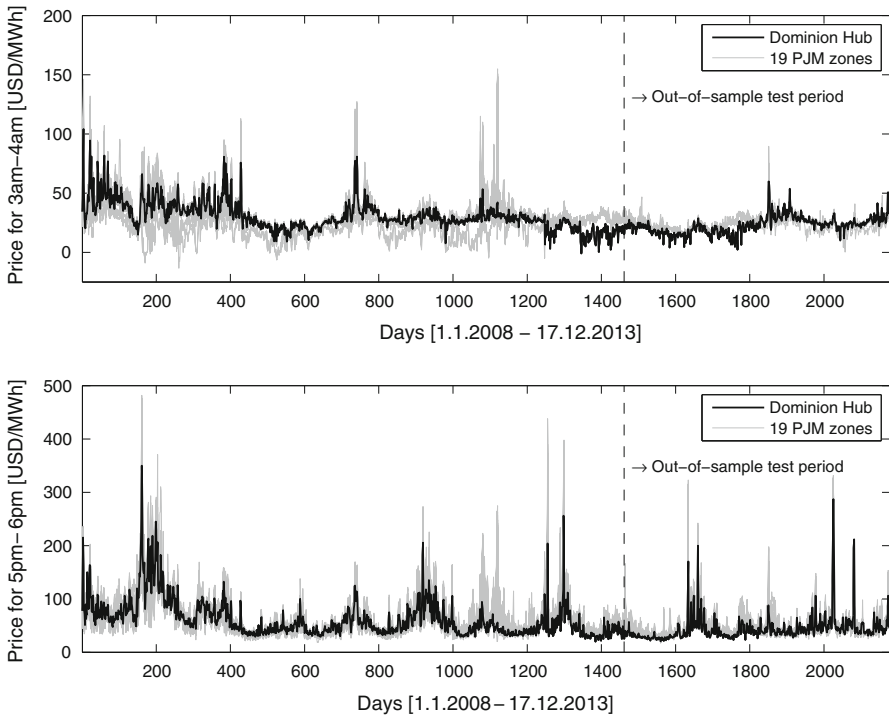
Fig. 2 Hourly day-ahead prices for the Dominion Hub and 19 PJM zones from the nearly 6 years period 1.1.2008–17.12.2013. Note the different price scales for the typical off-peak (3–4 am, *top*) and on-peak (5–6 pm, *bottom*) load periods. To show the inter-zone price variability all zonal prices are plotted in *gray*. In some zones even negative prices can be observed for off-peak hours

smoother (see e.g. Nowotarski et al. 2013). Finally, we keep the lag order, q, and the set of deterministic variables, D_t, constant for all types of models.

3.1 The benchmark

We choose an AR(q) model of daily day-ahead prices as the benchmark because of its widespread use in the literature and its relatively good performance in predicting electricity prices (given its simplicity; see e.g. Conejo et al. 2005; Weron 2006; Misiorek et al. 2006). It uses only aggregated, daily data and, hence, is suitable for comparison of all models studied in this paper.

In this model—denoted later in the text as **AR**—we describe the daily day-ahead price P_t by:

$$P_t = \alpha D_t + \sum_{i=1}^{p} \beta_i P_{t-i} + \varepsilon_t, \tag{1}$$

where D_t is a 3×1 vector of exogenous, deterministic variables, α is a 1×3 vector of parameters and β_i are the autoregressive parameters. We choose the lag order to

Table 1 Model types and notation

Name	Model
Simple autoregressive models	
AR	The benchmark AR model of daily prices (P_t)
AR-H	Average of 24 AR models of hourly prices (P_{kt})
Factor models with idiosyncratic components	
PC-HL	Calibrated to hourly and location specific prices in panel-HL
PC-H	Calibrated to hourly PJM Dominion hub prices in panel-H
PC-L	Calibrated to daily location specific prices in panel-L

In all cases we are forecasting the daily day-ahead electricity price in the PJM Dominion Hub and the forecasting horizon ranges from one (one step-ahead forecasts) to 60 days (60 step-ahead forecasts)

be $q = 7$, which is in line with the approach of Kristiansen (2012) and Weron and Misiorek (2008), who also used a lag order of 7 days when forecasting California and Nord Pool day-ahead prices. The same lag order is applied to all other autoregressive models analyzed in this paper.

