

ERC Working Papers in Economics 21/01

April / 2021

**Application of Bagging in Day-Ahead Electricity Price
Forecasting and Factor Augmentation**

Kadir Özen

Barcelona Graduate School of Economics, Barcelona, Spain

E-mail: kadir.ozen@barcelonagse.eu

Dilem Yıldırım

Middle East Technical University, Department of Economics, Ankara, Turkey

E-mail: dilem@metu.edu.tr

Application of Bagging in Day-Ahead Electricity Price Forecasting and Factor Augmentation

Kadir Özen*

Dilem Yildirim

Barcelona Graduate School of Economics

Department of Economics

Middle East Technical University

Abstract

The electricity price forecasting (EPF) is a challenging task not only because of the uncommon characteristics of electricity but also because of the existence of many potential predictors with changing predictive abilities over time. Particularly, how to account for all available factors and extract as much information as possible is the key to the production of accurate forecasts. To address this long-standing issue in a way that balances complexity and forecasting accuracy while facilitating the traceability of the predictor selection procedure, the method of Bootstrap Aggregation (bagging), which is a variant shrinkage estimation approach for the estimation of large scale models, is proposed in this paper. To forecast day-ahead electricity prices in a multivariate context for six major power markets we construct a large scale pure-price model (in addition to some stochastic models that are commonly applied in the literature) and apply the bagging approach in comparison with

*Corresponding author.

Kadir Özen: *Barcelona Graduate School of Economics, Ramon Trias Fargas, 25-27, 08005 Barcelona, Spain. kadir.ozen@barcelonagse.eu*, **Dilem Yildirim:** *Middle East Technical University, Department of Economics, 06800 Ankara, Turkey. dilem@metu.edu.tr*.

the popular Least Absolute Shrinkage and Selection Operator (LASSO) estimation method. Our forecasting study reveals that with its superior forecasting performance and its computationally simple algorithm, the bagging emerges as a strong competitor to the commonly applied LASSO approach for the short-term EPF. Further analysis for the variable selection for the bagging and LASSO approaches suggests that the differentiation in the forecast performances of two approaches might be due to, inter alia, their structural differences in the explanatory variables selection process. Moreover, to account for the intraday hourly dependencies of day-ahead electricity prices, all our models are augmented with latent factors, and a substantial improvement is observed only in the forecasts from models covering a relatively limited number of predictors, while almost no improvement is obtained in the forecasts from the large scale model estimated through LASSO and bagging techniques.

JEL Classification: C22; C38; C51; C53; Q47.

Keywords: Bagging; Shrinkage methods; Electricity price forecasting; Multivariate modeling; Forecast encompassing; Factor models.

1 Introduction

The dynamics of electricity price formation have become a complex phenomenon due to the power market liberalization processing since the early 1990s. With the ongoing liberalization process, the traditionally government-controlled power sectors have turned into open and competitive environments where commodities are bought, sold, and traded under specific market rules. Compared to other commodities, electricity is a very special commodity due to its rare characteristics. It has a non-storable nature, which requires a constant balance between production and consumption, and eliminates any possibility of inventories to be used to create any price arbitrages ([Shahidehpour et al. \(2002\)](#); [Weron \(2014\)](#)). Secondly, although price elasticity of electricity demand has displayed a slight increase over recent years probably due to amplified energy-saving awareness, electricity demand is price inelastic in the short-run since most consumers are indifferent to or unaware of the current electricity price ([Weron and Misiorek \(2008\)](#); [Zhu et al.](#)

(2018)). Moreover, the price of electricity is extremely dependent on the weather conditions and consumption time, e.g. hour of the day, day of the week, and time of the year (Weron (2014); Lago et al. (2018)). High intensity of business and everyday activities boost electricity demand and prices. Finally, electricity prices are subject to sudden and unexpected short-lived spikes due to factors including transmission problems, real-time supply-demand balancing, and the increasing penetration of renewable energy resources, which stimulates the dependency of electricity prices on weather conditions (G. P. and S. (2013); Lago et al. (2018)). All these uncommon characteristics have complicated electricity price formation and make accurate and reliable short-term electricity price forecasts quite valuable for number of institutions spanning from power plant operators to market operators and from transmission system planners to power portfolio managers.

Over the past decade, many different sophisticated approaches have been proposed to forecast day-ahead electricity prices (for a comprehensive review, see Aggarwal et al. (2009); Weron (2014)). Among them, time series models and regression analysis appear to be the most acclaimed modeling techniques in the electricity price forecasting (EPF) literature. Commonly applied time series models include not only the basic autoregressive (AR) and autoregressive moving average (ARMA) specifications but also AR and ARMA models with exogenous variables (ARX, ARMAX) (Weron and Misiorek (2008); Kristiansen (2012)) or with nonlinear dynamics (Crespo Cuaresma et al. (2004); Misiorek et al. (2006)), regime-switching models (Mount et al. (2006); Kanamura and Ohashi (2008)), vector autoregressions (VAR) (Ziel et al. (2015)), generalized autoregressive conditional heteroscedasticity (GARCH) based models (Garcia et al. (2005); Knittel and Roberts (2005)) and so on. However, the most challenging issue in the electricity price forecasting is the existence of many potential predictors, which complicates the identification of the most influential and relevant factors. This issue becomes more crucial especially in out-of-sample forecasting exercises where relevant predictors might turn into predictors with less or even no forecasting power. To tackle this issue, one of the most commonly used strategies is to utilize shrinkage estimation methods. These methods facilitate not only to account for all available factors and extract as much information as possible rather than being confined to a specific subset of predictors selected through traditional variable selection methods but also to track the predictor selection procedure.

In this respect, the **Least Absolute Shrinkage and Selection Operator** (LASSO) due to [Tibshirani \(1996\)](#), became a quite popular estimation tool in the EPF literature for handling models with many potential factors (e.g. [Ziel et al. \(2015\)](#); [Ludwig et al. \(2015\)](#); [Uniejewski et al. \(2016, 2018\)](#)). Furthermore, more recently, in an extensive empirical electricity price forecasting exercise, [Ziel and Weron \(2018\)](#) and [Uniejewski et al. \(2019\)](#) have suggested that the LASSO appears to be the best model compared to various other benchmark models. However, despite its highly improved forecast accuracy, the LASSO is subject to a number of limitations including the need for a dedicated numerical optimization procedure and its considerable dependency on the regularization parameter. Due to that dependency, a grid of regularization parameters might be required to specify the best performing model, which, in turn, can exacerbate the already complicated optimization procedure ([Uniejewski et al. \(2019\)](#)).

An alternative strategy, which is the main focus of this paper, is to use the method of **Bootstrap Aggregation** (bagging). Bagging is a statistical method originally developed by [Breiman \(1996a,b\)](#) and its theoretical advancements are well established by studies including [Hall et al. \(1995\)](#), [Bühlmann and Yu \(2002\)](#), [Andrews \(2004\)](#), [Friedman and Hall \(2007\)](#), and [Lee et al. \(2010\)](#). Fundamentally, as described by ([Inoue and Kilian, 2008](#), p. 551): “Bagging involves generating a large number of bootstrap resamples of the original forecasting problem, applying a pre-test model selection rule to each of the resamples, and averaging the forecasts from the models selected by the pre-test on each bootstrap sample”. Averaging across resamples eliminates the instability of the model selection decision rules, which are inherently sensitive to small alterations in the data, and allow one to harvest as much information as possible from available predictors. Averaging the parameter estimates across bootstrap samples in each time period also address the parameter instability problem. Although a number of recent studies suggest that bagging is a promising tool for improving forecast accuracy of economic and financial variables such as inflation, employment growth and stock returns (e.g., [Inoue and Kilian \(2008\)](#); [Rapach and Strauss \(2010\)](#); [Jordan et al. \(2017\)](#)), to the best of our knowledge, bagging has not been applied in the EPF literature so far¹. This finding is interesting because

¹There are also some complex algorithms in the scope of machine learning applications of the electricity price forecasting which include bagging as a tool but share a perspective that is completely different

bagging, which is asymptotically a shrinkage forecast (Huang and Lee (2010); Jin et al. (2014); Stock and Watson (2012)), might manifest itself as a computationally simple one to overcome the issue of having many potential predictors with changing predictive abilities over time as well as balancing the complexity and forecasting accuracy while facilitating the traceability of the predictor selection procedure.

In that sense, the contribution of our paper is twofold. Our first and principal contribution is to provide a thorough analysis of the use of the bagging technique for forecasting day-ahead electricity prices in a multivariate setting. To provide an extensive empirical study, we consider six datasets from key electricity markets and perform bagging in comparison with a number of stochastic models that are commonly applied in the EPF literature. These models include the **mean_{now}** model that is based on the mean values of the past prices, the **naive** model of Nogales et al. (2002) that is based on the similar-day principle and simple Autoregressive (AR) models. Besides these three benchmark models, we employ **EXPERT** model, which is originally proposed by Misiolek et al. (2006) and manifest itself being abstract and parsimonious while including AR, periodic and nonlinear effects together. Finally, we construct a large scale pure-price model and apply the bagging approach in comparison with the popular LASSO estimation method. We explore the performance of the bagging technique in comparison to all constructed models based on different forecast performance evaluation criteria and Diebold-Mariano tests, which facilitate us to investigate whether bagging can significantly improve the forecast accuracy and any structural differences in terms of mostly selected predictors with bagging and LASSO².

Our second contribution is addressing an important consideration that is often overlooked in forecasting day-ahead electricity prices in a multivariate framework. Unlike the univariate fashion, the multivariate modeling approach, where each hour of the day is modeled separately, is motivated by the fact that in many day-ahead electricity markets, continuous trading is not allowed; instead, agents are supposed to submit their

than that we present here both in terms of methodology and the data. Due to the breadth of the EPF literature, a detailed discussion of those studies is not provided in this paper but the interested reader is referred to the most recent study of Agrawal et al. (2019).

²The replication material and the forecasting toolbox in Gauss Aptech Programming language can be downloaded from [WillBeAvailableSoon](#).

bids and offers for delivery of electricity for all hours of the next day before a specific market closing time. Hence, the prices of all the hours of the next day are determined simultaneously at once. Although the multivariate framework is more coherent with the data generation of electricity prices, its one important limitation is that it lacks intra-day hourly dependencies of day-ahead electricity prices. In the EPF literature, there is only a small number of studies accounting for these dependencies. In that sense, [Vehviläinen and Pyykkönen \(2005\)](#), [Härdle and Trück \(2010\)](#), [Nowotarski et al. \(2014\)](#), [Raviv et al. \(2015\)](#), [Maciejowska and Weron \(2015, 2016\)](#) and [Ziel \(2016\)](#) address that dependence through factor models, which have been broadly utilized in various economic and financial applications but have not been exploited in the EPF literature, except above-mentioned few studies. Our modeling and implementation strategy for factor models, however, differs from existing studies in various dimensions. In our application, we extend our models to incorporate the intra-day hourly dependencies of day-ahead electricity prices by augmenting all models using factors estimated through a Principal Component Analysis (PCA) with data dependent selection procedures. Then, we explore whether factor-augmentation can provide a further improvement in the forecast accuracy.

The remainder of the paper is organized as follows. In Section 2, we describe our data sets and the data transformation method we employed to stabilize the variance of day-ahead hourly electricity prices. Then, in Section 3, after providing a comprehensive discussion of the forecasting models including the bagging technique and factor augmentation, we review forecast performance evaluation criteria and the Diebold-Mariano test designed to explore significant differences in the forecasting performances. In Section 4, we present the empirical results, discuss forecast performances of the models, and provide the occurrence tables which indicate the number of times a specific predictor is selected in each model. Section 5 concludes the paper.

2 Data

Our dataset, summarized in [Table 1](#), consists of hourly data for six electricity day-ahead price series from five major power markets, including Nordic Power Exchange Nord Pool for system price (NP.SYS) and for United Kingdom (NP.N2EX), Common-

wealth Edison (ComEd) zone in the Pennsylvania-New Jersey-Maryland (PJM) market (PJM.COMED), OTE and OMIE which manage the Czech Republic (OTE.CZ) and Iberian market (OMIE.SP), respectively. Our last dataset comes from the price track of the Global Energy Forecasting Competition 2014 (GEFCom2014), which provides locational marginal prices at an hourly resolution. While the source of the data is not publicized by the organizers of the competition, the reader is referred to [Hong et al. \(2016\)](#) for details of this series. The GEFCom2014 dataset covers a 3-year period from January 1, 2011 to December 17, 2013, the remaining datasets cover a 6-year and 9-month period from January 1, 2013 to September 19, 2019.

Table 1: Summary of the day-ahead electricity price series

Electricity Market	Acronym	#of Data Points	oos	Source
Nord Pool (system price)	NP.SYS	58872/2453	35280/1470	nordpoolgroup.com
Nord Pool,UK	NP.N2EX	58872/2453	35280/1470	nordpoolgroup.com
PJM,USA	PJM.COMED	58872/2453	35280/1470	dataminer2.pjm.com
OTE price for the Czech Republic	OTE.CZ	58872/2453	35280/1470	ote-cr.cz
OMIE price for Spain	OMIE.SP	58872/2453	35280/1470	m.omie.es
GEFCom2014 competition data	GEFCom2014	25968/1082	8400/350	Hong et al. (2016)

Note: The table reports the summary of the day-ahead electricity price series considered in the paper. The total number of data points (# of Data Points) and the length of the out-of-sample forecast period (oos) are given in univariate/multivariate settings, respectively. The GEFCom2014 dataset covers a 3-year period from Jan 1, 2011 to Dec 17, 2013, the remaining datasets – a 6-year and 9-month period from Jan 1, 2013 to Sep 19, 2019. NP.N2EX price is in terms of GBP/MWh, PJM.COMED and GEFCom2014 is in terms of USD/MWh and remaining series are in terms of EUR/MWh.

In our analysis, approximately the first three (two) years are set as the estimation or training period for five major markets (GEFCom2014) (i.e. from 1 January 2013 to 10 September 2015 for five major markets and from 1 January 2011 to 1 January 2013 for GEFCom2014). Once all considered models are estimated using data from the training period, we produce forecasts for all 24 h of the next day. Then, keeping the estimation window length constant, which is 983-day and 732-day for five major markets and GEFCom, respectively, the window is rolled forward by one day and we reestimate all models and predict each hour of the second day. This process is repeated until the forecasts for the last day of the out-of-sample period (19 September 2019 for five major markets and 17 December 2013 for GEFCom2014) are computed. Before conducting our empirical analysis, our price series (excluding GEFCom2014) are preprocessed to account

for missing values and changes to/from daylight-saving-time. In that sense, being in line with many existing studies, including [Weron \(2006\)](#), [Nowotarski et al. \(2014\)](#), and [Uniejewski et al. \(2016\)](#), the missing values (including the clock change in March) are substituted by the simple average of two neighbor hours. The ‘doubled’ hours (including the clock change in October) are also averaged and substituted for the corresponding hour.

2.1 Data transformation

Similar to many electricity price series, our hourly prices exhibit price spikes and high volatility. This critical issue requires careful consideration since it might render the observed model estimates and statistical inferences derived from those estimates prone to severe problems and hence reduce the forecast accuracy. Compared to other forecasting methodologies, this problem could be much more substantial for our bagging technique which relies on pre-test model selection rules that require reliable statistical inferences. To reduce spike severity and stabilize the variance of the electricity price series, many transformation methods have been proposed (for a comprehensive discussion of data transformation techniques utilized in EPF literature, see [Uniejewski et al. \(2018\)](#)). Among them, the logarithmic transform is the most popular approach due to its simplicity. However, the log-transform is not feasible when dealing with our datasets including negative or zero values, which is an increasingly observed phenomenon in electricity price series due to the increased penetration of renewable energy in electricity generation with negligible electricity generation costs.

In our analysis, similar to [Ziel and Weron \(2018\)](#), we prefer to employ the area (or inverse) hyperbolic sine (asinh) transformation which has a logarithmic tail behavior like the log-transform but can handle both positive and negative values. In that respect, given that $Y_{d,h}$ represents the original day-ahead price at day d and hour h , the asinh transformation is given as:

$$y_{d,h} = \operatorname{asinh}(\tilde{Y}_{d,h}) = \log(\tilde{Y}_{d,h} + \sqrt{\tilde{Y}_{d,h}^2 + 1}) \quad (1)$$

where $y_{d,h}$ denotes the asinh-transformed price $\tilde{Y}_{d,h}$ is the normalized price obtained from

the median normalization by using the formula:

$$\tilde{Y}_{d,h} = \frac{Y_{d,h} - a}{b} \quad (2)$$

where a and b are called shift and scale parameter, respectively. The shift parameter, a , is the median of the estimation period and the scale parameter, b , is the median absolute deviation (MAD) around the sample median in the estimation period adjusted by $\frac{1}{z_{0.75}}$. This adjustment is required to ensure asymptotically normal consistency of the standard deviation. Once asinh-transformed prices ($y_{d,h}$) are computed, we estimate all our models using the transformed data, and derive the forecasts. After computation of forecasts $\hat{y}_{d,h}$, inverse transformation (hyperbolic sine) is applied to derive the day-ahead electricity price forecasts as:

$$\hat{Y}_{d,h} = b \cdot \sinh(\hat{y}_{d,h}) + a \quad (3)$$

3 Econometric methodology

In this section, we describe all models we employed to forecast day-ahead electricity prices. In that sense, we start with the overview of three benchmark models (the **mean_{how}**, **naive**, and **AR** models), move on to **EXPERT** models and the large scale pure-price model with LASSO approach, continue with the introduction of the bagging procedure and after that with the factor augmentation that aims to exploit the intra-day dependencies of prices. Finally, we conclude this section with the discussion of evaluation metrics that are needed to evaluate and compare forecast performances.