3.2 Autoregressive models of hourly prices

Since the daily day-ahead prices P_t are the arithmetic average of the hourly prices P_{kt}, we can model separately each hour $k = 1, \ldots, 24$ with an AR(q) process:

$$P_{kt} = \alpha_k D_t + \sum_{i=1}^{q} \beta_{ik} P_{k,t-i} + u_{kt}, \tag{2}$$

and obtain the daily price P_t by taking their average. This model is denoted later in the text as **AR-H**, see also Table 1.

Note, that within this approach we estimate separately 24 models for hourly prices. This can be interpreted as a restricted VAR(q) model, with diagonal parameter matrices B_i and uncorrelated residuals u_t:

$$Y_t = A D_t + \sum_{i=1}^{q} B_i P_{t-i} + u_t,$$

where $Y_t = [P_{1t}, \ldots, P_{24t}]'$, $u_t = [u_{1t}, \ldots, u_{24t}]'$, A is a 1×3 vector of mean parameters and B_i are 24×24 matrices of autoregressive parameters.

The restricted VAR(q) model uses information about hourly prices but does not exploit their correlation structure. Since all hours during the day are correlated with each other, it seems reasonable to model them jointly. However, if we decide to model them together, the large number of parameters to estimate may result in over-fitting, yielding small in-sample residuals but large out-of-sample errors. For example, in a VAR(q) model of hourly data for one location, there will be $1 + 24q$ parameters in each equation.

Table 2 Data characteristics
and the number of factors and
the explained variability for each
of the three used panels

	Panel-HL	Panel-H	Panel-L
Data panel characteristics			
Data frequency	Hourly	Hourly	Daily
Number of locations	20	1	20
Total number of variables	480	24	20
Factor model characteristics			
Number of factors	8	2	3
Information criterion	IC_2	IC_3	IC_3
Explained variability	96.3 %	95.1 %	97.3 %

3.3 Factor models

If we want to explore the structure of electricity prices, we need to use some dimension reduction methods. In this study, we propose to apply factor models, with factors estimated as principal components. If we treat the electricity day-ahead prices across locations and hours as a panel then we can use the approach described in Bai (2003), Bai and Ng (2002) and Stock and Watson (2002). It was shown that the principal component (PC) estimation method is consistent for large dimensional models (where both of the dimensions: time and the number of series) tends to infinity. In the largest panel, we observe 480 variables, which should be sufficient to approximate the true factors.

The main assumption of the factor models is that all variables $P_{kt}, k = 1, \ldots, 480$ for Panel-HL (respectively 24 and 20 for Panel-H and Panel-L, see Table 2), co-move and depend on a small set of common factors $F_t = [F_{1t}, \ldots, F_{Nt}]'$. The individual series P_{kt} can be modeled as a linear function of N principal components F_t and stochastic residuals v_{kt}:

$$P_{kt} = \Lambda_k F_t + v_{kt}, \tag{3}$$

where the loads $\Lambda_k = [\Lambda_{k1}, \ldots, \Lambda_{kN}]$ describe the relation between the factors F_t and the panel variables P_{kt}. Note, that these loads are not 'power system loads', but model parameters as in Bai (2003). It was shown in Stock and Watson (2002) and Bai (2003) that the eigenvectors corresponding to the N largest eigenvalues of the matrix $P'P$ multiplied by $\sqrt{T}$ are consistent estimators of the common factors F_t.

The number of common factors can be chosen on the basis of information criteria or the fraction of total variability explained. Here, we use the information criteria IC_2 and IC_3 proposed by Bai and Ng (2002). The resulting choice of the number of factors and the explained variability are provided in Table 2.

Once the disaggregated models are estimated, then the daily electricity prices can be obtained by averaging the hourly prices. The resulting models are denoted by **PC-HL**, **PC-H** and **PC-L**, depending on the data panel used for calibration, respectively Panel-HL, Panel-H and Panel-L, see Table 2.

In order to predict future values of hourly prices, we need to forecast both, the common factors F_{nt} and the idiosyncratic components v_{kt}. Although the factors are contemporaneously orthogonal, due to normalization assumptions, they may be still

inter-temporally correlated. Hence, it seems reasonable to model them jointly. Moreover, they may depend on some other variables, such as the deterministic variables (D_t). At the same time, the idiosyncratic components can be only weakly correlated across periods and therefore can be modeled separately, for each hour. Moreover, they cannot have the same seasonal pattern because all the co-movement between hours is captured by the factors.

In this study, the common factors are assumed to follow a vector autoregressive VAR(q) model:

$$F_t = \Phi D_t + \sum_{i=1}^{p} \Theta_i F_{t-i} + \zeta_t, \tag{4}$$

where Φ denotes a $N \times 3$ matrix of deterministic coefficients and Θ_i are $N \times N$ matrices of autoregressive parameters. To describe and forecast the idiosyncratic components we use a simple autoregressive AR(q) structure:

$$v_t = \sum_{i=1}^{p} \gamma_i v_{t-i} + \epsilon_t, \tag{5}$$

which does not include deterministic nor fundamental variables.

4 Forecasting performance

4.1 Evaluation of the forecasting performance

In this section, we examine, whether using the intra-day (*from hourly day-ahead prices*) and inter-zone information improves the forecast accuracy. We use an AR(q) model of daily day-ahead prices as the benchmark.

We consider different forecast horizons. One step-ahead forecasts are typically used for forecast comparison in power market studies (Weron 2006). However, other forecast horizons are also very important for risk management and derivatives pricing applications. Hence, we consider here short- and mid-term forecast horizons. For each time point t and forecast horizon $\tau = 1, \ldots, 60$ days, we compute a point forecast $\hat{P}_{t+\tau|t}$ of the daily PJM Dominion Hub price $P_{t+\tau}$ based on the information available at time t. The forecasting performance is compared using the mean absolute percentage error:

$$\text{MAPE}(\tau) = \frac{1}{T} \sum_{t=1}^{T} \frac{\left| \hat{P}_{t+\tau|t} - P_{t+\tau} \right|}{P_{t+\tau}}, \tag{6}$$

with $T = 717$ days, which corresponds to the out-of-sample test period from 1.1.2012 to 17.12.2013. Note that unlike in many other electricity price forecasting studies (for a discussion see e.g. Weron 2006), using MAPE is not controversial here since—as can be seen in Fig. 1—the daily PJM Dominion Hub price P_t is significantly above zero in the considered time period. Moreover, the main conclusions of our empirical study also hold if the mean absolute error (MAE) or the root mean square error (RMSE) are taken into account.

The point forecasts are evaluated on the basis of the Diebold–Mariano (DM) test, see Diebold and Mariano (1995). It allows to compare pairs of models and indicates, which of the two statistically outperforms the other. It may happen that, although one of the models has a lower MAPE, the differences between competing models are so small that they are statistically insignificant.

For each forecasting technique, we calculate the loss differential series $d_t = L(\varepsilon_{Model,t}) - L(\varepsilon_{Benchmark,t})$, with the loss function $L(\varepsilon_t) = |\varepsilon_t|/P_t$. We then conduct the DM tests for significance of differences. Note that we perform one-sided DM tests, with the null hypothesis $H_0 : E(d_t) \leq 0$. Hence, when the p value is smaller than the chosen significance level (e.g. $\alpha = 5\%$), we can conclude that the proposed model is better than the benchmark and when the p value is larger than $1 - \alpha$ (e.g. 95%) the opposite holds.

4.2 The forecasting scheme

We estimate model parameters using information provided by a rolling calibration window of a constant length. The window spans 4 years—initially from 1.1.2008 to 31.12.2011. For each daily forecast, we roll the calibration window forward by 1 day to ensure that all models are estimated on a sample of the same size. For instance, to forecast the price for 2.1.2012 the models are calibrated on data from the period 2.1.2008–1.1.2012. We use the last 2 years to evaluate the forecasting performance, see Figs 1 and 2. For each of the 717 days in the out-of-sample test period (from 1.1.2012 to 17.12.2013), the estimated parameters are used to compute the $\tau = 1, \ldots, 60$ step-ahead forecasts of daily prices for the PJM Dominion Hub.