Before starting to discuss the forecast models, it is worth mentioning that we implement a multivariate modeling framework where each hour of the day $h = 1, \dots, 24$ is treated as a separate series and the prices for all hours of the next day are forecasted at once using rolling window. The multivariate approach, which is generally favored in studies for short-term predictions (e.g. [Nowotarski et al. \(2014\)](#); [Weron \(2014\)](#)), is inspired by the facts that in many day-ahead electricity markets there is one-auction each day for the next 24 hours and each hour has a distinct price profile due to differences in hourly demand, costs and operational constraints. However, it ignores potential interdependen-

cies between the hours of the day, which suggests that the price at hour j depends not only on the price of previous days at hour j but could also depend on a different hour i . In that sense, while one might be in favor of the use of the univariate approach, which has the advantage of incorporating that dependence by setting a single model for the entire data and performing a recursive scheme to produce forecasts, it has the drawback that it could produce inaccurate forecasts depending on the magnitude of the effect of accumulation of errors, arising during the recursive scheme (Ziel and Weron (2018)). For that reason, our modeling is implemented in a multivariate framework and further factor augmentations are introduced to evaluate whether or not accounting for interdependencies between the hours of the day provides any substantial improvement in the forecast accuracy in electricity prices.

3.1 Benchmark models

Our first conventional benchmark is the **mean_{how}** model. In this model, the forecast of the electricity price for day d and hour h is set to the weekly mean of hourly frequency of the asinh-transformed price ($\bar{y}_{how,d,h}$) observed for each 983-day (732-day for GEFCom2014) estimation window. Formally, the **mean_{how}** model is given as:

$$y_{d,h} = \bar{y}_{how,d,h} + \epsilon_{d,h} = \sum_{j=1}^{168} \beta_j how_{d,h}^j + \epsilon_{d,h} \quad (4)$$

where $\epsilon_{d,h}$ is the zero-mean iid disturbance term and $how_{d,h}^j$ is the *hour-of-the-week dummy* defined as:

$$how_{d,h}^j = \begin{cases} 1, & \text{if } 24(d-1) + h \text{ is the } j\text{-th hour of the week,} \\ 0, & \text{oth.} \end{cases} \quad (5)$$

where j stands for the hours of the week from 1 (Monday, $h = 1$) to 168 (Sunday, $h = 24$).

Our second benchmark model, denoted by **naive**, has been introduced to the EPF literature by Nogales et al. (2002) and belongs to the class of similar-day approaches of Weron (2014). Briefly, based on the similar-day principle, this model takes the previous

week's (the previous day's) same hour as the forecast price for Saturday, Sunday and Monday (for the rest of weekdays). As highlighted by many studies including [Uniejewski et al. \(2016\)](#), [Nogales et al. \(2002\)](#), [Conejo et al. \(2005\)](#), the forecast performance of the **naive** model can outperform some of the advanced models especially if they are poorly calibrated. Formally, we can write the **naive** model as:

$$Y_{d,h} = \begin{cases} Y_{d-7,h}, & \text{if } dow_{d,h}^l = 1 \text{ for } l = 0, 1, 6, \\ Y_{d-1,h}, & \text{oth.} \end{cases} \quad (6)$$

where $dow_{d,h}^l$ is the *day-of-the-week dummy* given as:

$$dow_{d,h}^l = \begin{cases} 1, & \text{if } d \text{ is the } l\text{-th day of the week,} \\ 0, & \text{oth.} \end{cases} \quad (7)$$

where l stands for the days of the week from 0 (Sunday) to 6 (Saturday).

The last benchmark model we employed is the popular simple AR model that models the linear dependence of each price to the corresponding hour of previous days. Being in line with [Ziel and Weron \(2018\)](#), we construct the AR model in the demeaned form by using the weekly mean of hourly frequency for asinh-transformed prices $\bar{y}_{how,d,h}$ as

$$y_{d,h} = \bar{y}_{how,d,h} + \psi_{h,0} + \sum_{i=1}^{p_h} \psi_{h,i} (y_{d-i,h} - \bar{y}_{how,d-i,h}) + \epsilon_{d,h} \quad (8)$$

where p_h is the required number of lags included in the model to ensure an iid structure for the disturbance term $\epsilon_{d,h}$. We denote this model by **AR_{how}** and select the appropriate lag length according to the Akaike Information Criterion (AIC) with a maximum autoregressive order of 8. That is, for each estimation window, we select the lag length that minimizes $AIC = \log(\frac{\mathbf{e}'\mathbf{e}}{n}) + 2\frac{k}{n}$, where \mathbf{e} is the vector of residuals obtained from ordinary least squares (OLS) estimation of the model, k is the total number of parameters in the model and n is the length of the estimation window.

3.2 Autoregressive expert model

Expert models have been originally proposed by [Misiorek et al. \(2006\)](#) and have gained momentum in the EPF literature, with several applications, including [Weron \(2006\)](#), [Weron and Misiorek \(2008\)](#), [Kristiansen \(2012\)](#), [Nowotarski et al. \(2014\)](#), [Ziel \(2016\)](#), [Maciejowska et al. \(2016\)](#), [Uniejewski et al. \(2016\)](#), [Uniejewski and Weron \(2018\)](#), and [Ziel and Weron \(2018\)](#). These models manifest themselves as being abstract and parsimonious and since they are built upon field knowledge of experts they are called as expert models. In our analysis, we employ the expert_{DoW,p,nl} model of [Ziel and Weron \(2018\)](#)³:

$$\begin{aligned}
 y_{d,h} = & \beta_{h,1} + \underbrace{\beta_{h,2}y_{d-1,h} + \beta_{h,3}y_{d-2,h} + \beta_{h,4}y_{d-7,h}}_{\text{autoregressive effects}} + \underbrace{\beta_{h,5}y_{d-1,\min} + \beta_{h,6}y_{d-1,\max}}_{\text{non-linear effects}} \\
 & + \underbrace{\beta_{h,7}y_{d-1,24}}_{\text{last-hour effect}} + \underbrace{\sum_{i=1}^6 \beta_{h,7+i} \text{dow}_{d,h}^i}_{\text{weekday dummies}} \\
 & + \underbrace{\sum_{i=1}^6 \beta_{h,13+i} \text{dow}_{d,h}^i y_{d-1,h} + \sum_{i=1}^6 \beta_{h,19+i} \text{dow}_{d,h}^i y_{d-1,24}}_{\text{periodic effects}} + \epsilon_{d,h}
 \end{aligned} \tag{9}$$

In this setting, $y_{d-1,h}$, $y_{d-2,h}$, and $y_{d-7,h}$ stand for the autoregressive effects of the previous day's same hour. $y_{d-1,\min} = \min_{h=1, \dots, 24} \{y_{d-1,h}\}$ and $y_{d-1,\max} = \max_{h=1, \dots, 24} \{y_{d-1,h}\}$ are minimum and maximum of previous day's hourly prices and, they are included to account for the previous day's extreme price levels. Since the minimum and maximum are nonlinear functions, these effects are called as the nonlinear effects. The lagged price, $y_{d-1,24}$, which represents the price of the last hour of the previous day, accounts for the potential dependence of early morning hours on the previous day's midnight price. Finally, the *day-of-the-week dummies* and periodic effects emphasize the short-term seasonality components of the hourly electricity prices. In this setup, one should note that for $h = 24$, there occurs a multicollinearity problem between $y_{d-1,24}$ and $y_{d-1,h}$, and, between $\text{dow}_{d,h}^i y_{d-1,h}$ and $\text{dow}_{d,24}^i y_{d-1,24}$ for $i = 1, \dots, 6$ and we handle this problem by

³This version is the the most generic version of the Expert model employed in the main text of [Ziel and Weron \(2018\)](#), please see Eq. (A.1) in their Appendix. Also see **mAR1hm** and **AR2hm** models of [Uniejewski et al. \(2016\)](#)

dropping the corresponding variables. Consequently, we estimate 25 parameters (18 for $h = 24$) through the application of OLS. We call this model **EXPERT**.

3.3 Large scale models and shrinkage estimation procedures

As mentioned before the most challenging issue in an electricity price forecasting exercise is the existence of many potential predictors. One way of selecting a set of the most informative predictors is to rely on expert knowledge and past experiences, as in the expert models. Selecting a subset of available predictors, however, may cause loss of information and unstable and unreliable forecasts due to the changing performance of the predictors over time (see e.g. [Stock and Watson \(2003\)](#)). Further forecast improvements can be gained by taking advantage of more predictors that are available on the market at the time of forecasting. In that sense, we utilize the following Large Scale pure-price model inspired by the $24lasso_{DoW,p,nl}^{HQC}$ model of [Ziel and Weron \(2018\)](#):

$$\begin{aligned}
y_{d,h} = & \underbrace{\sum_{i=1}^8 \sum_{j=1}^{24} \mu_{h,i,j,1} y_{d-i,j}}_{\text{price autoregressive effects}} + \underbrace{\sum_{i=1}^8 \mu_{h,i,1,2} y_{d-i,min} + \sum_{i=1}^8 \mu_{h,i,1,3} y_{d-i,max}}_{\text{price non-linear effects}} \\
& + \underbrace{\sum_{i=1}^7 \mu_{h,i,1,4} dow_{d,h}^i}_{\text{weekday dummies}} + \underbrace{\sum_{i=1}^7 \mu_{h,i,1,5} dow_{d,h}^i y_{d-1,avg}}_{\text{average price effects}} \\
& + \underbrace{\sum_{i=1}^6 \mu_{h,i,1,6} dow_{d,h}^i y_{d-1,h} + \sum_{i=1}^6 \mu_{h,i,1,7} dow_{d,h}^i y_{d-1,24}}_{\text{price periodic effects}} + \epsilon_{d,h}.
\end{aligned} \tag{10}$$

In this equation, the first component represents the autoregressive effects of all hourly prices for all days up to 8 days before the current day, the second and the third components together account for the previous days' extreme price levels, the fourth component represents the day-of-the-week effect, the fifth component stands for the average price effect ($y_{d,avg}$ is the average hourly day-ahead price at time d), last two terms describe the periodic effects. The model contains six components and 234 explanatory variables for $h = 1, \dots, 23$. For the last hour of the day, we drop the second term in “*price periodic effects*”, since it creates a multicollinearity with the first term, thus we have 228 variables

in total. For estimation of such a large scale model, we will utilize two different shrinkage estimation methods. The first one is the popular LASSO approach and the second one is the method of bagging, which has not been embodied in the EPF literature so far. In the next two subsections, we discuss the details of these estimation procedures.

3.3.1 The LASSO approach

The LASSO is a shrinkage estimation procedure, which estimates the full model with all predictors and then shrinks coefficients of less effective predictors towards zero by utilizing a specific algorithm. It has been applied in various electricity price forecasting exercises and showed superior forecasting performance compared to other conventional non-shrinkage methods used in the EPF literature.

To describe the Lasso approach, let us re-write the large scale model given in Eq. 10 in matrix notation as:

$$\mathcal{Y}_h = \mathcal{X}_h \boldsymbol{\beta}_h + \mathbf{u}_h \quad (11)$$

where \mathcal{Y}_h indicates an R -vector of observations for a given estimation window of R , $\mathcal{Y}_h = \begin{bmatrix} y_{d,h} & \dots & y_{d+R-1,h} \end{bmatrix}'$, \mathcal{X}_h denote an $R \times k$ matrix that consists of columns of explanatory variables, $\mathcal{X}_h = \begin{bmatrix} \mathbb{X}'_{d,h} & \dots & \mathbb{X}'_{d+R-1,h} \end{bmatrix}'$, where $\mathbb{X}_{d,h}$ is a k -vector of transformed explanatory variables at day d for a given hour h , $\boldsymbol{\beta}_h$ is a k -vector of coefficients with typical element β_i , $i = 1, 2, \dots, k$, and \mathbf{u}_h denote R -vector with typical iid disturbance $\epsilon_{d,h}$, $d = 1, 2, \dots, T - R$. Given this representation, the lasso estimator given by [Hastie et al. \(2015\)](#) has the form⁴:

$$\hat{\tilde{\boldsymbol{\beta}}}_{h,\lambda} = \underset{\tilde{\boldsymbol{\beta}}_h}{\operatorname{argmin}} \{ \|\tilde{\mathcal{Y}}_h - \tilde{\mathcal{X}}_h \tilde{\boldsymbol{\beta}}_h\|_2^2 + \lambda \|\tilde{\boldsymbol{\beta}}_h\|_1 \} \quad (12)$$

where $\tilde{\mathcal{Y}}_h$ and $\tilde{\mathcal{X}}_h$ be the columnwise standardized analogs of \mathcal{Y}_h and \mathcal{X}_h , respectively by setting the variance equal to one and the mean equal to zero. λ indicates the regularization parameter. For $\lambda = 0$, Eq. 12 turns into the objective function of the simple OLS estimation and we obtain the OLS estimator of $\tilde{\boldsymbol{\beta}}_{h,\lambda}$. As λ increases more coefficients

⁴ $\|\cdot\|_2$ is the usual Euclidean norm and $\|\tilde{\boldsymbol{\beta}}_h\|_1 = \sum_{i=1}^k \tilde{\beta}_i$

shrink to zero and in the limiting case, where $\lambda \rightarrow \infty$, all parameters are forced to be zero.

In application, selecting an appropriate value for λ is crucial as the LASSO estimation results are quite sensitive to the choice of the regularization parameter. We employ in-sample information criterion (*IC*) selection procedure. In that sense, following [Uniejewski et al. \(2019\)](#), we run the LASSO estimation for each value of $\lambda_i = 10^{-\frac{19-i}{6}}$ for $i = 1, \dots, 10$ and select the optimum parameter λ that yields the smallest Hannan-Quinn Information Criterion (HQC), $HQC = \mathbf{e}'\mathbf{e} + 2\log(\log(n))(k-l)\frac{\mathbf{e}'\mathbf{e}}{n-(k-l)}$, where \mathbf{e} is the vector of residuals obtained from the LASSO estimation of the model in Eq. 10, k is the total number of parameters in the model, l refers to the number of parameters that are set to be zero in the LASSO estimation and n is the length of the estimation window. In this in-sample procedure, HQC is preferred due to its superior performance compared to the AIC and Bayesian Information Criterion (BIC) ([Ziel and Weron \(2018\)](#)). Throughout our discussion, the LASSO estimation with the regularization parameter being selected according to HQC is denoted by **LASSO_LS^{HQC}**.

3.3.2 The bagging approach

In an environment where there are many potential predictors with changing predictive abilities over time, Bootstrap Aggregation, i.e. bagging procedure, can be considered as a computationally simpler alternative to the popular LASSO approach. The key idea of bagging is generating multiple information set through bootstrap resamples of the original data, which helps to improve the forecast accuracy as the group of forecasts produced by the resampled versions of the original data are expected to have similar biases but a reduced variance as illustrated by [Dantas and Cyrino Oliveira \(2018\)](#).

The bagging procedure begins with forming pre-test (PT) forecasts. To compute the PT forecasts, we first estimate the unrestricted model given in Eq. 11 for a given estimation window R . Then, let t_i denote the t -statistic for which null of β_i is equal to zero in the model, where β_i is equal to i -th element of $\boldsymbol{\beta}_h$, $i = 1, 2, \dots, k$, we perform two-sided t -tests on each predictor of the model by specifying a critical value t_c ⁵. We

⁵Being in line with many studies on bagging, including [Inoue and Kilian \(2008\)](#) and [Rapach and Strauss \(2010\)](#), we prefer to use t -statistics as the pre-test strategy. The t -statistics for the OLS estimates

form a $R \times l$ ($1 \leq l \leq k$) pretested (PT) predictor matrix \mathcal{X}_h^{PT} by discarding the i -th column of \mathcal{X}_h , if $|t_i| < t_c$ and estimate the following compact model by OLS⁶:

$$\mathcal{Y}_h = \mathcal{X}_h^{PT} \boldsymbol{\eta}_h + \boldsymbol{\zeta}_h \quad (13)$$

where $\boldsymbol{\eta}_h$ is an l -vector of coefficients with typical element η_j , $j = 1, 2, \dots, l$, and $\boldsymbol{\zeta}_h$ is an l -vector with typical iid disturbances. Then, PT forecast of $y_{d+1,h}$ that is based on the most recent information available on day d is denoted by $\hat{y}_{d+1,h|d}^{PT}$ and it is calculated as:

$$\hat{y}_{d+1,h|d}^{PT} = \begin{cases} 0, & \text{if } |t_i| < t_c \forall i, \\ \mathbb{X}_{d+1,h}^{PT \prime} \hat{\boldsymbol{\eta}}_h, & \text{oth.} \end{cases} \quad (14)$$

where $\hat{\boldsymbol{\eta}}_h$ is the OLS estimator of $\boldsymbol{\eta}_h$ in Eq. (13).

Once we derived the PT forecasts, we continue with constructing bootstrap aggregated forecasts through following algorithm:

- (i) Construct the $R \times (k + 1)$ matrix by combining \mathcal{Y}_h and \mathcal{X}_h as given below:

$$\begin{bmatrix} y_{d,h} & \mathbb{X}'_{d,h} \\ y_{d+1,h} & \mathbb{X}'_{d+1,h} \\ \vdots & \vdots \\ y_{d+R-1,h} & \mathbb{X}'_{d+R-1,h} \end{bmatrix}$$

- (ii) Following the overlapping blocking scheme of [Kunsch \(1989\)](#), generate bootstrap samples by drawing blocks of size m randomly with replacement from the above matrix

of Eq. (11) are computed using [Newey and West \(1987\)](#) heteroskedasticity and autocorrelation consistent (HAC) standard errors.

⁶In the empirical analysis, unless otherwise stated, we consider as t_c values, 2.58, 1.96, and 1.65 for 1%, 5% and, 10% significance levels, respectively.

and paste them end-to-end to form the block bootstrap samples⁷:

$$\begin{bmatrix} y_{d,h}^* & \mathbb{X}_{d,h}^{*'} \\ y_{d+1,h}^* & \mathbb{X}_{d+1,h}^{*'} \\ \vdots & \vdots \\ y_{d+R-1,h}^* & \mathbb{X}_{d+R-1,h}^{*'} \end{bmatrix}$$

(iii) For each bootstrap sample estimate Eq. (11), apply the pretesting procedure as explained above and estimate the Eq. (13) and compute the forecast. Forecast of each randomly generated sample is given in (15). That is, the forecast generated from each bootstrap sample has the form:

$$\hat{y}_{d+1,h|d}^{*PT} = \begin{cases} 0, & \text{if } |t_i^*| < t_c \forall i, \\ \mathbb{X}_{d+1,h}^{*PT} \hat{\boldsymbol{\eta}}_h^*, & \text{oth.} \end{cases} \quad (15)$$

where $\hat{y}_{d+1,h|d}^{*PT}$, $\mathbb{X}_{d,h}^{*PT'}$, t_i^* , and, $\hat{\boldsymbol{\eta}}_h^*$ are bootstrap analogs of $\hat{y}_{d+1,h|d}^{PT}$, $\mathbb{X}_{d,h}^{PT'}$, t_i , and, $\hat{\boldsymbol{\eta}}_h$.

(iv) Form the bagging forecast, $\hat{y}_{d+1,h|d}^{ba}$, as the expectation of the bootstrap pre-tested forecasts across bootstrap samples as:

$$\hat{y}_{d+1,h|d}^{ba} = \mathbb{E}^* \hat{y}_{d+1,h}^{*PT} \quad (16)$$

where \mathbb{E}^* denotes the expectation operator with respect to the bootstrap probability measure. In application, the expectation operator is replaced with its sample analog and the bagging forecast is derived as:

$$\hat{y}_{d+1,h|d}^{ba} = \frac{1}{B} \sum_{b=1}^B \hat{y}_{d+1,h,b|d}^{*PT} \quad (17)$$

⁷In time series application, data dependency is an important phenomenon which effects the performance of bagging. To address the possible data dependency problem block bootstrap is proposed instead of regular bootstrap. Given that m is the block length, in time series forecasting applications m is set equal to forecast horizon (Gonçalves and Kilian (2004)). In our empirical analysis, we forecast for the next day's day-ahead price, that is why following Inoue and Kilian (2008) and Rapach and Strauss (2010) we set $m = 1$ implying we are effectively considering regular bootstrap. For a textbook treatment of time series applications of bootstrap we refer Davison and Hinkley (1997) and, Chapter 8 particularly for dependent data applications.

where $\hat{y}_{d+1,h,b|d}^{*PT}$ is the pretested forecast computed using the b -th bootstrap sample and B indicates the number of replications. In theory, $B \rightarrow \infty$ would produce a perfect approximation, however, in practice, we set $B = 100$ since it tends to give a reasonable approximation, as underlined by [Inoue and Kilian \(2008\)](#) and [Rapach and Strauss \(2010\)](#). Once $\hat{y}_{d+1,h|d}^{ba}$ is derived one can easily substitute into Eq. (3), and obtain the day-ahead electricity price forecast on day $d + 1$ and hour h conditional on the most recent information available on day d , $\hat{Y}_{d+1,h|d}^{ba}$. Throughout our discussion, we abbreviate the bagging approach with the pre-testing strategy performed at 1, 5 and 10 percent significance levels as **BA_LS1**, **BA_LS5**, and **BA_LS10**, respectively.

3.4 Factor-Augmented models

As mentioned before, the major drawback of multivariate modeling is that it lacks intraday hourly dependencies. To circumvent that issue, one might think of using the univariate approach. However, as discussed before, this procedure might suffer from serious inaccuracy in terms of forecasting due to accumulated errors. Another alternative could be constructing a VAR model. The use of a VAR model, however, will lead to a dramatic increase in the total number of parameters needed to be estimated. That increase, which is called as the “curse-of-dimensionality” by [Raviv et al. \(2015\)](#), dissipates the degrees of freedom resulting in large estimation uncertainty and it might render the estimation procedure infeasible especially for small samples. One way to overcome this “curse-of-dimensionality” problem is employing factor models, which aim to reduce the dimension by using a limited number of common factors to reflect the variability in the data. In our study, to account for the intra-day dependencies but without being exposed to that curse, we employ factor-augmented models. In essence, given that estimated factors intuitively represent the co-movements of hourly prices in different degrees, we augment the **AR_{hour}**, the **EXPERT**, the **LASSO_LS^{HQC}**, and the **BA_LS1** with the factors computed from three panel of intra-day prices selected via in-sample Bayesian Information Criterion through the PCA approach.

Factor models were previously employed in the EPF literature by [Vehviläinen and Pyykkönen \(2005\)](#), [Härdle and Trück \(2010\)](#), [Nowotarski et al. \(2014\)](#), [Raviv et al. \(2015\)](#),

Maciejowska and Weron (2015, 2016) and Ziel (2016). While these studies aim to forecast electricity prices by using direct forecasts of factors, which are computed mainly from AR and ARX type models, we prefer to augment our existing models with factors so that we can continue to harvest as much information as possible from available predictors but at the same time account for intra-day hourly dependencies of day-ahead electricity prices. Moreover, in our analysis, we utilize an in-sample information criterion, namely BIC, to specify the number of factors out of 24, whereas the current literature constructs models with fixed, pre-determined number of factors.

Methodologically, we estimate the the latent factors for each estimation window using the panel of raw (not transformed) price series. That is, for an estimation window of length R , we first define the \mathbf{Y} matrix as $\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 & \dots & \mathbf{Y}_{24} \end{bmatrix}_{R \times 24}$ where a typical column $\mathbf{Y}_h = \begin{bmatrix} Y_{d,h} & \dots & Y_{d+R-1,h} \end{bmatrix}'$ represents electricity prices of all days in the estimation window at hour h . Then, the principal component decomposition of \mathbf{Y} is given as:

$$\mathbf{Y} = \mathbf{F}\mathbf{\Lambda} + \boldsymbol{\epsilon} \quad (18)$$

where, with its n -th column being $\mathbf{F}_n = \begin{bmatrix} F_{d,n} & \dots & F_{d+R-1,n} \end{bmatrix}'$, \mathbf{F} covers the common latent factors, $\mathbf{\Lambda}$ is the 24×24 factor loadings matrix, and $\boldsymbol{\epsilon}$ is the idiosyncratic variation of prices at different hours, $n = 1, \dots, 24$. By arranging the factors according to the descending order of their corresponding eigenvalues, we estimate the factors - common drivers of price changes - through the standard PCA approach. Let $\mathbf{f}_n = \begin{bmatrix} f_{d,n} & \dots & f_{d+R-1,n} \end{bmatrix}'$, be the *asinh* transformed version of the \mathbf{F}_n , we respectively form the factor-augmented versions of the **AR_{how}**, the **EXPERT**, and the large scale pure-price models as

$$y_{d,h} = \bar{y}_{how,d,h} + \psi_{h,0} + \sum_{i=1}^{p_h} \psi_{h,i}(y_{d-i,h} - \bar{y}_{how,d-i,h}) + \sum_{n=1}^N \alpha_n f_{d-1,n} + \epsilon_{d,h} \quad (19)$$

$$\begin{aligned}
y_{d,h} = & \beta_{h,1} + \underbrace{\beta_{h,2}y_{d-1,h} + \beta_{h,3}y_{d-2,h} + \beta_{h,4}y_{d-7,h}}_{\text{autoregressive effects}} + \underbrace{\beta_{h,5}y_{d-1,min} + \beta_{h,6}y_{d-1,max}}_{\text{non-linear effects}} \\
& + \underbrace{\beta_{h,7}y_{d-1,24}}_{\text{last-hour effect}} + \underbrace{\sum_{i=1}^6 \beta_{h,7+i}dow_{d,h}^i}_{\text{weekday dummies}} \\
& + \underbrace{\sum_{i=1}^6 \beta_{h,13+i}dow_{d,h}^i y_{d-1,h} + \sum_{i=1}^6 \beta_{h,19+i}dow_{d,h}^i y_{d-1,24}}_{\text{periodic effects}} + \underbrace{\sum_{n=1}^N \alpha_n f_{d-1,n}}_{\text{intra-day effects}} + \epsilon_{d,h}
\end{aligned} \tag{20}$$

$$\begin{aligned}
y_{d,h} = & \underbrace{\sum_{n=1}^{24} \alpha_n f_{d-1,n}}_{\text{intra-day effects}} + \underbrace{\sum_{i=1}^8 \sum_{j=1}^{24} \mu_{h,i,j,1} y_{d-i,j}}_{\text{price autoregressive effects}} + \underbrace{\sum_{i=1}^8 \mu_{h,i,1,2} y_{d-i,min} + \sum_{i=1}^8 \mu_{h,i,1,3} y_{d-i,max}}_{\text{price non-linear effects}} \\
& + \underbrace{\sum_{i=1}^7 \mu_{h,i,1,4} dow_{d,h}^i}_{\text{weekday dummies}} + \underbrace{\sum_{i=1}^7 \mu_{h,i,1,5} dow_{d,h}^i y_{d-1,avg}}_{\text{price average effects}} \\
& + \underbrace{\sum_{i=1}^6 \mu_{h,i,1,6} dow_{d,h}^i y_{d-1,h} + \sum_{i=1}^6 \mu_{h,i,1,7} dow_{d,h}^i y_{d-1,24}}_{\text{price periodic effects}} + \epsilon_{d,h}.
\end{aligned} \tag{21}$$

where N represents the total number of factors included in the models and α_n is the marginal effect of the n -th factor on the price. For the augmented **AR_{how}** (**fAR_{how}**) and the augmented expert model, **EXPERT** (**fEXPERT**), the number of factors is chosen according to the in-sample BIC, while all twenty-four factors are included in the augmented large scale pure-price model. In our analysis, if the estimation of the augmented large scale pure-price model is carried out by the LASSO approach, we abbreviate it as **fLASSO_LS^{HQC}**. The estimation results obtained from the bagging approach with the pre-testing strategy performed at 1, 5 and 10 percent significance levels are denoted as **fBA_LS1**, **fBA_LS5**, and **fBA_LS10**, respectively.

3.5 Forecast evaluation metrics

Forecast performance evaluation techniques are vast in econometric forecasting literature. Particularly for the EPF literature, we refer reader to [Hyndman and Koehler \(2006\)](#); [Nowotarski et al. \(2014\)](#); [Nowotarski and Weron \(2016\)](#). In our analyses, we compare forecast performances of the models discussed in the previous sections in terms of three widely used statistics in the EPF literature, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean of Weekly-weighted Mean Absolute Error ($\overline{\text{WMAE}}$). We checked for the sensitivity of the forecast performances of the models to alternative evaluation metrics and observed that the results remained almost unchanged. Evaluation techniques employed in this study are formulated as follows:

$$\text{MAE} = \frac{1}{24P} \sum_{i=1}^P \sum_{j=1}^{24} |\hat{\epsilon}_{i,j}| \quad (22)$$

$$\text{RMSE} = \sqrt{\frac{1}{24P} \sum_{i=1}^P \sum_{j=1}^{24} (\hat{\epsilon}_{i,j})^2} \quad (23)$$

$$\overline{\text{WMAE}} = \frac{1}{P/7} \sum_{k=1}^{P/7} \left(\frac{\sum_{i=7k-6}^{7k} \sum_{j=1}^{24} |\hat{\epsilon}_{i,j}|}{\sum_{i=7k-6}^{7k} \sum_{j=1}^{24} Y_{i,j}} \right) \quad (24)$$

where, P is the number of days in the out-of-sample period, and $\hat{\epsilon}_{i,j} = \hat{Y}_{i,j|i-1} - Y_{i,j}$ with $\hat{Y}_{i,j|i-1}$ and $Y_{i,j}$ is the predicted price and realized price for day i and hour j , respectively.

Note that the calculation of the $\overline{\text{WMAE}}$ requires the out-of-sample period to cover a multiple of a week and therefore while computing the $\overline{\text{WMAE}}$ only full weeks are taken into account. In the analyses, our out-of-sample period covers 1470 days for five major markets (350 days for GEFCom2014) and that corresponds to 210 weeks (50 weeks) in total. Regarding the performances of the forecast evaluation metrics we employ, the MAE and the $\overline{\text{WMAE}}$ are more robust to outliers compared to the RMSE, while the $\overline{\text{WMAE}}$ has the further advantage to eliminate the adverse effects of close to zero and negative electricity prices due to the normalization ([Uniejewski et al. \(2019\)](#); [Marcjasz et al. \(2020\)](#)). We also evaluate hourly forecast performances of all considered models in terms of the hourly MAE (MAE_h) and the hourly RMSE (RMSE_h).

While the RMSE, MAE, and $\overline{\text{WMAE}}$ provide informative signals on the comparative performance of the forecasts of the models, they fail to suggest statistically significant conclusions. Therefore, we further evaluate the forecasts by using the forecast encompassing test of Diebold and Mariano (1995) (abbreviated DM). If, say, there are two competing models, Model A and Model B, the DM test simply tests whether or not forecasts of one model statistically outperform the other one. To do that, given the covariance-stationary loss differential series for the hypothetical models A and B:

$$\Delta_{A,B,d,h} = \left| \hat{\epsilon}_{A,d,h} \right| - \left| \hat{\epsilon}_{B,d,h} \right| \quad (25)$$

where, $\hat{\epsilon}_{X,d,h}$ is the prediction error of model X , $X = A, B$ at time d for a given hour h . one needs to compute the p -values of two one-sided tests for each dataset: having the nulls of $H_0 : \mathbb{E}(\Delta_{A,B,d,h}) \leq 0$ implies that Model- A encompasses B , whereas null of $H_0 : \mathbb{E}(\Delta_{A,B,d,h}) \geq 0$ implies Model- B encompasses A for hour h . Rejection of both nulls at the same time means failure of models to outperform each other for hour h .

We also consider the “*multivariate*” DM test which allows us to derive a single statistic for each forecasting model instead of 24. The multivariate loss differential series has the form:

$$\Delta_{A,B,d} = \sum_{i=1}^{24} \left| \hat{\epsilon}_{A,d,i} \right| - \sum_{i=1}^{24} \left| \hat{\epsilon}_{B,d,i} \right|. \quad (26)$$

Like its standard version, we test two null hypotheses $H_0 : \mathbb{E}(\Delta_{A,B,d}) \leq 0$ and $H_0 : \mathbb{E}(\Delta_{A,B,d}) \geq 0$ to decide on whether or not one model outperforms the other one considering all 24 h of the day.

4 Empirical results

In this section, we present the results of our empirical forecasting exercise for day-ahead electricity prices of five major power markets and GEFCom2014. In the first subsection, we first compare forecast performances of all models according to MAE, RMSE, and $\overline{\text{WMAE}}$ and then continue with the DM tests to provide statistical inferences for the assessment of the performances. In the final subsection, we provide the occurrence tables that present the selected variables by the models providing promising forecast

performances.

4.1 Forecast results

Conducting our out-of-sample study over the out-of-sample period of 1470-day for five major markets and 350-day for GEFCom2014, we present the forecast performances of the **mean_{how}**, **naive**, **AR_{how}**, **EXPERT**, **LASSO_LS^{HQC}**, **BA_LS1**, **BA_LS5**, and **BA_LS10** models in the first eight columns of Table 2. While Panel A of Table 2 gives the values of MAE, Panel B and C provide the values of RMSE and $\overline{\text{WMAE}}$, respectively. Table 2 clearly indicates that shrinkage estimation methods (i.e. bagging and LASSO) outperform the benchmark models (i.e. **mean_{how}**, **naive**, **AR_{how}**) and the **EXPERT** model, irrespective of whether MAE, RMSE, or $\overline{\text{WMAE}}$ is used. This finding is not surprising and it is in line with the existing literature (Nowotarski and Weron (2016); Ziel and Weron (2018); Uniejewski et al. (2018, 2019)). Though it provides lower forecast accuracy compared to the large scale shrinkage estimated models, it is noteworthy to mention that, in sharp contrast to other benchmark models, the **EXPERT** model shows a remarkable forecast performance and it appears to be a promising model in terms of all evaluation metrics.