Once the parameters of **AR**, see Eq. (1), and **AR-H**, see Eq. (2), models are estimated, the forecasts of future prices are computed sequentially. The hourly prices are aggregated into the daily ones by simple averaging. For the factor models, the procedure is more complicated. First, for each time window factors F_t and loads Λ_n are estimated from relation (3). Then, the factors are used to estimate the parameters of a vector autoregressive VAR(q) model, see Eq. (4). Once the models are estimated, the factor forecasts $\hat{F}_{t+\tau|t}$ are computed sequentially. Next, an analogous approach is applied to the estimated idiosyncratic components $\hat{v}_{kt}$. For each time window, the parameters of an autoregressive AR(q) model, see Eq. (5), are calibrated and used in sequential forecasting of future values of the idiosyncratic component $\hat{v}_{k,t+\tau|t}$. Finally, when both, common factors and idiosyncratic components, are predicted, they are used to estimate future values of the PJM Dominion Hub prices, according to formula (3). For **PC-HL** and **PC-H** models, the forecasts of the daily price are obtained by averaging the hourly price forecasts. The output of the **PC-L** model is already a daily price forecast.

4.3 Results

The models are compared for different forecast horizons: the first set of values ($\tau = 1, 2, \ldots, 7$) represents short-term forecasts, the second ($\tau = 14, 30, 45, 60$)

Table 3 MAPE errors for all five models and forecast horizons $\tau = 1, \ldots, 7, 14, 30, 45, 60$ days

τ	Benchmark	Multivariate models			
	AR (%)	AR-H (%)	PC-HL (%)	PC-H (%)	PC-L (%)
1	9.09	8.99	8.39	9.04	9.20
2	13.03	12.94	12.43	12.93	12.63
3	14.83	14.66	14.35	14.65	14.26
4	15.35	15.18	15.04	15.13	14.80
5	15.62	15.48	15.43	15.38	15.08
6	15.80	15.71	15.62	15.56	15.36
7	15.92	15.93	15.82	15.65	15.53
14	18.95	18.99	18.34	18.32	18.39
30	23.61	23.61	22.15	22.63	22.81
45	26.38	26.43	24.08	25.57	25.55
60	29.13	29.17	26.49	28.53	28.13

The out-of-sample test period includes 717 days, from 1.1.2012 to 17.12.2013. Compare with Fig. 3

corresponds to mid-term forecasts. Generally, models which explore the structure of the market, should perform better for longer forecast horizons.

The point forecasting results are summarized in Table 3 where MAPE errors for all five models and a selection of forecast horizons ranging from 1 to 60 days are presented. In Fig. 3 the difference between each model's MAPE errors and the MAPE for the benchmark **AR** model are plotted. Values lower than zero indicate a better forecasting performance with respect to the benchmark. Conversely, values higher than zero indicate a worse forecasting performance with respect to the **AR** model.

All multivariate factor models perform better than the benchmark for all reported forecast horizons τ, except for one case—one step-ahead predictions of the **PC-L** model, which is calibrated to daily prices from the 19 PJM zones. The richest factor model, i.e. **PC-HL**, is the best in one step-ahead predictions, better by 0.7 % than the
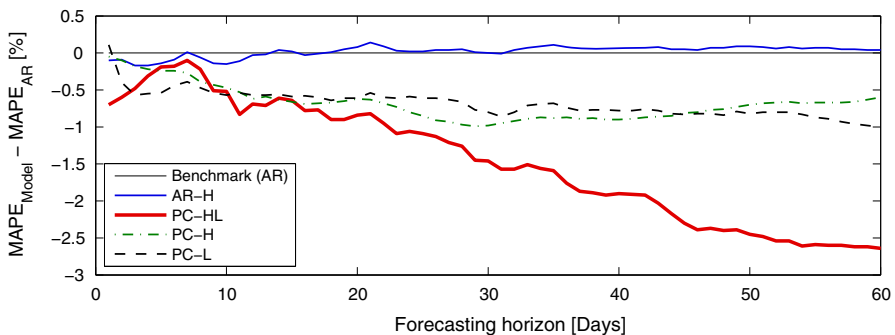