Regarding the comparison between the approaches of bagging and LASSO, we observe that, in terms of the MAE, the bagging approach outperforms the LASSO for five datasets: NP.SYS, PJM.COMED, OTE.CZ, OMIE.SP and GEFCom2014, while the LASSO appears to be the best model for only one data set: NP.N2EX. As discussed in the above sections, we apply the bagging approach by performing the pre-testing strategy at different significance levels. From Panel A of Table 2, we get the result that setting a low significance level in the pre-testing strategy (i.e. keeping the predictor in the model only if it is highly significant) leads to an increase in the performance of the bagging approach. That is, according to the MAE results, while **BA_LS1** provides better performance than the **LASSO_LS^{HQC}** for five datasets (NP.SYS, PJM.COMED, OTE.CZ, OMIE.SP, and GEFCom2014), **BA_LS5** and **BA_LS10** outperform the **LASSO_LS^{HQC}** for four (PJM.COMED, OTE.CZ, OMIE.SP, and GEFCom2014) and three datasets (PJM.COMED, OMIE.SP, and GEFCom2014), respectively. Concerning the RMSE and

Table 2: Forecasting results for bagging and other conventional day-ahead price forecasting models

Panel (A): Mean Absolute Errors (MAE)

Market	Original Versions						Factor-augmented Versions					
	mean _{how}	naive	AR _{how}	EXPERT	LASSO_LS ^{HQC}	BA_LSI	BA_LS5	BA_LS10	fAR _{how}	fEXPERT	fLASSO_LS ^{HQC}	fBA_LSI
NP.SYS	8.306	2.991	2.625	2.122	2.104	2.042	2.130	2.228	2.375	2.156	2.166	2.054
NP.N2EX	10.154	5.951	5.055	4.768	4.618	4.620	4.668	4.779	4.917	4.745	4.634	4.653
PJM.COMED	7.811	5.113	4.077	3.594	3.461	3.325	3.313	3.349	4.183	3.462	3.367	3.388
OTE.CZ	9.918	7.466	7.329	5.366	5.306	5.246	5.262	5.413	6.671	5.304	5.354	5.344
OMIE.SP	10.136	5.615	5.103	3.860	3.974	3.798	3.799	3.879	4.381	3.829	3.990	3.826
GEFCom2014	15.065	9.463	8.194	7.348	7.257	7.047	7.068	7.197	8.153	7.357	7.333	7.091

Panel (B): Root Mean Squared Error (RMSE)

NP.SYS	11.210	5.667	4.713	4.187	4.215	4.232	5.014	5.807	4.358	4.521	4.808	4.255
NP.N2EX	18.780	18.560	13.902	13.809	13.573	13.576	13.595	13.682	13.819	14.454	13.778	13.663
PJM.COMED	14.105	10.743	8.432	8.192	8.029	7.351	7.074	7.017	25.425	8.475	7.219	7.451
OTE.CZ	13.785	11.287	10.447	8.070	8.004	7.941	7.937	8.144	9.635	8.005	8.077	8.026
OMIE.SP	12.901	8.305	7.141	5.399	5.447	5.333	5.307	5.398	6.162	5.362	5.467	5.358
GEFCom2014	31.220	18.082	16.117	15.387	15.336	13.813	13.493	13.511	17.385	15.597	15.082	14.076

Panel (C): Mean of Weekly-weighted Mean Absolute Error ($\overline{\text{WMAE}}$)

NP.SYS	40.495	15.924	13.834	11.032	10.847	10.664	11.106	11.612	12.319	11.058	11.085	10.742
NP.N2EX	36.188	21.570	18.295	17.208	16.699	16.723	16.899	17.308	17.789	17.102	16.753	16.847
PJM.COMED	42.348	26.860	21.230	18.502	17.630	17.221	17.210	17.472	19.824	17.596	17.468	17.446
OTE.CZ	43.707	35.611	34.410	25.146	24.798	24.696	24.784	25.461	31.236	24.813	24.968	25.143
OMIE.SP	38.675	21.231	19.439	14.669	14.966	14.386	14.361	14.622	16.595	14.533	15.002	14.427
GEFCom2014	37.329	26.026	22.275	19.473	19.302	19.341	19.598	20.094	21.471	19.531	19.616	19.125

Note: The table reports the model forecasting results calculated for full out-of-sample period as defined by Eq. (22), (23), and (24), respectively. Best performing model result is indicated with boldface in each dataset.

$\overline{\text{WMAE}}$ values, given in Panel B and Panel C of Table 2, we observe similar pictures. More specifically, while the LASSO approach outperforms the bagging technique only for NP.SYS and NP.N2EX (NP.N2EX and GEFCom2014) in terms of the RMSE ($\overline{\text{WMAE}}$), the bagging (**BA_LS1** and **BA_LS5**) provides higher forecast accuracy for the majority of the markets.

To provide further insight into the comparison of forecast accuracies of the LASSO and bagging approaches, we plot the MAE values from **BA_LS1** and **LASSO_LS^{HQC}** for each 24 h of the day in Figure 1. Overall, we see that forecast errors from the bagging and LASSO approaches follow a similar pattern over the day, though bagging appears to provide higher forecast accuracy during the working hours (05:00-18:00) in almost all markets but especially in NP.SYS, PJM.COMED, and OMIE.SP. For these markets, later in the text, using hourly DM test results we statistically prove that **BA_LS1** clearly encompasses the **LASSO_LS^{HQC}** for 10 to 16 hours of the day. The main reason behind the observed superior forecast performance of the bagging approach during working hours could be the structural differences in the selected explanatory variables by the bagging and LASSO approaches, which will be discussed in detail in the next subsection. To assess whether or not factor-augmentation, which aims to account for intra-day hourly dependencies of day-ahead electricity prices, provides any further improvement in the forecast accuracy we report the forecast results of the factor-augmented models, **fAR_{how}**, **fEXPERT**, **fLASSO_LS^{HQC}**, and **fBA_LS1** in the last four columns of Table 2⁸. The results indicate that, overall, the factor augmentation provides an improvement in the forecast performances of the **AR_{how}** and **EXPERT** models, though that improvement is slightly less substantial when the RMSE is used as the evaluation metric. While factor augmentation seems to increase the forecast accuracy of the **AR_{how}** and **EXPERT** models, the same cannot be deduced for shrinkage methods. Regarding the LASSO approach, we observe that factor-augmentation leads to an improvement in the forecast performance for only one or two markets (PJM.COMED, GEFCom2014), depending on the loss metric used. In the case of the bagging approach, on the other

⁸In terms of the bagging approach, we report only the forecast results from the augmented-bagging technique with the pre-testing strategy performed at 1 percent significance level due to space limitations and also the outperformance of the **BA_LS1** over **BA_LS5** and **BA_LS10**. However, the forecast results of **fBA_LS5** and **fBA_LS10** are available from the authors upon request.

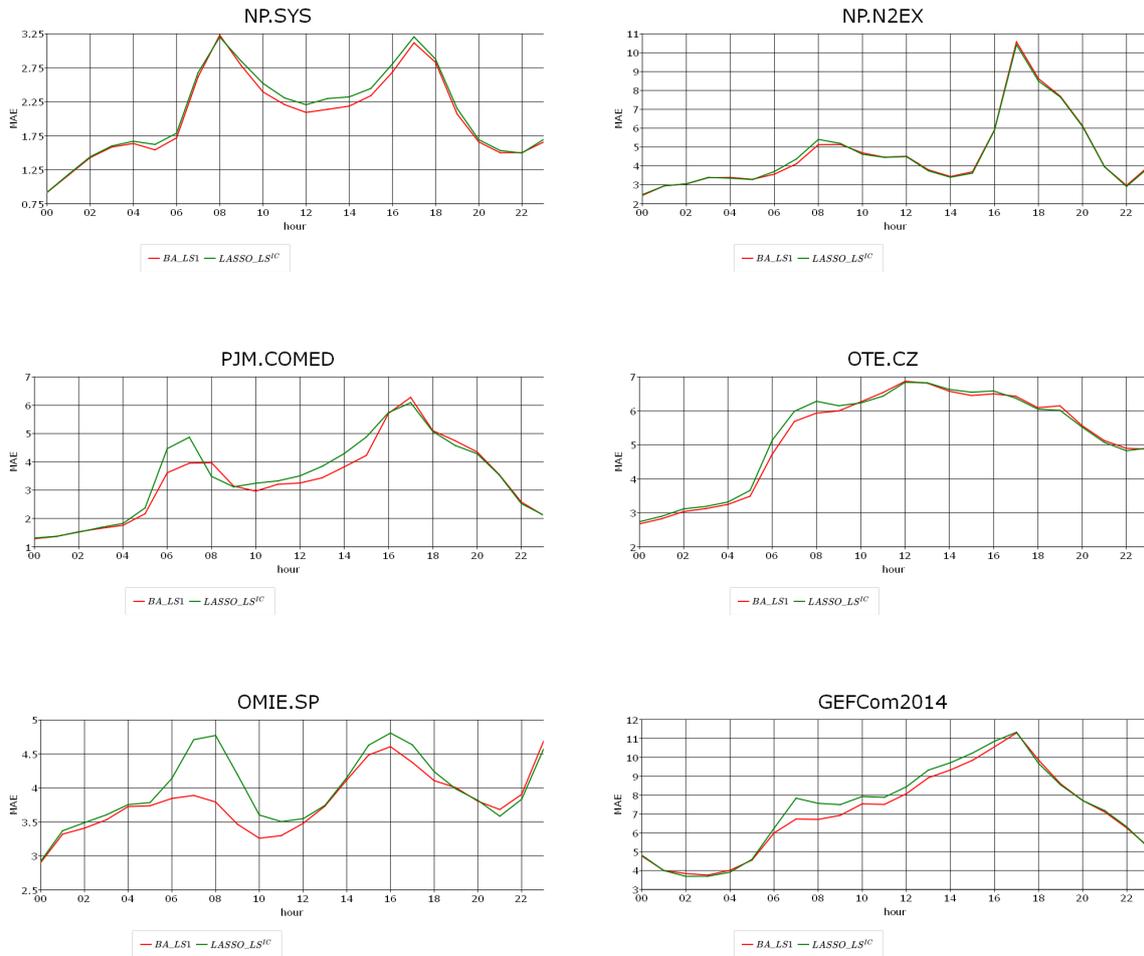


Figure 1: Hourly Mean Absolute Error for $LASSO_LS^{IC}$ and BA_LS1 models as defined in Section 3.5 for full out-of-sample period. Results are given for each dataset in separate panels.

hand, **fBA_LS1** appears to be inferior to its original version for all markets in terms of all evaluation metrics. Given that the motivation behind the inclusion of the factors in the models is to capture intra-day dependencies of electricity prices, our finding that factor augmentation results in inferior forecasts for shrinkage methods might imply that large scale shrinkage methods already include intra-day dependencies in their original versions even if they are formed in a multivariate context. Therefore, one might not need to deal with the complexity of the implementation of a univariate framework to derive forecasts from a large scale model. However, for the multivariate **AR_{how}** and **EXPERT** models, the factor-augmentation approach could be a computationally simple alternative to the univariate modeling framework.

So far, our out-of-sample forecasting exercise has revealed the superior forecasting performance of the shrinkage estimation procedures in comparison to other constructed models. Moreover, comparing two shrinkage estimation approaches, bagging and LASSO, the bagging technique seems to outperform the LASSO for the majority of the markets, in both daily and hourly time scales. All these conclusions, however, are deduced from evaluation metrics, which can provide signals on the comparative performance of the forecasts of the models but cannot suggest a statistical inference. To be able to provide statistical inferences regarding comparative forecast performances we continue with the DM tests, as explained in Section 3.5. In that sense, we first apply the multivariate DM test, which compares forecasts on a daily basis, and then we perform the standard DM test to derive a better understanding of how the forecast performances change across the hours⁹.

Figure 2 shows the multivariate DM test results for each dataset. We utilize a heat map to represent the range of the p -values for the null hypothesis that the model on the X-axis encompasses the model on the Y-axis. In the map, as a p -value gets closer to zero (0.10), the difference between the forecasts of the model on the X-axis and the forecasts of the model on the Y-axis becomes more (less) substantial and the corresponding cell turns red (green). That is, for example, for a specific market if the first row of the map is completely green, while the first column is completely red (excluding the diagonal

⁹Since we observe similar patterns for different loss metrics all DM test results (both multivariate and standard versions) are based on Mean Absolute Error.

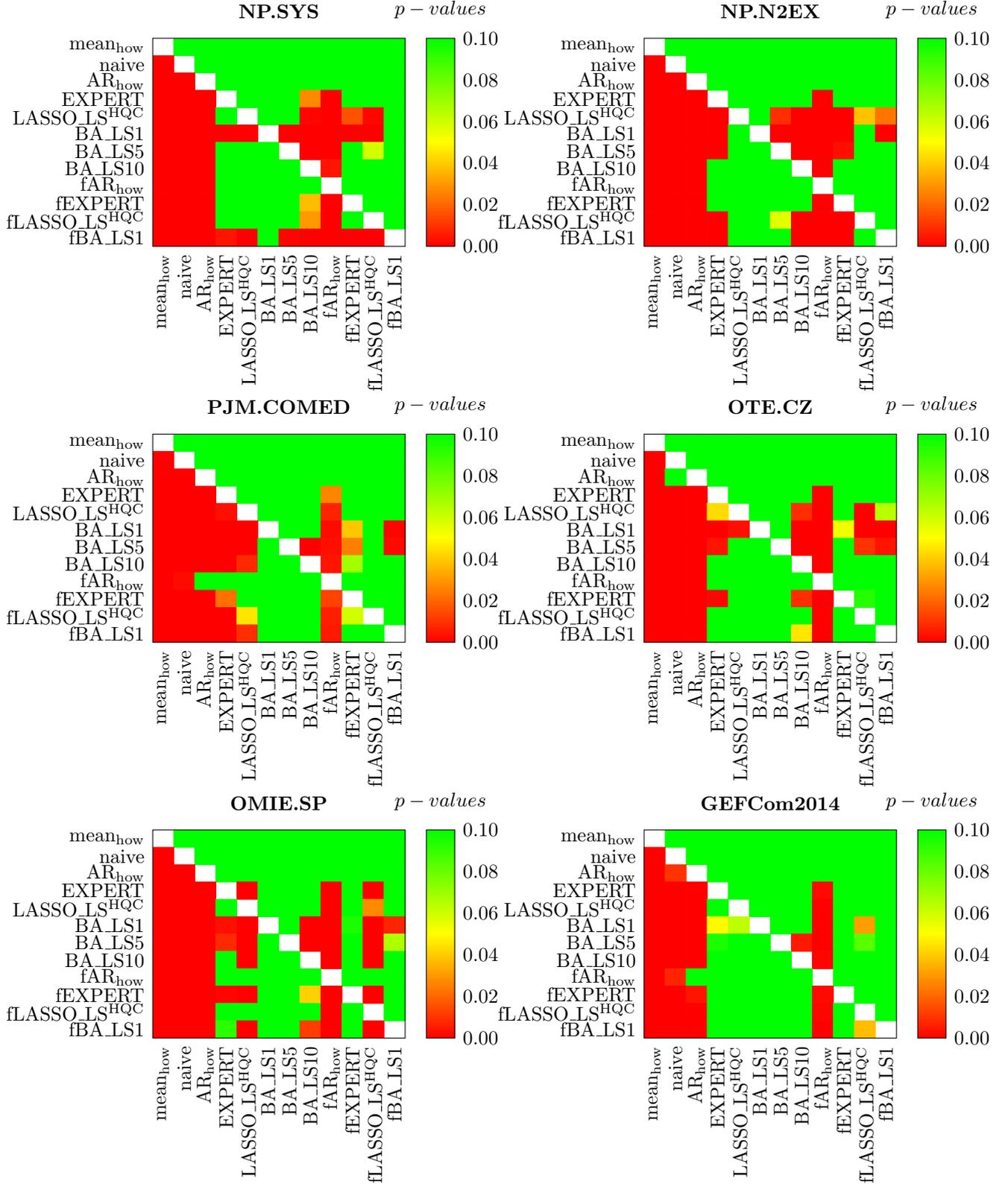


Figure 2: Multivariate DM test results for each dataset as defined by the loss differential series given in Eq. (26). p -values are given for the null hypothesis of $H_0 : \mathbb{E}(\Delta_{X,Y,d}) \leq 0$ which implies that the model on the X -axis encompasses the model on the Y -axis. A heat-map is used to indicate the range of p -values with the corresponding colorbars.