Fig. 3 The difference in percent between each model's MAPE errors and the MAPE for the benchmark **AR** model. Values lower than zero indicate a better forecasting performance with respect to the benchmark. See also Table 3

Table 4 Diebold–Mariano test p values for point forecasts

τ	AR versus	AR	AR	AR	AR-H	AR-H	AR-H	PC-HL	PC-HL	PC-H
	AR-H	PC-HL	PC-H	PC-L	PC-HL	PC-H	PC-L	PC-H	PC-L	PC-L
1	0.055	0.006	0.273	0.828	0.012	0.745	0.971	0.991	0.998	0.897
2	0.091	0.017	0.132	0.000	0.030	0.473	0.005	0.961	0.763	0.010
3	0.005	0.045	0.016	0.000	0.132	0.473	0.001	0.849	0.370	0.002
4	0.007	0.141	0.003	0.000	0.314	0.263	0.002	0.612	0.201	0.009
5	0.021	0.253	0.002	0.000	0.432	0.109	0.002	0.430	0.115	0.015
6	0.092	0.270	0.002	0.000	0.384	0.030	0.004	0.409	0.182	0.072
7	0.555	0.373	0.001	0.002	0.358	0.000	0.002	0.286	0.168	0.194
14	0.742	0.026	0.000	0.000	0.016	0.000	0.000	0.476	0.554	0.690
30	0.479	0.000	0.000	0.000	0.000	0.000	0.000	0.929	0.973	0.914
45	0.740	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.430
60	0.715	0.000	0.000	0.000	0.000	0.000	0.000	1.000	1.000	0.001

Note that when the p value is smaller than the chosen significance level (e.g. $\alpha = 5\,\%$), we can conclude that the proposed model (*lower row*) is better than the reference model (*upper row*) and when the p value is larger than $1 - \alpha$ (e.g. 95 %) the opposite holds

benchmark. It also beats all competitors for forecast horizons of 15 days or more. In particular, for $\tau \geq 52$ days it is better than the benchmark by over 2.5 %. The gains from using the two other factor models are less spectacular—they oscillate between 0.5 and 1 % for $\tau \geq 10$ days. **PC-H** is slightly better in the intermediate range between 12 and 44 days, while **PC-L** for $\tau = 2, .., 11$ and $\tau \geq 45$ days.

On the other hand, the simple restricted vector autoregressive model **AR-H** is slightly better than the benchmark only in the short-term. For almost all horizons in excess of two weeks it is slightly worse than the benchmark. This supports our hypothesis that knowledge about the intra-day (*from hourly day-ahead prices*) and/or inter-zone correlation structure of the electricity prices helps to forecast in the long run.

The Diebold–Mariano test p values for point forecasts are presented in Table 4. When the p value is smaller than the chosen significance level (e.g. $\alpha = 5\,\%$), we can conclude that the proposed model (model names in the *lower row*) is better than the reference model (model names in the *upper row*) and when the p value is larger than $1 - \alpha$ (e.g. 95 %) the opposite holds. The richest **PC-HL** model significantly outperforms the benchmark **AR** model at the 5 % level for $\tau \leq 3$ and $\tau \geq 9$ days. At the same time it is never outperformed by the benchmark. **PC-H** significantly outperforms the benchmark at the 5 % level for $\tau \geq 3$ days, but never significantly outperforms **PC-HL**, even at the 10 % level. The factor model calibrated to inter-zone data, i.e **PC-L**, behaves very much alike.

The restricted vector autoregressive model **AR-H** significantly outperforms the benchmark at the 5 % level only for $\tau = 3, 4, 5, 9, 10$; at the same time it significantly underperforms at the same level for $\tau \geq 21$ days. **AR-H** is generally also significantly

worse at the 5 % level than the factor models. Except for the one step-ahead predictions and the **PC-L** model it never outperforms the factor models, even at the 10 % level.