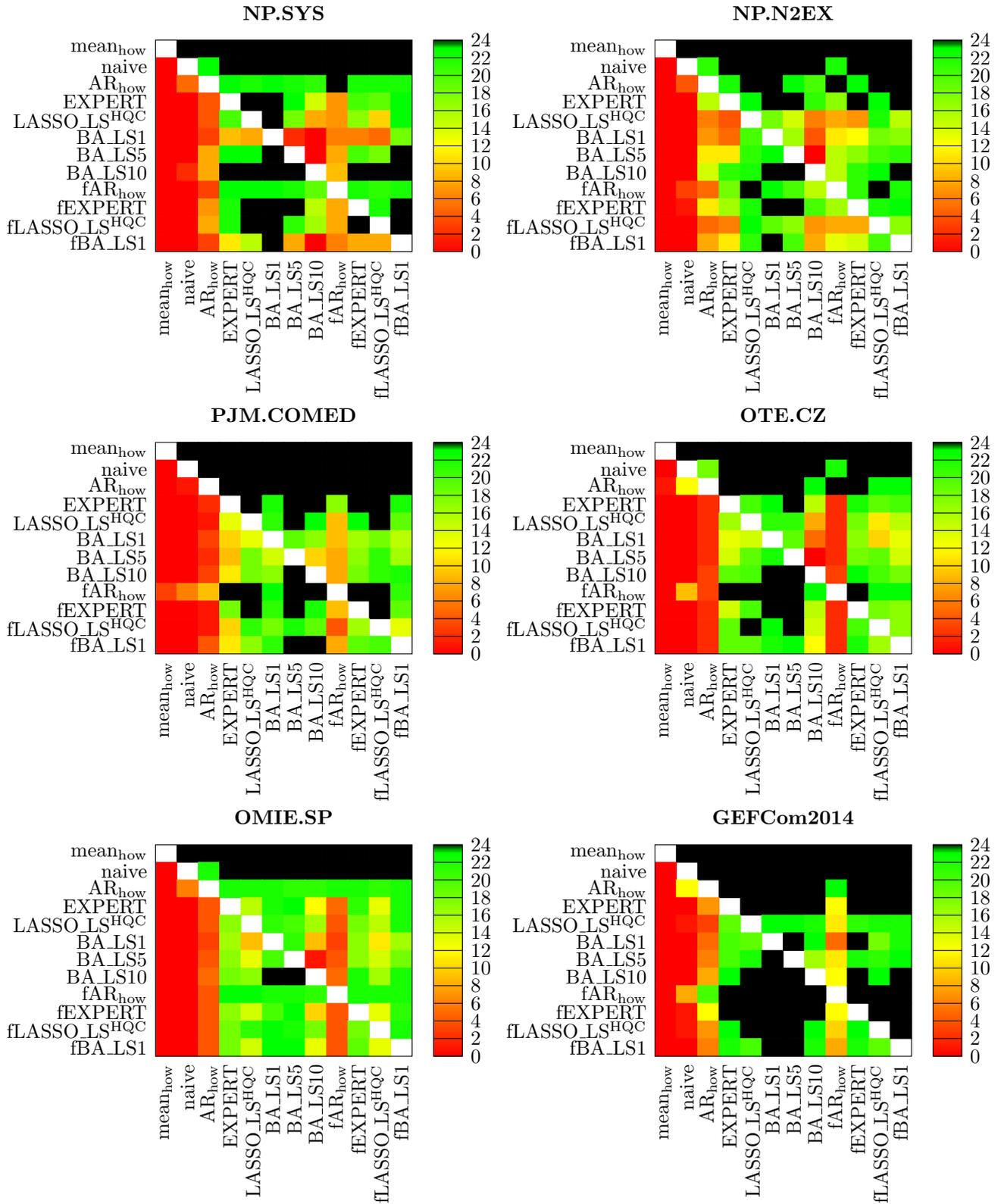


Figure 3: Hourly DM test results for each dataset as defined by the loss differential series given in Eq. (25). Each cell indicates total number of hours out of 24 for which model on the X-axis encompasses model on the Y-axis at 5% significance level. A heat-map is used to indicate the range of hours with the corresponding colorbars.

part), then we observe the inference that the forecasts of the **mean_{how}** model are significantly outperformed by those of all other models for that specific market. A careful inspection of Figure 2 reveals important inferences, almost all of which are consistent with the conclusions drawn from Table 2. Firstly, it appears that the multivariate DM tests statistically confirm that shrinkage estimation methods outperform the benchmark models (i.e. **mean_{how}**, **naive**, **AR_{how}**) for all markets. Moreover, our finding that factor-augmentation improves the forecast accuracy of the **AR_{how}** and **EXPERT** models for all markets is supported by the DM tests as well. Regarding the forecasting performances of the bagging and LASSO approaches, it appears that the **BA_LS1**, seems to beat the **LASSO_LS^{HQC}** for the markets NP.SYS, OTE.CZ, OMIE.SP and GEFCom 2014. For the NP.N2EX market, forecasts from the original LASSO, i.e. **LASSO_LS^{HQC}**, and from its factor-augmented version, i.e. **fLASSO_LS^{HQC}**, are not significantly worse than those of **BA_LS1**. This finding reveals the statistical insignificance of our previously observed inference that the LASSO provides lower forecast errors than the bagging technique for the market NP.N2EX in terms of all evaluation metrics. For PJM.COMED, on the other hand, while **LASSO_LS^{HQC}** fails to beat the **BA_LS1**, the forecasts from the **fLASSO_LS^{HQC}** are not significantly worse than those of **BA_LS1**. Concerning the effect of factor-augmentation on the bagging and LASSO, we observe that augmentation does not provide a significant forecasting outperformance for all markets except PJM.COMED. Overall, being in line with the results of Table 2, the bagging approach appears to provide significantly higher forecast accuracy for the majority of the markets according to the multivariate DM tests.

To delve deeper into the statistical forecast comparison between bagging and LASSO approaches and to evaluate the forecast performances of the models across hours, we present the standard DM test results conducted for each hour of the day in Figure 3. Each cell in the heat map indicates the total number of hours for which the model on the X-axis encompasses the model on the Y-axis at the 5 percent significance level. If the forecasts of a model on the X-axis outperforms the forecasts of a model on the Y-axis for all 24h, the corresponding cell turns black and if the forecasts of a model on the X-axis are not significantly more accurate than those of a model on the Y-axis for any hour, the corresponding cell turns red. For NP.SYS, PJM.COMED, and OMIE.SP our results

show that **BA_LS1** clearly encompasses the **LASSO_LS^{HQC}** for 10 to 16 hours of the day (while the reverse is not true). The hours of the day for which **BA_LS1** beats the **LASSO_LS^{HQC}** approximately correspond to working hours where total energy demand is high and this finding is consistent with the inference drawn previously from Figure 1. For OTE.CZ, it seems from Figure 3 that for the hours approximately from 03:00 to 09:00, **BA_LS1** beats the **LASSO_LS^{HQC}**, while the two approaches show statistically similar forecast performances for the markets NP.N2EX and GEFCom2014 for almost all hours of the day.

4.2 Occurrence of variables

In this section, we aim to analyze the structures of the models that provide the most promising forecasts, i.e. **BA_LS1**, **LASSO_LS^{HQC}**, and **fEXPERT**. In theory, bagging and LASSO are both shrinkage estimation methods (Stock and Watson (2012)) and therefore they are expected to produce asymptotically similar forecasting performances. In finite samples, however, they can produce substantially different forecast performances, as illustrated in our forecasting exercise where the bagging approach attains superior forecasting performance compared to the LASSO in most of the markets, while their performances are not statistically different for the rest few markets. The underlying force for the differentiation in the forecasting performances of the bagging and LASSO approaches can be attributed to the selected parameters during the estimation procedure. In an attempt to better characterize the structural differences between **BA_LS1** and **LASSO_LS^{HQC}**, we construct Tables 3 to 7 to present the mean occurrences of the explanatory variables in both models. In each table, while the left panel gives the results for the LASSO, the right panel presents the results for the bagging approach. To generate occurrence tables we first count the number of times a specific explanatory variable is selected by the LASSO approach in each out-of-sample point, convert them in percentage points, and then average over all 6 datasets. Similarly, for bagging results are formed by counting the total number of variables selected by the t -test at the 1 percent significance level in each out-of-sample point and averaging those numbers over all datasets. We utilize heat maps to ease the interpretation of the tables so that more (less) commonly

selected explanatory variables are shown by greenish (reddish) cells.

While our results for the occurrence matrix of the LASSO are mostly in parallel with the findings of [Ziel and Weron \(2018\)](#) and other studies in the EPF literature, some substantial differences emerge between the variables selected by the bagging and LASSO approaches. The first striking difference comes from the widely known previous day's same hour effect that can be traced from the diagonal elements of the day: d-1 matrix in [Table 3](#). While the previous day's same hour price appears to be a quite relevant variable in the LASSO, it is less apparent in the bagging approach except for hours between 17:00 and 21:00. Another difference between LASSO and bagging approaches comes from the previous day's last hour effect, i.e. last row of the day: d-1 matrix in [Table 3](#). This effect shows high relevance in the LASSO for almost all hours even if its appearance follows a slightly decreasing pattern towards the evening hours. In the the bagging, on the other hand, it carries relevant information mainly for the night and early morning hours from 01:00 to 07:00, whereas it appears to be relatively less important during the working hours. The form of the appearance of the previous day's last hour price in the bagging approach seems to be consistent with the expectation that the high relevance of the previous day's last hour price for the early hours of the day should follow a decreasing pattern as some other factors in the price determination process kicks in during the day. Moreover, recall that when we compared the hourly forecasting performances of the bagging and LASSO approaches in [Figure 1](#), we deduced that their forecasting performances mainly differ during the working hours with bagging providing higher forecast accuracy than the LASSO. In this regard, one reason for the superior forecasting performance of the bagging might be the observed way of the integration of the previous day's last hour effect in the bagging approach. Regarding the other previous days' same hour effects, we see from [Tables 4, 5, and 6](#) that as the time difference increases the autoregressive effect becomes less relevant both for bagging and LASSO, though it is still more important for the LASSO.

Another remarkable difference between the bagging and LASSO is the appearances of the day-of-the-week dummies, the non-linear minimum and maximum effects, and the periodic effects, presented in [Table 7](#). As seen in the third row of the Table, in the LASSO approach, Monday, Sunday, and Saturday dummies pop up for almost all hours, whereas

in the bagging approach these dummies commonly appear only during the hours from 06:00 to 20:00, and Monday dummy is less apparent. Regarding the non-linear minimum and maximum effects, presented in the first two rows of Table 7, it seems that while the previous two days' minimum and maximum prices are more important for the LASSO approach, especially in the first six hours of the day, only the previous day's minimum price appears to be relevant for the bagging approach during the night and morning hours from 01:00 to 08:00. Similarly, the relevance of the periodic effects in the bagging approach is less apparent compared to the case observed in the LASSO approach, as seen in the last two rows of Table 7.

The observed differences in the mean occurrence tables give us grounds to believe that the differentiation in the forecast performances of the LASSO and bagging approaches mainly arise from their structural differences in the explanatory variables selection process. Overall, the mean occurrence tables indicate that some specific explanatory variables show higher appearances in the LASSO approach. One reason behind this finding could be the fact that the LASSO technique deals with the large scale models by shrinking some coefficients and setting some others to exactly zero, which in turn might lead higher number of appearances of some specific variables. The bagging approach, however, follows a completely different estimation strategy by dropping the insignificant variables in one pseudo-sample and giving them another chance for being selected in some other artificially generated sample.

Finally, turning to the selected factors in the **fEXPERT** model, we form Table 8 by choosing the optimum number of factors for each out-of-sample point according to BIC, counting the total number of selected factors, and finally averaging them over all datasets. According to Table 8, the first three factors show the highest relevance for almost all hours of the day, while the fourth and the fifth factors appear to be effective especially for the night and early morning hours from 03:00 to 06:00. This result implies that the forecast improvement of the **fEXPERT** model over its original version comes mainly from the inclusion of the first three factors and some additional higher degree factors for the relevant hours. Furthermore, our finding, highlighting the importance of the first three factors, is in line with that of [Raviv et al. \(2015\)](#) who observe that intra-day hourly dependencies of day-ahead electricity prices can be represented well through

a small number of factors without being exposed to the high dimensionality issue¹⁰.

¹⁰We also analyzed the mean occurrences of factors for the **fAR_{how}** model and obtain an approximately similar pattern to that of the **fEXPERT** model. The corresponding table is not reported here due to space considerations but it is available upon request.

Table 3: Mean occurrences of the LASSO_LS^{HQC} and BA_LSI model parameters in the left and right panel, respectively. A heat-map is used to indicate the range of occurrence values between 0 (→ red) and 100 (→ green). Continued in Table 4.

h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
$\beta_{h_1,1,1}$	37	40	27	37	38	43	42	22	17	22	24	14	21	22	15	8	7	7	12	16	11	09	45	39
$\beta_{h_1,1,2}$	37	39	14	1	3	25	13	12	8	6	3	4	1	4	2	0	0	0	0	0	5	6	1	12
$\beta_{h_1,1,3}$	17	43	38	19	6	11	10	11	13	3	5	3	2	7	6	4	2	10	7	4	9	5	13	
$\beta_{h_1,1,4}$	25	42	42	22	8	2	4	9	7	7	5	3	2	3	2	3	1	3	2	1	1	1	1	1
$\beta_{h_1,1,5}$	14	23	32	46	68	31	8	23	18	17	7	2	2	3	0	2	6	12	3	1	1	1	1	1
$\beta_{h_1,1,6}$	13	24	35	31	51	70	52	31	16	10	22	31	31	35	25	21	0	15	17	18	20	22	17	
$\beta_{h_1,1,7}$	27	35	27	22	24	18	44	49	30	27	11	13	16	27	35	40	39	30	4	4	3	6	15	14
$\beta_{h_1,1,8}$	18	9	5	5	13	14	70	59	32	7	1	17	19	22	20	23	10	23	14	7	15	9	13	
$\beta_{h_1,1,9}$	20	16	11	8	13	20	16	13	47	32	7	1	17	19	22	20	23	10	23	14	7	15	9	13
$\beta_{h_1,1,10}$	5	11	12	11	10	10	11	6	28	6	0	0	1	15	16	25	35	10	3	14	14	13	15	
$\beta_{h_1,1,11}$	3	2	8	8	12	12	14	10	9	26	35	9	7	4	8	14	15	15	9	3	5	9	14	8
$\beta_{h_1,1,12}$	2	11	11	17	20	12	15	21	25	24	39	38	27	12	9	10	19	26	22	7	2	3	10	6
$\beta_{h_1,1,13}$	0	2	9	12	11	9	4	3	5	13	28	47	41	30	5	4	4	3	14	3	1	2	7	3
$\beta_{h_1,1,14}$	0	2	1	1	1	1	0	3	1	6	36	52	36	8	1	1	5	7	3	0	0	10	4	4
$\beta_{h_1,1,15}$	23	14	9	6	6	10	16	23	7	12	18	18	41	51	53	48	10	3	4	3	2	0	9	2
$\beta_{h_1,1,16}$	6	15	17	22	23	19	21	18	23	20	13	35	54	64	78	86	97	18	9	18	12	6	9	11
$\beta_{h_1,1,17}$	40	38	37	16	15	12	11	18	30	27	41	54	52	56	54	52	76	97	32	17	21	24	22	10
$\beta_{h_1,1,18}$	53	58	70	69	66	61	51	67	54	46	47	29	18	38	34	36	29	24	16	47	96	27	6	10
$\beta_{h_1,1,19}$	33	35	42	34	39	56	83	84	82	70	58	45	38	34	36	29	24	16	47	96	27	6	10	
$\beta_{h_1,1,20}$	39	52	64	67	77	91	87	90	87	89	77	64	62	61	53	48	30	35	37	57	100	65	23	12
$\beta_{h_1,1,21}$	18	24	30	43	49	42	53	67	79	82	71	73	67	57	52	50	58	39	35	37	57	100	65	23
$\beta_{h_1,1,22}$	48	41	48	45	46	56	57	50	38	54	80	85	83	79	67	62	61	62	68	67	53	54	88	100
$\beta_{h_1,1,23}$	100	100	100	100	98	87	80	74	82	96	95	81	77	75	69	68	71	62	68	67	53	54	88	100
$\beta_{h_2,1,1}$	22	10	6	12	20	15	17	13	7	7	12	22	27	31	20	11	17	17	21	18	24	29	43	44
$\beta_{h_2,1,2}$	2	23	6	1	3	2	14	1	3	6	4	3	3	5	13	6	8	7	6	8	5	4	7	4
$\beta_{h_2,1,3}$	1	7	34	19	9	11	10	12	7	1	3	7	6	3	2	0	1	1	0	2	2	4	1	1
$\beta_{h_2,1,4}$	9	10	27	37	29	19	7	5	6	1	2	1	3	3	1	0	1	3	6	5	3	3	3	6
$\beta_{h_2,1,5}$	4	14	12	19	33	28	15	19	10	9	4	0	1	1	7	1	0	1	2	4	5	4	6	1
$\beta_{h_2,1,6}$	8	6	4	4	3	19	17	11	11	1	3	6	17	17	14	15	1	1	4	6	1	1	7	3
$\beta_{h_2,1,7}$	17	4	4	9	6	11	4	10	10	3	5	6	12	11	15	15	13	14	22	10	5	10	3	3
$\beta_{h_2,1,8}$	5	8	5	8	5	12	16	16	11	13	21	17	22	25	24	5	14	11	6	9	9	9	9	3
$\beta_{h_2,1,9}$	2	4	6	7	5	9	5	8	13	9	5	2	7	8	10	13	3	2	3	4	1	0	7	5
$\beta_{h_2,1,10}$	6	4	5	6	4	3	4	11	8	11	8	0	1	0	2	3	7	2	1	0	3	0	5	5
$\beta_{h_2,1,11}$	4	8	6	9	6	4	8	3	11	18	21	10	8	4	6	7	5	1	3	5	3	12	9	10
$\beta_{h_2,1,12}$	7	8	4	2	6	5	12	13	9	16	19	24	13	6	6	5	2	1	1	1	3	6	4	7
$\beta_{h_2,1,13}$	11	8	6	10	16	17	11	9	6	4	4	2	0	3	2	1	1	3	6	4	1	0	3	12
$\beta_{h_2,1,14}$	2	4	6	13	13	9	17	29	17	2	4	3	1	2	6	11	5	18	16	7	1	6	16	18
$\beta_{h_2,1,15}$	10	4	1	6	6	18	11	15	21	25	31	8	11	10	16	24	34	14	8	9	10	9	6	17
$\beta_{h_2,1,16}$	8	9	9	10	18	18	13	12	11	11	12	19	27	31	38	38	44	5	17	5	4	3	4	4
$\beta_{h_2,1,17}$	29	29	27	20	22	26	20	24	13	10	15	15	19	18	20	31	6	47	11	33	29	25	25	
$\beta_{h_2,1,18}$	38	26	22	24	15	25	17	19	16	12	10	5	4	5	11	4	20	31	32	25	53	25	18	19
$\beta_{h_2,1,19}$	36	41	27	25	26	27	21	15	9	7	7	2	3	4	4	5	8	11	34	10	31	7	9	9
$\beta_{h_2,1,20}$	30	54	62	61	62	55	32	19	7	19	18	12	11	15	2	13	3	10	8	19	8	40	46	44
$\beta_{h_2,1,21}$	41	43	54	62	54	51	50	39	26	28	22	30	23	28	20	23	14	23	33	21	29	43	44	44

h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
$\beta_{h_1,1,1}$	40	33	27	27	23	21	19	22	22	24	27	22	24	27	28	29	28	24	20	17	22	25	23	24	20
$\beta_{h_1,1,2}$	10	12	10	12	11	11	12	11	10	12	13	16	18	19	21	18	14	12	11	10	10	13	15	11	11
$\beta_{h_1,1,3}$	10	13	14	15	16	16	17	18	17	15	15	16	15	16	17	15	13	12	11	12	12	13	12	14	13
$\beta_{h_1,1,4}$	13	15	17	16	15	16	15	13	14	17	15	13	12	11	12	12	15	13	12	11	11	13	14	11	11
$\beta_{h_1,1,5}$	16	14	15	17	26	15	14	10	10	12	9	8	7	8	10	11	11	10	11	11	10	8	9	9	9
$\beta_{h_1,1,6}$	14	15	13	20	36	21	22	16	13	11	12	14	15	16	14	15	16	14	13	13	14	19	16	15	
$\beta_{h_1,1,7}$	14	17	15	14	17	18	18	15	14	15	12	12	11	12	11	12	11	12	11	11	11	11	11	10	10
$\beta_{h_1,1,8}$	15	16	19	18	16	14	17	18	19	12	11	11	14	15	14	14	11	12	11	11	11	11	11	11	10
$\beta_{h_1,1,9}$	13	15	16	14	14	13	17	21	22	18	21	17	16	17	17	14	12	12	19	17	17	17	17	14	14
$\beta_{h_1,1,10}$	19	16	15	17	14	17	21	17	14	19	15	18	17	17	16	20	23	21	22	16	17	18	19	19	20
$\beta_{h_1,1,11}$	12	15	12	14	15	12	14	17	18	16	20	13	17	16	16	19	21	23	25	26	29	25	31	26	26
$\beta_{h_1,1,12}$	12	15	11	11	11	14	19	18	17	14	16	14	12	14	16	19	21	23	21	20	21	20	17	12	12
$\beta_{h_1,1,13}$	14	12	11	11	10	10	13	12	10	9	10	9	9	13	12	13	12	15	15	12	15	15	12	13	15
$\beta_{h_1,1,14}$	13	11	13	13	11	11	12	17	15	13	12	13	12	13	12	13	15	11	12	14	15	16	14	16	16
$\beta_{h_1,1,15}$	15	18	15	16	14	19	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
$\beta_{h_1,1,16}$	16	13	15	16	14	15	21	22	26	24	25	25	27	24	35	47	19	13	18	14	13	23	23	16	21
$\beta_{h_1,1,17}$	19	18	21	21	23	27	29	27	24	19	21	14	14	15	14	15	17	24	46	18	20	18	13	12	12
$\beta_{h_1,1,18}$	11	11	11	12	11	16	27	23	21	22	21	20	20	20	20	20	15	17	24	46	18	20	18	13	12
$\beta_{h_1,1,19}$	14	23	24	25	29	28	29	27	36	36	38	31	28	29	28	29	24	23	17	19	47	25	17	16	16
$\beta_{h_1,1,20}$	12	15	14	15	17	13	17	22	26																

Table 4: Mean occurrences of the LASSO_LS^{HQC} and BA_LS1 model parameters in the left and right panel, respectively. A heat-map is used to indicate the range of occurrence values between 0 (→ red) and 100 (→ green). Continued in Table 5.

	h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
day: d = 3	$\beta_{h,3,1,1}$	10	0	2	5	2	4	1	7	13	8	5	9	11	15	22	21	17	12	11	11	15	15	17	17	8
	$\beta_{h,3,2,1}$	31	32	13	2	3	11	4	1	2	16	15	12	10	14	12	11	7	6	16	4	3	4	3	4	3
	$\beta_{h,3,3,1}$	8	11	23	13	11	15	13	9	1	0	9	8	10	9	10	14	15	2	1	4	8	7	1	4	8
	$\beta_{h,3,4,1}$	6	16	31	30	33	27	9	7	3	6	6	4	9	11	7	4	4	0	3	0	5	2	1	3	0
	$\beta_{h,3,5,1}$	5	6	9	19	29	21	10	4	2	3	1	4	3	4	2	1	1	1	1	1	1	0	1	1	0
	$\beta_{h,3,6,1}$	4	6	5	9	32	24	13	8	3	1	2	2	3	12	10	9	10	9	14	6	8	10	7	3	10
	$\beta_{h,3,7,1}$	23	30	27	28	36	30	55	25	28	25	11	11	12	14	13	11	6	9	14	6	8	10	7	6	8
	$\beta_{h,3,8,1}$	3	6	8	4	8	9	16	29	21	12	8	1	0	0	3	7	7	2	9	7	6	3	2	3	3
	$\beta_{h,3,9,1}$	13	12	8	7	10	5	4	7	12	8	4	9	2	0	5	5	7	2	5	5	4	8	8	4	4
	$\beta_{h,3,10,1}$	5	9	8	5	3	3	7	8	3	11	27	12	12	9	9	5	3	1	5	4	9	12	3	3	3
	$\beta_{h,3,11,1}$	13	21	24	17	13	22	6	7	9	8	20	30	18	16	12	16	3	8	10	7	2	8	7	7	7
	$\beta_{h,3,12,1}$	6	5	2	2	3	3	5	3	1	7	8	14	4	2	3	2	2	3	2	3	2	2	19	7	7
	$\beta_{h,3,13,1}$	7	1	1	2	5	6	18	17	8	0	2	4	16	18	21	15	13	9	3	11	6	3	5	12	12
	$\beta_{h,3,14,1}$	3	10	9	10	5	2	11	11	7	5	4	5	11	18	2	6	7	6	2	1	5	0	2	2	2
	$\beta_{h,3,15,1}$	8	14	11	12	12	9	6	5	7	5	10	14	20	9	26	2	3	8	9	3	2	2	2	2	2
	$\beta_{h,3,16,1}$	5	4	9	12	6	18	25	10	21	29	23	6	7	11	10	9	18	0	7	4	1	3	5	5	5
	$\beta_{h,3,17,1}$	6	0	6	1	6	5	9	17	11	8	8	15	16	16	23	25	13	63	5	15	8	9	10	7	9
	$\beta_{h,3,18,1}$	4	6	12	7	10	20	24	17	10	5	12	7	2	3	4	6	2	2	2	65	8	4	2	6	2
	$\beta_{h,3,19,1}$	3	2	3	3	7	11	8	7	9	1	2	0	3	2	1	2	9	10	68	8	4	2	3	3	3
	$\beta_{h,3,20,1}$	14	20	25	26	23	15	9	11	21	14	12	6	8	8	7	4	0	0	8	41	16	11	11	11	11
	$\beta_{h,3,21,1}$	12	17	24	22	22	17	9	7	11	9	7	8	13	16	18	16	11	6	6	5	7	21	48	25	25
	$\beta_{h,3,22,1}$	17	18	23	19	26	25	25	7	14	11	2	9	11	5	8	14	8	7	1	2	8	12	27	27	27
	$\beta_{h,3,23,1}$	24	12	6	5	5	8	1	3	7	3	3	0	0	1	6	11	9	3	1	3	3	8	8	20	24
	$\beta_{h,3,24,1}$	13	14	12	6	6	6	3	6	6	5	6	9	2	2	4	8	6	7	17	6	13	12	23	22	32
$\beta_{h,3,25,1}$	11	16	18	15	16	17	16	14	6	1	0	1	2	1	5	6	6	4	2	1	8	5	3	7	7	
$\beta_{h,3,26,1}$	5	16	27	38	27	16	13	10	7	7	4	5	7	5	6	6	4	5	6	4	1	4	1	4	4	
$\beta_{h,3,27,1}$	9	17	20	18	30	19	18	16	10	7	7	4	7	8	6	7	7	3	3	6	4	7	9	4	4	
$\beta_{h,3,28,1}$	17	28	16	11	17	37	30	18	18	6	4	4	8	4	4	3	7	1	7	3	8	1	7	9	9	
$\beta_{h,3,29,1}$	9	11	13	13	15	14	20	17	14	12	6	0	4	8	9	10	10	3	9	7	3	3	7	4	4	
$\beta_{h,3,30,1}$	6	8	5	9	8	8	27	19	16	12	5	3	3	4	0	1	3	4	6	6	6	5	5	5	5	
$\beta_{h,3,31,1}$	12	10	10	8	8	13	10	8	20	22	14	0	2	4	7	11	8	8	4	8	15	6	3	3	3	
$\beta_{h,3,32,1}$	4	15	9	7	9	12	17	15	14	11	16	32	29	18	23	18	12	3	8	5	10	11	10	7	7	
$\beta_{h,3,33,1}$	6	11	14	12	15	15	11	13	13	0	3	13	13	0	3	5	9	11	8	11	8	2	0	5	4	
$\beta_{h,3,34,1}$	3	8	11	17	14	9	2	7	4	3	1	0	1	4	12	13	11	5	9	7	11	1	1	7	3	
$\beta_{h,3,35,1}$	12	9	16	13	18	8	12	10	8	3	4	6	19	19	11	5	1	5	9	4	1	2	3	1	1	
$\beta_{h,3,36,1}$	9	6	7	3	4	6	4	10	5	25	24	14	14	9	11	12	13	9	7	5	9	8	8	6	6	
$\beta_{h,3,37,1}$	7	8	6	9	10	8	9	5	4	9	19	8	14	15	9	6	28	5	13	8	11	8	7	8	7	
$\beta_{h,3,38,1}$	15	11	12	14	16	7	11	16	8	10	11	8	8	5	6	15	44	14	14	13	8	7	3	10	10	
$\beta_{h,3,39,1}$	8	14	11	12	7	11	17	32	22	15	6	2	4	2	4	9	14	24	54	17	16	7	3	3	3	
$\beta_{h,3,40,1}$	8	4	11	12	7	15	12	14	14	5	1	2	3	0	0	2	8	5	42	4	5	5	5	5	5	
$\beta_{h,3,41,1}$	9	16	16	23	22	19	14	20	29	24	19	3	5	1	4	1	5	2	1	0	0	5	6	11	41	
$\beta_{h,3,42,1}$	9	15	18	15	16	16	12	11	7	18	16	12	10	17	11	10	14	22	18	9	10	2	1	18	27	
$\beta_{h,3,43,1}$	17	22	19	16	16	12	11	7	18	16	12	10	17	11	10	14	22	18	9	10	2	1	18	27	27	
day: d = 4	$\beta_{h,4,1,1}$	12	12	10	11	10	10	11	11	12	10	9	9	11	12	13	14	12	11	12	13	14	12	10	10	11
	$\beta_{h,4,2,1}$	15	16	14	15	15	14	14	11	13	10	11	13	10	11	12	13	14	12	11	12	13	14	12	10	11
	$\beta_{h,4,3,1}$	10	14	18	16	15	12	13	9	10	7	6	10	11	12	13	16	14	13	16	14	13	15	17	14	
	$\beta_{h,4,4,1}$	13	14	16	17	15	11	14	14	18	16	14	18	16	13	16	13	16	13	16	14	13	12	12	13	14
	$\beta_{h,4,5,1}$	13	11	10	11	13	8	14	15	18	20	22	18	19	18	17	12	12	12	12	12	12	12	13	9	11
	$\beta_{h,4,6,1}$	14	14	13	13	11	12	13	12	10	11	12	13	14	14	14	14	14	14	14	14	14	14	16	13	11
	$\beta_{h,4,7,1}$	13	14	10	10	11	15	20	14	13	15	14	13	14	14	14	14	14	14	14	14	14	16	13	11	12
	$\beta_{h,4,8,1}$	10	10	12	12	13	11	10	17	18	14	14	12	13	10	12	9	10	10	12	9	10	10	11	13	9
	$\beta_{h,4,9,1}$	13	14	10	10	11	15	20	14	13	15	14	13	14	14	14	14	14	14	14	14	14	16	13	11	12
	$\beta_{h,4,10,1}$	10	10	12	12	13	11	10	17	18	14	14	12	13	10	12	9	10	10	12	9	10	10	11	13	9
	$\beta_{h,4,11,1}$	13	14	13	14	11	13	11	13	11	12	15	14	14	11	11	15	16	14	13	12	12	13	12	14	14
	$\beta_{h,4,12,1}$	10	10	12	12	13	11	10	17	18	14	14	12	13	10	12	9	10	10	12	9	10	10	11	13	9
	$\beta_{h,4,13,1}$	8	12	13	14	13	14	13	14	10	9	12	10	7	10	9	10	13	11	13	11	11	12	12	1	

Table 5: Mean occurrences of the LASSO_LS^{HQC} and BA_LS1 model parameters in the left and right panel, respectively. A heat-map is used to indicate the range of occurrence values between 0 (\rightarrow red) and 100 (\rightarrow green). Continued in Table 6.