5 Conclusions

In this paper we have examined whether using intra-day (from hourly day-ahead prices) and/or inter-zone data can improve forecasts of daily day-ahead electricity prices for the PJM Dominion Hub. As a benchmark we have used a univariate autoregressive AR model (of order $q = 7$ and calibrated to daily data). The multivariate competitors include a restricted vector autoregressive model and three factor models. The largest factor model is calibrated to a panel of hourly day-ahead prices for 20 locations in the PJM market (19 zones and one hub, 480 variables in total), the intermediate model uses hourly day-ahead prices for the PJM Dominion Hub (24 variables), whereas the smallest model utilizes only daily prices, but across 20 locations (20 variables).

The results show that all three considered multivariate factor models perform better than the benchmark for all reported forecast horizons τ, except for one case—one step-ahead predictions of the PC-L model, calibrated to daily prices from the 19 PJM zones. In the mid-term the restricted VAR fails to provide accurate price predictions, however, the gains from using the richest factor model (calibrated to hourly zonal prices) are even more visible. Moreover, all three factor models provide improvement over the restricted VAR model for forecast horizons 2 days or more. This indicates that exploring the intra-day and/or inter-zone structure of electricity prices leads not only to more precise mid-term forecasts, but also to more precise short-term price forecasts of daily spot prices. On the other hand, only a joint exploration of both, the hourly and zonal structure allows to obtain much better longer term (15 days or more) predictions.

Our results are in line with the findings of Dempster et al. (2008), who reported inter-zonal Granger causality between electricity prices in 11 markets of the WSCC region (Western U.S.), and Maciejowska and Weron (2013) and Raviv et al. (2013), who stressed the importance of considering disaggregated (hourly) electricity prices for forecasting of aggregated (daily) prices. In a general forecasting setup, Stock and Watson (2002) showed that using the information contained in a large panel of data led to better forecasts of the individual series. To some extent this explains why the factor model calibrated to the richest panel (Panel-HL, see Tables 1, 2) performs better than the models calibrated to the smaller panels (Panel-H and Panel-L). However, as Lütkepohl (2011) warns, the inclusion of too many disaggregates can result in estimation error and specification error which ultimately lead to an efficiency loss. Apparently in our case this has not happened.

Despite the very recent inflow of relevant publications, the literature on the application of multivariate models to forecasting electricity prices is still relatively scarce. This paper makes an important contribution by showing that hourly and zonal prices can be efficiently used to forecast average daily prices for one of the major hubs in the PJM market, both in the short- and in the mid-term horizons. This study can be extended in a number of ways. Firstly, other linear models and model specifications can be used: ARMAX and VARMAX models to incorporate fundamental variables,

like fuel prices, reserve margin data or weather variables, separate factors for peak and off-peak hours, etc. Secondly, other dimension reduction approaches can be taken, for instance, utilizing dynamic semiparametric factor models (DSFM; see e.g. Park et al. 2009). Thirdly, more sophisticated and more realistic trend-seasonal components can be considered, as they may lead to a yet better performance in the longer term (see e.g. Nowotarski et al. 2013). Finally, interval forecasts can be computed to provide more valuable information for power market participants.

References

Bai J (2003) Inferential theory for factor models of large dimensions. Econometrica 71(1):135–171

Bai J, Ng S (2002) Determining the number of factors in approximate factor models. Econometrica 70(1):191–221

Bermingham C, D'Agostino A (2014) Understanding and forecasting aggregate and disaggregate price dynamics. Empir Econ 46:765–788

Bernhardt C, Klüppelberg C, Meyer-Brandis T (2008) Estimating high quantiles for electricity prices by stable linear models. J Energy Markets 1:3–19

Bierbrauer M, Menn C, Rachev ST, Trück S (2007) Spot and derivative pricing in the EEX power market. J Bank Financ 31:3462–3485

Chan K, Gray P (2006) Using extreme value theory to measure value-at-risk for daily electricity spot prices. Int J Forecast 22:283–300

Chen J, Deng S-J, Huo X (2008) Electricity price curve modeling and forecasting by manifold learning. IEEE Trans Power Syst 23:877–888

Conejo AJ, Contreras J, Espínola R, Plazas MA (2005) Forecasting electricity prices for a day-ahead pool-based electric energy market. Int J Forecast 21:435–462

Dempster G, Isaacs J, Smith N (2008) Price discovery in restructured electricity markets. Resourc Energy Econ 30:250–259