	h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24		
day: $\Delta t = 1$	$\beta_{h,5,1,1}$	26	17	12	8	7	10	5	6	4	0	0	1	0	2	0	0	0	0	0	0	0	0	2	2	10	25
	$\beta_{h,5,2,1}$	22	19	14	9	8	9	10	6	5	0	1	2	3	5	3	5	4	2	4	2	2	1	1	1	3	4
	$\beta_{h,5,3,1}$	4	6	15	10	2	1	3	6	10	5	1	2	3	5	3	5	4	2	4	2	2	1	1	1	10	12
	$\beta_{h,5,4,1}$	4	21	34	43	19	8	7	6	5	10	12	7	5	10	7	8	12	6	1	0	2	2	3	3	0	0
	$\beta_{h,5,5,1}$	3	3	8	12	28	17	6	1	3	0	1	0	1	3	4	8	3	0	1	2	1	2	1	2	1	0
	$\beta_{h,5,6,1}$	15	10	7	4	16	34	27	12	13	4	5	1	0	3	6	4	3	10	15	13	10	6	0	0	3	0
	$\beta_{h,5,7,1}$	15	10	11	17	22	19	27	9	7	6	2	1	4	7	5	3	2	6	25	25	15	0	6	0	3	0
	$\beta_{h,5,8,1}$	9	10	10	11	16	11	9	10	5	4	0	2	7	7	9	13	7	13	9	3	2	2	2	2	2	2
	$\beta_{h,5,9,1}$	11	10	11	12	10	6	11	10	7	12	13	9	9	11	9	2	7	7	9	8	9	4	3	3	3	3
	$\beta_{h,5,10,1}$	15	18	5	3	8	12	18	24	16	12	5	5	5	6	11	20	19	14	4	9	11	3	3	1	3	1
	$\beta_{h,5,11,1}$	13	4	0	1	3	5	0	1	6	6	9	7	5	8	11	6	7	5	5	3	2	6	2	6	2	6
	$\beta_{h,5,12,1}$	7	7	8	4	7	6	4	2	2	1	15	18	15	11	3	3	10	1	3	1	0	0	0	1	0	0
	$\beta_{h,5,13,1}$	2	4	6	13	15	14	9	6	1	7	8	1	7	4	4	3	0	1	1	4	3	2	0	0	0	0
	$\beta_{h,5,14,1}$	0	1	2	0	1	1	4	4	2	0	2	9	7	4	2	0	0	1	1	6	3	2	1	1	1	1
	$\beta_{h,5,15,1}$	10	4	18	19	16	6	6	13	7	3	1	6	10	9	4	0	1	1	4	2	2	1	1	4	2	1
	$\beta_{h,5,16,1}$	4	3	3	6	9	9	7	19	6	7	3	1	0	14	17	18	3	9	9	1	2	1	1	2	1	3
	$\beta_{h,5,17,1}$	8	1	3	7	10	21	18	14	22	27	12	2	1	7	18	21	39	6	3	7	2	0	1	7	7	
	$\beta_{h,5,18,1}$	15	11	6	2	4	7	10	7	5	6	4	3	3	4	7	7	31	33	8	7	6	8	10	9	7	
	$\beta_{h,5,19,1}$	26	27	29	31	31	28	27	20	19	13	11	4	8	9	11	17	12	25	69	19	5	6	4	3	3	
	$\beta_{h,5,20,1}$	5	11	9	5	4	13	18	6	5	9	7	4	4	8	6	3	3	4	14	56	16	0	2	4	4	
	$\beta_{h,5,21,1}$	2	2	5	10	7	11	2	2	6	1	2	1	2	2	2	3	9	12	5	40	9	4	0	0	0	
	$\beta_{h,5,22,1}$	6	9	12	13	16	10	8	11	3	5	15	2	6	5	7	5	4	3	1	43	8	8	8	8	8	
	$\beta_{h,5,23,1}$	10	8	13	14	18	21	14	14	11	2	2	1	7	6	2	2	1	4	2	0	7	40	16	16	16	
	$\beta_{h,5,24,1}$	22	27	21	21	26	17	16	17	9	2	6	8	20	16	13	17	21	9	14	7	0	1	7	13	13	
day: $\Delta t = 10$	$\beta_{h,6,1,1}$	27	6	4	3	6	9	8	11	13	7	11	7	3	2	1	2	1	0	1	0	1	2	2	6	6	10
	$\beta_{h,6,2,1}$	23	27	19	9	8	2	1	2	2	0	1	0	1	1	3	3	4	1	3	5	4	1	3	5	4	1
	$\beta_{h,6,3,1}$	1	5	21	15	10	2	3	3	5	1	0	1	3	6	4	6	9	5	4	1	0	1	4	1	4	0
	$\beta_{h,6,4,1}$	0	10	6	6	2	0	0	2	3	1	2	1	2	6	12	4	3	13	8	11	8	4	4	4	0	
	$\beta_{h,6,5,1}$	5	13	25	30	38	24	10	2	1	7	7	6	8	7	12	7	8	9	5	7	2	0	5	3	6	
	$\beta_{h,6,6,1}$	21	13	9	9	22	63	14	8	6	1	2	4	4	2	2	1	1	1	4	3	0	1	6	1	6	
	$\beta_{h,6,7,1}$	1	5	10	11	12	51	77	43	31	23	5	5	6	17	22	15	8	14	21	13	9	8	14	14	14	
	$\beta_{h,6,8,1}$	5	7	10	9	4	10	26	49	20	17	11	6	3	6	13	10	4	4	10	12	20	12	3	7	7	
	$\beta_{h,6,9,1}$	10	6	12	13	2	4	3	6	30	3	1	3	7	14	16	20	22	9	6	6	12	11	14	5	3	
	$\beta_{h,6,10,1}$	8	11	4	4	4	7	9	22	32	29	17	7	8	7	8	7	12	7	2	3	6	3	3	3	3	
	$\beta_{h,6,11,1}$	3	10	10	5	6	8	8	11	21	34	11	4	4	1	2	3	9	13	9	5	3	1	5	1	5	
	$\beta_{h,6,12,1}$	10	5	13	4	1	5	7	1	6	0	13	26	25	8	5	3	3	5	2	1	5	1	5	1	2	
	$\beta_{h,6,13,1}$	9	14	19	17	14	11	5	5	3	2	6	11	3	2	1	1	1	1	0	0	3	2	2	2	2	
	$\beta_{h,6,14,1}$	18	17	11	15	10	10	3	15	16	8	3	5	21	28	20	4	2	0	0	1	9	13	15	8	8	
	$\beta_{h,6,15,1}$	8	17	22	24	21	2	9	4	2	6	6	7	11	27	25	11	7	17	6	7	2	5	4	4	4	
	$\beta_{h,6,16,1}$	4	2	0	1	2	2	1	4	4	2	3	7	12	28	42	20	1	3	10	11	12	4	11	11	11	
	$\beta_{h,6,17,1}$	3	5	8	10	8	3	6	12	9	7	8	14	20	27	24	42	59	12	1	1	2	1	3	6	6	
	$\beta_{h,6,18,1}$	9	4	5	10	19	18	19	25	21	12	17	10	14	19	26	18	31	81	18	3	5	2	11	8	8	
	$\beta_{h,6,19,1}$	7	6	8	10	9	16	14	18	14	18	11	3	2	0	1	2	4	5	11	12	75	28	9	3	7	
	$\beta_{h,6,20,1}$	16	8	10	14	19	17	12	8	6	3	4	1	0	1	1	1	1	1	1	64	17	2	6	1	0	
	$\beta_{h,6,21,1}$	8	1	4	5	7	5	5	9	7	6	2	0	3	2	6	8	11	5	2	17	42	14	2	0	0	
	$\beta_{h,6,22,1}$	11	15	25	19	29	19	13	9	10	12	1	1	1	2	6	8	9	3	2	3	2	46	16	14	16	
	$\beta_{h,6,23,1}$	11	20	24	28	30	28	21	13	12	4	6	1	1	3	1	1	2	6	8	9	3	3	2	8	38	
	$\beta_{h,6,24,1}$	20	17	17	19	30	29	23	28	21	14	13	11	11	13	14	22	20	23	23	7	2	4	25	25	16	

Note: See Table 4.

Table 6: Mean occurrences of the LASSO_LS^{HQC} and BA_LS1 model parameters in the left and right panel, respectively. A heat-map is used to indicate the range of occurrence values between 0 (\rightarrow red) and 100 (\rightarrow green). Continued in Table 7.

h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
$\mu_{h,7,1,1}$	37	8	16	12	5	10	32	36	20	10	6	6	5	4	4	3	4	4	2	2	2	7	3	25	17
$\mu_{h,7,2,1}$	12	27	11	9	5	12	7	6	2	1	0	0	3	0	3	0	3	0	3	1	0	2	0	0	2
$\mu_{h,7,3,1}$	8	13	22	11	6	10	14	12	10	7	1	1	0	0	0	2	7	0	1	0	0	1	0	0	2
$\mu_{h,7,4,1}$	10	21	18	20	33	4	10	7	11	11	12	9	8	10	6	8	8	4	2	1	0	1	0	0	
$\mu_{h,7,5,1}$	14	11	18	20	33	4	1	3	9	4	5	6	6	8	10	4	1	0	2	3	1	1	0	1	
$\mu_{h,7,6,1}$	10	6	10	10	38	71	10	5	8	4	8	5	6	7	5	7	5	2	2	2	2	6	22	13	20
$\mu_{h,7,7,1}$	9	19	14	13	31	65	99	49	38	17	15	11	13	15	19	32	17	12	17	23	26	22	13	20	
$\mu_{h,7,8,1}$	20	10	5	2	8	26	52	95	70	39	19	10	12	10	21	22	4	23	17	18	12	6	6	3	
$\mu_{h,7,9,1}$	6	7	5	1	6	8	6	18	62	67	32	18	14	6	9	12	11	6	9	6	4	6	14	9	
$\mu_{h,7,10,1}$	1	2	3	1	6	3	2	3	4	30	42	18	10	9	6	7	8	9	6	10	14	9	7	20	
$\mu_{h,7,11,1}$	8	4	4	5	10	13	11	13	6	10	46	21	0	2	1	4	1	1	2	5	2	4	2	2	
$\mu_{h,7,12,1}$	2	8	1	2	3	9	13	12	4	9	36	10	9	7	8	5	1	2	1	5	3	2	2	2	
$\mu_{h,7,13,1}$	11	19	13	5	6	10	6	12	11	4	13	28	16	7	5	1	1	2	1	6	3	3	3	3	
$\mu_{h,7,14,1}$	3	5	4	2	4	6	3	4	1	7	25	37	41	30	7	3	5	2	3	2	10	9	7	7	
$\mu_{h,7,15,1}$	2	3	7	19	8	5	4	2	3	4	12	19	21	30	16	7	1	3	2	8	2	3	4	4	
$\mu_{h,7,16,1}$	5	3	1	3	4	6	6	11	5	8	14	41	51	62	66	37	10	6	9	6	4	5	3	4	
$\mu_{h,7,17,1}$	17	26	20	12	18	14	5	11	8	15	19	14	16	16	15	34	48	22	10	4	1	4	3	2	
$\mu_{h,7,18,1}$	7	11	17	25	30	30	18	20	14	4	4	1	4	4	19	11	6	10	64	22	9	2	0	1	
$\mu_{h,7,19,1}$	5	15	8	14	17	30	15	13	10	16	10	6	8	13	18	18	17	12	9	50	25	2	6	2	
$\mu_{h,7,20,1}$	3	5	2	1	2	3	7	17	13	6	2	1	2	9	5	9	11	6	9	37	14	9	8	8	
$\mu_{h,7,21,1}$	17	14	5	16	19	18	11	8	3	3	0	1	4	3	4	6	4	2	4	41	23	12	12	12	
$\mu_{h,7,22,1}$	7	22	22	24	27	32	27	13	9	14	6	8	7	9	18	16	14	9	0	2	48	19	19	19	
$\mu_{h,7,23,1}$	10	12	17	26	35	27	39	28	22	17	13	12	12	12	16	15	18	20	24	4	10	16	16	16	
$\mu_{h,8,1,1}$	10	12	23	33	31	24	19	17	14	10	7	3	6	6	7	8	7	8	10	11	15	10	20	15	
$\mu_{h,8,2,1}$	8	7	6	6	3	9	12	10	12	4	5	3	4	15	13	4	0	1	2	5	6	9	8	1	
$\mu_{h,8,3,1}$	2	13	14	11	3	5	0	1	1	2	1	0	0	0	1	0	2	1	1	4	6	5	5	5	
$\mu_{h,8,4,1}$	23	20	11	19	11	7	2	5	11	2	2	9	4	5	6	9	5	5	16	20	17	12	18	18	
$\mu_{h,8,5,1}$	10	16	11	19	32	18	10	12	8	7	11	9	8	6	5	3	1	4	2	1	2	1	4	4	
$\mu_{h,8,6,1}$	12	13	15	19	20	25	7	10	8	2	1	3	5	5	4	3	4	2	0	1	0	0	1	3	
$\mu_{h,8,7,1}$	13	10	11	13	18	15	17	15	8	15	14	8	19	26	25	25	16	7	9	10	7	9	7	7	
$\mu_{h,8,8,1}$	8	10	8	10	9	10	4	10	11	7	4	0	0	0	0	1	0	1	2	1	0	7	7	1	
$\mu_{h,8,9,1}$	3	8	6	3	7	1	1	3	2	1	3	3	2	4	10	11	13	12	3	4	19	12	10	12	
$\mu_{h,8,10,1}$	2	3	7	10	12	15	15	18	5	1	2	1	4	4	5	9	9	7	4	1	11	9	8	2	
$\mu_{h,8,11,1}$	9	10	9	12	7	10	20	25	18	8	2	1	0	0	0	5	6	10	7	5	5	6	1	3	
$\mu_{h,8,12,1}$	7	20	9	6	4	7	4	6	6	2	6	16	14	9	1	1	2	7	6	3	1	2	8	8	
$\mu_{h,8,13,1}$	9	7	9	6	6	2	2	3	4	8	9	3	2	3	4	7	10	4	5	6	1	4	0	4	
$\mu_{h,8,14,1}$	3	4	4	2	5	12	14	4	2	3	5	4	5	8	16	15	10	5	1	1	0	1	0	1	
$\mu_{h,8,15,1}$	7	6	7	6	6	11	14	7	6	1	0	2	5	8	16	15	10	5	1	1	1	2	5	5	
$\mu_{h,8,16,1}$	5	2	5	9	9	10	19	21	8	7	17	18	23	25	23	5	0	1	2	0	1	1	1	2	
$\mu_{h,8,17,1}$	5	7	6	7	4	6	9	4	6	9	5	8	13	14	16	23	39	13	14	28	14	22	22	17	
$\mu_{h,8,18,1}$	9	1	3	5	10	16	12	9	5	8	11	7	15	14	8	8	12	5	2	19	10	8	2	2	
$\mu_{h,8,19,1}$	5	10	18	20	17	22	24	11	1	7	11	10	7	5	8	6	37	18	3	2	3	7	2	2	
$\mu_{h,8,20,1}$	10	7	8	13	16	21	16	13	22	9	4	6	2	3	9	6	6	7	21	6	5	0	1	3	
$\mu_{h,8,21,1}$	14	7	9	8	8	19	7	22	16	12	7	5	4	6	9	10	9	6	9	6	6	0	1	3	
$\mu_{h,8,22,1}$	17	9	4	7	13	16	18	13	9	2	1	1	1	1	1	0	1	2	1	5	1	2	30	5	
$\mu_{h,8,23,1}$	13	29	35	34	36	30	21	17	16	12	4	10	11	5	6	8	9	9	8	1	28	20	20	20	
$\mu_{h,8,24,1}$	13	37	36	28	31	35	34	19	11	6	1	7	9	7	9	6	6	13	3	4	19	24	19	24	

Note: See Table 5.

Table 7: Mean occurrences of the LASSO_LS^{HQC} and BA_LS1 model parameters in the left and right panel, respectively. A heat-map is used to indicate the range of occurrence values between 0 (→ red) and 100 (→ green).

h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
$\beta_{h,1,15}$	58	74	87	86	80	59	35	25	22	16	4	0	1	3	7	6	12	5	8	6	6	5	3	2	
$\beta_{h,2,15}$	2	2	7	16	22	20	8	10	4	3	0	2	1	2	9	7	3	2	1	3	2	1	5	2	
$\beta_{h,3,15}$	12	20	31	34	28	31	8	9	10	4	3	2	3	5	6	2	1	3	3	1	3	9	10	14	16
$\beta_{h,4,15}$	13	18	21	30	26	33	15	17	14	10	6	11	10	9	15	12	22	20	12	10	9	5	17	9	
$\beta_{h,5,15}$	18	23	26	27	29	33	15	9	6	7	5	7	8	11	10	3	6	7	3	5	5	8	8	8	
$\beta_{h,6,15}$	9	26	25	34	39	14	18	25	19	12	5	3	10	11	21	20	12	5	10	16	14	5	18	15	
$\beta_{h,7,15}$	7	12	15	17	6	9	25	16	5	6	0	1	2	3	3	2	3	1	1	6	4	9	5	11	
$\beta_{h,8,15}$	15	7	9	7	17	19	7	22	23	14	18	11	7	7	9	8	12	13	16	13	9	14	11	11	
$\beta_{h,9,15}$	11	10	8	16	17	18	20	38	40	43	29	18	11	7	9	15	3	9	31	30	21	24	17	14	
$\beta_{h,10,15}$	32	35	46	48	52	41	33	32	36	35	36	41	38	29	30	27	20	3	3	23	20	20	23	3	
$\beta_{h,11,15}$	7	10	18	16	13	21	20	22	15	14	21	11	13	9	3	4	9	9	16	25	14	25	24	20	
$\beta_{h,12,15}$	23	24	20	19	26	24	15	23	15	5	13	5	6	7	15	6	17	16	17	6	11	8	12	6	
$\beta_{h,13,15}$	23	23	20	11	13	29	22	17	11	24	28	26	22	13	14	5	4	6	12	18	14	5	6	8	
$\beta_{h,14,15}$	17	18	13	16	14	14	7	15	26	19	10	6	9	10	8	6	7	4	12	13	4	9	10	16	
$\beta_{h,15,15}$	15	12	16	16	16	40	32	13	12	12	16	14	20	15	22	18	13	4	4	7	4	9	7	11	
$\beta_{h,16,15}$	9	17	6	6	6	18	25	21	24	20	7	3	7	8	7	8	11	6	9	12	3	5	7	7	
$\beta_{h,17,15}$	80	72	70	70	95	98	87	98	99	100	99	84	100	90	83	93	97	84	77	54	51	52	27	38	
$\beta_{h,18,15}$	46	45	52	44	39	77	91	77	72	71	75	75	78	75	86	84	85	85	87	73	72	66	46	46	
$\beta_{h,19,15}$	29	33	40	40	45	33	39	43	41	39	43	50	32	25	29	28	37	41	33	31	22	22	18	27	
$\beta_{h,20,15}$	33	32	33	23	25	26	24	31	25	20	18	18	19	21	25	26	30	45	47	47	19	13	9	11	
$\beta_{h,21,15}$	24	20	15	22	25	41	25	24	28	26	32	28	28	28	34	34	25	31	49	44	38	32	18	18	
$\beta_{h,22,15}$	42	44	40	35	33	34	23	23	44	44	40	31	35	40	45	65	58	78	93	91	86	50	42	42	
$\beta_{h,23,15}$	79	81	78	69	69	100	100	100	100	100	97	99	100	100	100	100	86	88	86	83	67	72	74	74	
$\beta_{h,24,15}$	14	23	34	33	32	41	29	28	15	4	3	5	8	8	9	11	11	12	17	5	21	33	24	21	
$\beta_{h,25,15}$	43	39	31	24	20	25	37	35	31	18	11	6	4	3	5	8	13	17	26	21	11	12	11	16	
$\beta_{h,26,15}$	35	37	39	20	21	29	21	21	26	17	13	4	7	8	8	5	11	5	11	12	9	18	21	15	
$\beta_{h,27,15}$	16	17	15	9	18	15	15	16	23	6	12	6	12	19	13	11	15	12	7	18	9	6	8	5	
$\beta_{h,28,15}$	23	24	25	27	36	32	24	23	17	4	3	1	3	2	1	3	2	5	13	7	13	14	11	7	
$\beta_{h,29,15}$	22	28	29	25	29	24	23	17	21	21	16	13	18	20	27	28	25	19	21	27	29	41	29	22	
$\beta_{h,30,15}$	37	40	78	65	57	60	36	52	40	38	37	35	41	39	32	26	37	57	56	53	41	45	46	37	
$\beta_{h,31,15}$	60	50	42	36	37	56	30	23	30	22	28	24	20	22	21	15	33	28	44	43	32	31	6	30	
$\beta_{h,32,15}$	27	24	25	41	39	29	42	28	28	13	21	23	30	36	35	42	51	38	44	59	34	31	25	32	
$\beta_{h,33,15}$	42	43	41	49	49	37	40	59	57	46	39	32	35	36	48	47	52	60	45	36	29	13	21	29	
$\beta_{h,34,15}$	23	24	30	31	33	46	68	72	63	98	45	45	45	37	44	51	46	54	44	36	21	46	25	26	
$\beta_{h,35,15}$	36	29	31	34	27	40	68	74	82	79	70	67	71	59	58	59	57	43	45	49	57	42	41	39	
$\beta_{h,36,15}$	37	40	33	36	37	47	68	67	51	37	22	19	26	20	33	35	31	20	26	34	29	4	1	16	
$\beta_{h,37,15}$	62	58	44	46	28	40	35	29	19	26	13	13	16	19	23	25	20	27	23	15	19	28	19	16	
$\beta_{h,38,15}$	36	35	49	54	47	65	43	41	39	48	47	42	46	39	38	32	45	53	67	56	45	40	31	31	
$\beta_{h,39,15}$	20	22	28	27	27	36	50	32	34	34	33	29	25	18	17	24	24	19	21	22	21	15	20	17	
$\beta_{h,40,15}$	50	36	34	32	26	27	30	26	22	27	24	34	35	33	29	33	29	33	29	33	30	13	14	25	
$\beta_{h,41,15}$	31	32	25	30	27	29	38	38	34	20	25	24	25	24	25	24	26	26	31	39	35	36	11	11	
$\beta_{h,42,15}$	47	29	26	35	33	31	22	26	39	36	17	18	14	17	31	17	32	23	21	35	32	28	17	17	
$\beta_{h,1,10}$	28	33	30	45	45	37	32	30	24	21	17	12	13	14	13	13	14	19	15	15	13	16	15	17	14
$\beta_{h,2,10}$	15	13	14	13	14	15	14	19	10	9	9	11	12	11	12	11	12	14	14	22	11	12	13	10	12
$\beta_{h,3,10}$	13	16	13	10	15	12	13	9	13	11	9	7	8	9	10	9	8	8	9	10	11	14	17	16	
$\beta_{h,4,10}$	18	17	11	9	8	9	17	15	13	12	14	14	16	15	16	17	18	19	15	14	16	17	16	14	
$\beta_{h,5,10}$	11	11	9	8	9	11	17	13	12	14	14	16	15	16	17	18	19	15	14	12	10	11	11	12	
$\beta_{h,6,10}$	13	13	16	14	16	15	16	17	14	13	13	13	13	14	13	14	13	14	13	11	12	14	12	13	
$\beta_{h,7,10}$	15	18	16	18	18	18	18	17	14	13	11	12	17	16	15	15	14	13	13	14	12	13	15	12	
$\beta_{h,8,10}$	17	18	16	18	12	12	17	15	14	10	8	7	9	9	9	9	9	10	8	8	11	14	12		
$\beta_{h,9,10}$	17	16	17	17	17	17	16	20	15	19	18	16	16	16	15	14	11	13	15	15	16	21	20	20	
$\beta_{h,10,10}$	41	11	11	12	12	14	20	27	20	15	14	15	15	12	14	14	18	21	21	16	10	13	15	11	
$\beta_{h,11,10}$	17	19	19	19	14	9	11	11	13	13	14	15	15	18	17	16	13	12	13	12	13	16	13	14	
$\beta_{h,12,10}$	42	11	12	13	14	15	13	15	11	12	13	12	12	14	14	14	14	14	12	11	12	13	10	12	
$\beta_{h,13,10}$	9	11	10	9	12	11	12	13	17	15	14	14	14	13	14	14	16	15	10	12	11	11	11	11	
$\beta_{h,14,10}$	9	11	14	16	16	18	18	17	14	14	14	14	15	18	17	16	15	14	16	15	15	13	12	12	
$\beta_{h,15,10}$	10	12	12	1																					

Table 8: Mean occurrences of factors at the **fEXPERT** model across all six datasets and full out-of-sample test period in percentage points. A heat-map is used to indicate the range of occurrence values between 0 (\rightarrow red) and 100 (\rightarrow green).

h	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
α_1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100
α_2	82	96	100	100	100	100	86	100	99	90	81	85	88	89	83	77	81	62	53	62	49	51	62	55	55
α_3	78	95	100	100	96	97	83	90	87	84	75	77	83	84	80	76	81	59	48	54	43	47	48	43	43
α_4	65	84	94	100	96	97	74	76	75	67	58	64	73	70	67	63	62	35	31	26	21	24	25	27	27
α_5	47	68	84	95	87	90	65	70	62	64	53	47	58	58	60	57	51	35	21	23	17	18	21	21	21
α_6	23	37	64	77	74	62	60	53	55	43	27	15	35	42	30	29	26	7	2	5	6	6	0	7	7
α_7	13	26	54	68	64	48	47	48	39	36	24	11	23	24	26	23	17	7	2	5	5	1	0	7	7
α_8	10	24	50	60	56	43	27	31	35	27	17	9	17	18	20	21	15	5	2	5	5	0	0	2	2
α_9	10	21	47	59	55	36	20	26	32	20	11	7	11	12	13	18	15	3	2	4	5	0	0	2	2
α_{10}	10	20	44	57	53	33	16	24	23	11	6	5	8	10	13	17	10	2	1	4	5	0	0	1	1
α_{11}	7	11	29	50	48	27	15	19	12	7	5	4	7	9	12	16	8	0	1	4	5	0	0	1	1
α_{12}	6	11	22	42	39	24	15	19	12	6	4	3	7	7	11	16	8	0	0	4	1	0	0	1	1
α_{13}	6	8	17	35	35	20	14	11	6	1	1	2	3	4	10	15	6	0	0	1	1	0	0	1	1
α_{14}	6	8	15	31	33	20	14	10	5	1	1	1	0	0	4	12	3	0	0	0	1	0	0	0	0
α_{15}	6	6	14	31	29	17	6	2	2	1	1	1	0	0	4	7	3	0	0	0	0	0	0	0	0
α_{16}	4	5	10	30	28	14	4	1	1	1	1	1	0	0	4	6	2	0	0	0	0	0	0	0	0
α_{17}	1	2	4	12	14	5	0	0	1	1	0	1	0	0	4	6	2	0	0	0	0	0	0	0	0
α_{18}	1	1	3	7	7	1	0	0	0	0	0	1	0	0	2	3	2	0	0	0	0	0	0	0	0
α_{19}	1	1	3	5	7	1	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0
α_{20}	0	0	0	1	3	0	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0
α_{21}	0	0	0	1	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
α_{22}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
α_{23}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0
α_{24}	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Note: The table reports the mean occurrences of factors in Eq. (20) with the number of factors being specified by the BIC.

5 Conclusion

There is an enormous literature concerning forecasting day-ahead electricity prices and it is growing *pari passu* with the availability of more sophisticated forecasting procedures. However, the most challenging issue of the literature, which is the integration of many potential predictors with changing predictive abilities over time into the models, still requires further attention. In that sense, in this study, we have introduced a new shrinkage method to the EPF literature, namely bagging, and applied this procedure in a multivariate framework to forecast day-ahead electricity prices for six major power markets in comparison with a number of stochastic models that are commonly applied in the literature.

Overall, our findings suggest that bagging is a very competitive and promising tool for the estimation of large scale models and for improving forecast accuracy of day-ahead electricity prices with its superior forecasting performance and its computationally simple algorithm. More specifically, we observe that the bagging approach significantly outperforms all benchmark models for all markets, while it outperforms the popular LASSO approach for four markets, in both daily and hourly scales. For the rest two markets, where bagging fails to beat the LASSO in terms of forecast accuracy, we observe that forecasts from bagging are not significantly worse than those of LASSO. We further deduce that the forecast performances of the bagging and LASSO differentiate from each other across hours of the day as well so that bagging appears to provide higher forecast accuracy during the working hours in almost all markets. When analyzing the selected explanatory variables or the occurrence of the variables in both approaches, we deduce that differentiation in the forecast performances of the LASSO and bagging approaches mainly arise from their structural differences in the explanatory variables selection process.

Although we perform our forecasting study in a multivariate framework, we are aware of the limitation that the multivariate setting lacks intra-day hourly dependencies of day-ahead electricity prices. To account for these dependencies, we further augment all our models with latent factors observed from a standard PCA approach. In that respect, our results suggest that the inclusion of the latent factors provides a substantial improve-

ment in the forecasts from models covering a specific subset of explanatory variables, while almost no improvement is obtained in the forecasts from the LASSO and bagging approaches. This finding implies that large scale models already contain, to some extent, the intra-day dependencies of electricity prices even if they are formed in a multivariate context, and therefore, there is no need to deal with the complexity of the implementation of a univariate framework.

Declaration of Competing Interest

None.

Acknowledgments

The authors are grateful to Ayşegül Çorakçı and Uğur Soytaş for their thoughtful comments and suggestions. Responsibility for any errors and omissions, however, lies entirely with the authors.

References

- Aggarwal, S. K., Saini, L. M., and Kumar, A. (2009). Electricity price forecasting in deregulated markets: A review and evaluation. *International Journal of Electrical Power and Energy Systems*.
- Agrawal, R. K., Muchahary, F., and Tripathi, M. M. (2019). Ensemble of relevance vector machines and boosted trees for electricity price forecasting. *Applied Energy*.
- Andrews, D. W. (2004). The block-block bootstrap: Improved asymptotic refinements. *Econometrica*.
- Breiman, L. (1996a). Bagging predictors. *Machine Learning*.
- Breiman, L. (1996b). Heuristics of instability and stabilization in model selection. *Annals of Statistics*.
- Bühlmann, P. and Yu, B. (2002). Analyzing bagging.
- Conejo, A. J., Contreras, J., Espínola, R., and Plazas, M. A. (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting*.
- Crespo Cuaresma, J., Hlouskova, J., Kossmeier, S., and Obersteiner, M. (2004). Forecasting electricity spot-prices using linear univariate time-series models. *Applied Energy*.
- Dantas, T. M. and Cyrino Oliveira, F. L. (2018). Improving time series forecasting: An approach combining bootstrap aggregation, clusters and exponential smoothing. *International Journal of Forecasting*, 34(4):748–761.
- Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*.
- Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*.
- Friedman, J. H. and Hall, P. (2007). On bagging and nonlinear estimation. *Journal of Statistical Planning and Inference*.

- G. P., G. and S., V. (2013). Determinants of Electricity Price in Competitive Power Market. *International Journal of Business and Management*.
- Garcia, R. C., Contreras, J., van Akkeren, M., and Garcia, J. B. C. (2005). A GARCH forecasting model to predict day-ahead electricity prices. *IEEE Transactions on Power Systems*.
- Gonçalves, S. and Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. *Journal of Econometrics*.
- Hall, P., Horowitz, J. L., and Jing, B. Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika*.
- Härdle, W. K. and Trück, S. (2010). The dynamics of hourly electricity prices The dynamics of hourly electricity prices ★.
- Hastie, T., Tibshirani, R., and Wainwright, M. (2015). *Statistical learning with sparsity: The lasso and generalizations*.
- Hong, T., Pinson, P., Fan, S., Zareipour, H., Troccoli, A., and Hyndman, R. J. (2016). Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond.
- Huang, H. and Lee, T. H. (2010). To combine forecasts or to combine information? *Econometric Reviews*, 29(5):534–570.
- Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*.
- Inoue, A. and Kilian, L. (2008). How useful is bagging in forecasting economic time series? A case study of U.S. Consumer Price Inflation. *Journal of the American Statistical Association*, 103(482):511–522.
- Jin, S., Su, L., and Ullah, A. (2014). Robustify Financial Time Series Forecasting with Bagging. *Econometric Reviews*, 33(5-6):575–605.

- Jordan, S. J., Vivian, A., and Wohar, M. E. (2017). Forecasting market returns: bagging or combining? *International Journal of Forecasting*.
- Kanamura, T. and Ohashi, K. (2008). On transition probabilities of regime switching in electricity prices. *Energy Economics*.
- Knittel, C. R. and Roberts, M. R. (2005). An empirical examination of restructured electricity prices. *Energy Economics*.
- Kristiansen, T. (2012). Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy*.
- Kunsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics*.
- Lago, J., De Ridder, F., Vranx, P., and De Schutter, B. (2018). Forecasting day-ahead electricity prices in Europe: The importance of considering market integration. *Applied Energy*.
- Lee, T.-h., Tu, Y., and Ullah, A. (2010). Bagging nonparametric and semiparametric forecasts with constraints. (951).
- Ludwig, N., Feuerriegel, S., and Neumann, D. (2015). Putting Big Data analytics to work: Feature selection for forecasting electricity prices using the LASSO and random forests. *Journal of Decision Systems*.
- Maciejowska, K., Nowotarski, J., and Weron, R. (2016). Probabilistic forecasting of electricity spot prices using Factor Quantile Regression Averaging. *International Journal of Forecasting*.
- Maciejowska, K. and Weron, R. (2015). Forecasting of daily electricity prices with factor models: utilizing intra-day and inter-zone relationships. *Computational Statistics*, 30(3):805–819.
- Maciejowska, K. and Weron, R. (2016). Short- and Mid-Term Forecasting of Baseload Electricity Prices in the U.K.: The Impact of Intra-Day Price Relationships and Market Fundamentals. *IEEE Transactions on Power Systems*, 31(2):994–1005.

- Marcjasz, G., Uniejewski, B., and Weron, R. (2020). Beating the naïve-combining lasso with naïve intraday electricity price forecasts. *Energies*.
- Misiorek, A., Trueck, S., and Weron, R. (2006). Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models. *Studies in Nonlinear Dynamics and Econometrics*.
- Mount, T. D., Ning, Y., and Cai, X. (2006). Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics*.
- Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent. Technical report.
- Nogales, F. J., Contreras, J., Conejo, A. J., and Espínola, R. (2002). Forecasting next-day electricity prices by time series models. *IEEE Transactions on Power Systems*.
- Nowotarski, J., Raviv, E., Trück, S., and Weron, R. (2014). An empirical comparison of alternative schemes for combining electricity spot price forecasts. *Energy Economics*.
- Nowotarski, J. and Weron, R. (2016). On the importance of the long-term seasonal component in day-ahead electricity price forecasting. *Energy Economics*, 57:228–235.
- Rapach, D. E. and Strauss, J. K. (2010). Bagging or combining (or both)? An analysis based on forecasting U.S. employment growth. *Econometric Reviews*, 29(5):511–533.
- Raviv, E., Bouwman, K. E., and van Dijk, D. (2015). Forecasting day-ahead electricity prices: Utilizing hourly prices. *Energy Economics*.
- Shahidehpour, M., Yamin, H., and Li, Z. (2002). *Market Operations in Electric Power Systems*.
- Stock, J. H. and Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices.
- Stock, J. H. and Watson, M. W. (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business and Economic Statistics*, 30(4):481–493.

- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso Author (s): Robert Tibshirani Source : Journal of the Royal Statistical Society . Series B (Methodological), Vol . 58 , No . 1 (1996), Published by : Wiley for the Royal Statistical Society Stable URL. 58(1):267–288.
- Uniejewski, B., Marcjasz, G., and Weron, R. (2019). Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO. *International Journal of Forecasting*, (xxxx).
- Uniejewski, B., Nowotarski, J., and Weron, R. (2016). Automated variable selection and shrinkage for day-ahead electricity price forecasting. *Energies*, 9(8).
- Uniejewski, B. and Weron, R. (2018). Efficient forecasting of electricity spot prices with expert and LASSO models. *Energies*.
- Uniejewski, B., Weron, R., and Ziel, F. (2018). Variance Stabilizing Transformations for Electricity Spot Price Forecasting. *IEEE Transactions on Power Systems*, 33(2):2219–2229.
- Vehviläinen, I. and Pyykkönen, T. (2005). Stochastic factor model for electricity spot price - The case of the Nordic market. *Energy Economics*.
- Weron, R. (2006). *Modeling and forecasting electricity loads and prices: A statistical approach*.
- Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 30(4):1030–1081.
- Weron, R. and Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*.
- Zhu, X., Li, L., Zhou, K., Zhang, X., and Yang, S. (2018). A meta-analysis on the price elasticity and income elasticity of residential electricity demand. *Journal of Cleaner Production*.

- Ziel, F. (2016). Forecasting Electricity Spot Prices Using Lasso: On Capturing the Autoregressive Intraday Structure. *IEEE Transactions on Power Systems*, 31(6):4977–4987.
- Ziel, F., Steinert, R., and Husmann, S. (2015). Efficient modeling and forecasting of electricity spot prices. *Energy Economics*, 47:98–111.
- Ziel, F. and Weron, R. (2018). Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks. *Energy Economics*, 70:396–420.