Diebold FX, Mariano RS (1995) Comparing predictive accuracy. J Bus Econ Stat 13:253–263

Elattar EE (2013) Day-ahead price forecasting of electricity markets based on local informative vector machine. IET Gener Trans Distrib 7:1063–1071

Eydeland A, Wolyniec K (2003) Energy and power risk management. Wiley, Hoboken

Garcia-Martos C, Rodriguez J, Sanchez M (2012) Forecasting electricity prices by extracting dynamic common factors: application to the Iberian market. IET Gener Trans Distrib 6:11–20

Härdle W K, Trück S (2010) The dynamics of hourly electricity prices. SFB 649 Discussion Paper 2010–013

Hendry DF, Hubrich K (2011) Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate. J Bus Econ Sta 29:216–227

Higgs H (2009) Modelling price and volatility inter-relationships in the Australian wholesale spot electricity markets. Energy Econ 31:748–756

Janczura J, Trück S, Weron R, Wolff R (2013) Identifying spikes and seasonal components in electricity spot price data: a guide to robust modeling. Energy Econ 38:96–110

Kaminski V (2013) Energy markets. Risk Books, London

Koopman S, Ooms M, Carnero A (2007) Periodic seasonal Reg-ARFIMA–GARCH models for daily electricity spot prices. J Am Stat Assoc 102:16–27

Kristiansen T (2012) Forecasting Nord Pool day-ahead prices with an autoregressive model. Energy Policy 49:328–332

Liebl D (2013) Modeling and forecasting electricity spot prices: a functional data perspective. Ann Appl Stat 7:1562–1592

Lütkepohl H (2005) New introduction to multiple time series analysis. Springer, Berlin

Lütkepohl H (2011) Forecasting nonlinear aggregates and aggregates with time-varying weights. Jahrbücher für Nationalökonomie und Statistik 231:107–133

Maciejowska K, Weron R (2013) Forecasting of daily electricity spot prices by incorporating intra-day relationships: evidence form the UK power market. In: IEEE Conference Proceedings—EEM13. doi:10.1109/EEM.2013.6607314

Misiorek A, Trück S, Weron R (2006) Point and interval forecasting of spot electricity prices: linear vs. non-linear time series models. Stud Nonlinear Dyn Econom 10(3), Article 2

Nowotarski J, Tomczyk J, Weron R (2013) Robust estimation and forecasting of the long-term seasonal component of electricity spot prices. Energy Econ 39:13–27

Park BU, Mammen E, Härdle W, Borak S (2009) Time series modelling with semiparametric factor dynamics. J Am Stat Assoc 104(485):284–298

Perevalov N, Maier P (2010) On the advantages of disaggregated data: insights from forecasting the U.S. economy in a data-rich environment. Bank of Canada, Working Paper 2010–10

Powell JL (1994) Estimation of semiparametric models. In: Engle R, McFadden D (eds) Handbook of econometrics, vol IV. Elsevier, Amsterdam, pp 2444–2521

Raviv E, Bouwman KE, van Dijk D (2013) Forecasting day-ahead electricity prices: utilizing hourly prices. Tinbergen Institute Discussion Paper 13–068/III. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2266312

Schlueter S (2010) A long-term/short-term model for daily electricity prices with dynamic volatility. Energy Econ 32:1074–1081

Stock JH, Watson MW (2002) Forecasting using principal components from a large number of predictors. J Am Stat Assoc 97(460):1167–1179

Vilar JM, Cao R, Aneiros G (2012) Forecasting next-day electricity demand and price using nonparametric functional methods. Electr Power Energy Syst 39:48–55

Weron R (2006) Modeling and forecasting electricity loads and prices: a statistical approach. Wiley, Chichester

Weron R (2014) Electricity price forecasting: a review of the state-of-the-art with a look into the future. Int J Forecast. 30(4). doi:10.1016/j.ijforecast.2014.08.008

Weron R, Misiorek A (2008) Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models. Int J Forecast 24:744–763

Wu H, Chan S, Tsui K, Hou Y (2013) A new recursive dynamic factor analysis for point and interval forecast of electricity price. IEEE Trans Power Syst 28:2352–2365