Automated Variable Selection and Shrinkage for Day-Ahead Electricity Price Forecasting

Authors:

Bartosz Uniejewski, Jakub Nowotarski, Rafa? Weron

Date Submitted: 2019-01-07

Keywords: lasso, elastic net, ridge regression, stepwise regression, autoregression, variable selection, day-ahead market, electricity price forecasting

Abstract:

In day-ahead electricity price forecasting (EPF) variable selection is a crucial issue. Conducting an empirical study involving state-of-the-art parsimonious expert models as benchmarks, datasets from three major power markets and five classes of automated selection and shrinkage procedures (single-step elimination, stepwise regression, ridge regression, lasso and elastic nets), we show that using the latter two classes can bring significant accuracy gains compared to commonly-used EPF models. In particular, one of the elastic nets, a class that has not been considered in EPF before, stands out as the best performing model overall.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):	LAPSE:2019.0130
Citation (this specific file, latest version):	LAPSE:2019.0130-1
Citation (this specific file, this version):	LAPSE:2019.0130-1v1

DOI of Published Version: https://doi.org/10.3390/en9080621

energies

MDPI

Article

Automated Variable Selection and Shrinkage for Day-Ahead Electricity Price Forecasting

Bartosz Uniejewski, Jakub Nowotarski and Rafał Weron *

Department of Operations Research, Wrocław University of Technology, 50-370 Wrocław, Poland; uniejewskibartosz@gmail.com (B.U.); jakub.nowotarski@pwr.edu.pl (J.N.)

* Correspondence: rafal.weron@pwr.edu.pl; Tel.: +48-71-320-4525

Academic Editor: Javier Contreras

Received: 5 July 2016; Accepted: 29 July 2016; Published: 5 August 2016

Abstract: In day-ahead electricity price forecasting (EPF) variable selection is a crucial issue. Conducting an empirical study involving state-of-the-art parsimonious expert models as benchmarks, datasets from three major power markets and five classes of automated selection and shrinkage procedures (single-step elimination, stepwise regression, ridge regression, lasso and elastic nets), we show that using the latter two classes can bring significant accuracy gains compared to commonly-used EPF models. In particular, one of the elastic nets, a class that has not been considered in EPF before, stands out as the best performing model overall.

Keywords: electricity price forecasting; day-ahead market; autoregression; variable selection; stepwise regression; ridge regression; lasso; elastic net

JEL: C14, C22, C51, C53, Q47

1. Introduction

Alongside short-term load forecasting, short-term electricity price forecasting (EPF) has become a core process of an energy company's operational activities [1]. The reason is quite simple. A 1% improvement in the mean absolute percentage error (MAPE) in forecasting accuracy would result in about 0.1%–0.35% cost reductions from short-term EPF [2]. In dollar terms, this would translate into savings of ca. $1.5 million per year for a typical medium-size utility with a 5-GW peak load [3].

As has been noted in a number of studies, be it statistical or computational intelligence, a key point in EPF is the appropriate choice of explanatory variables [1,4–11]. The typical approach has been to select predictors in an ad hoc fashion, sometimes using expert knowledge, seldom based on some formal validation procedures. Very rarely has an automated selection or shrinkage procedure been carried out in EPF, especially for a large set of initial explanatory variables.

Early examples of formal variable selection in EPF include Karakatsani and Bunn [12] and Misiorek [13], who used stepwise regression to eliminate statistically insignificant variables in parsimonious autoregression (AR) and regime-switching models for individual load periods. Amjady and Keynia [4] proposed a feature selection algorithm that utilized the mutual information technique. (for later applications, see, e.g., [11,14,15]). In an econometric setup, Gianfreda and Grossi [5] computed p-values of the coefficients of a regression model with autoregressive fractionally integrated moving average disturbances (Reg-ARFIMA) and in one step eliminated all statistically-insignificant variables. In a study concerning the profitability of battery storage, Barnes and Balda [16] utilized ridge regression to compute forecasts of the New York Independent System Operator (NYISO) electricity prices for a model with more than 50 regressors.

More recently, González et al. [17] used random forests to identify important explanatory variables among the 22 considered. Ludwig et al. [7] used both random forests and the least absolute shrinkage

and selection operator (i.e., lasso or LASSO) as a feature selection algorithm to choose the relevant out of the 77 available weather stations. In a recent neural network study, Keles et al. [11] combined the *k-nearest-neighbor* algorithm with backward elimination to select the most appropriate input variables out of more than 50 fundamental parameters or lagged versions of these parameters. Finally, Ziel et al. [9,18] used the lasso to sparsify very large sets of model parameters (well over 100). They used time-varying coefficients to capture the intra-day dependency structure, either using B-splines and one large regression model for all hours of the day [9] or, more efficiently, using a set of 24 regression models for the 24 h of the day [18].

However, a thorough study involving state-of-the-art parsimonious expert models as benchmarks, data from diverse power markets and, most importantly, a set of different selection or shrinkage procedures is still missing in the literature. In particular, to our best knowledge, elastic nets have not been applied in the EPF context at all. It is exactly the aim of this paper to address these issues. We perform an empirical study that involves:

- nine variants of three parsimonious autoregressive model structures with exogenous variables (ARX): one originally proposed by Misiorek et al. [19] and later used in a number of EPF studies [13,18,20–27], one which evolved from it during the successful participation of TEAM POLAND in the Global Energy Forecasting Competition 2014 (GEFCom2014; see [28–30]) and an extension of the former, which creates a stronger link with yesterday's prices and additionally considers a second exogenous variable (zonal load or wind power),
- three two-year long, hourly resolution test periods from three distinct power markets (GEFCom2014, Nord Pool and the U.K.),
- nine variants of five classes of selection and shrinkage procedures: single-step elimination of insignificant predictors (without or with constraints), stepwise regression (with forward selection or backward elimination), ridge regression, lasso and three elastic nets (with $\alpha = 0.25$, 0.5 or 0.75),
- model validation in terms of the robust weekly-weighted mean absolute error (WMAE; see [1]) and the Diebold–Mariano (DM; see [31]) test

and draw statistically-significant conclusions of high practical value.

The remainder of the paper is structured as follows. In Section 2, we introduce the datasets. Next, in Section 3, we first discuss the iterative calibration and forecasting scheme, then describe the techniques considered for price forecasting: a simple naive benchmark, nine variants of three parsimonious ARX-type model structures and five classes of selection and shrinkage procedures. In Section 4, we summarize the empirical findings. Namely, we evaluate the quality of point forecasts in terms of WMAE errors, run the DM tests to formally assess the significance of differences in the forecasting performance and analyze variable selection for the best performing elastic net model. Finally, in Section 5 wrap up the results and conclude.

2. Datasets

The datasets used in this empirical study include three spot market time series. The first one comes from the Global Energy Forecasting Competition 2014 (GEFCom2014), the largest energy forecasting competition to date [28]. The dataset includes three time series at hourly resolution: locational marginal prices, day-ahead predictions of system loads and day-ahead predictions of zonal loads and covers the period 1 January 2011–14 December 2013; see Figure 1. The origin of the data has never been revealed by the organizers. The full dataset is now available as supplementary material accompanying [28] (Appendix A); however, during the competition, the information set was being extended on a weekly basis to prevent 'peeking' into the future. The dataset was preprocessed by the organizers and does not include any missing or doubled values.

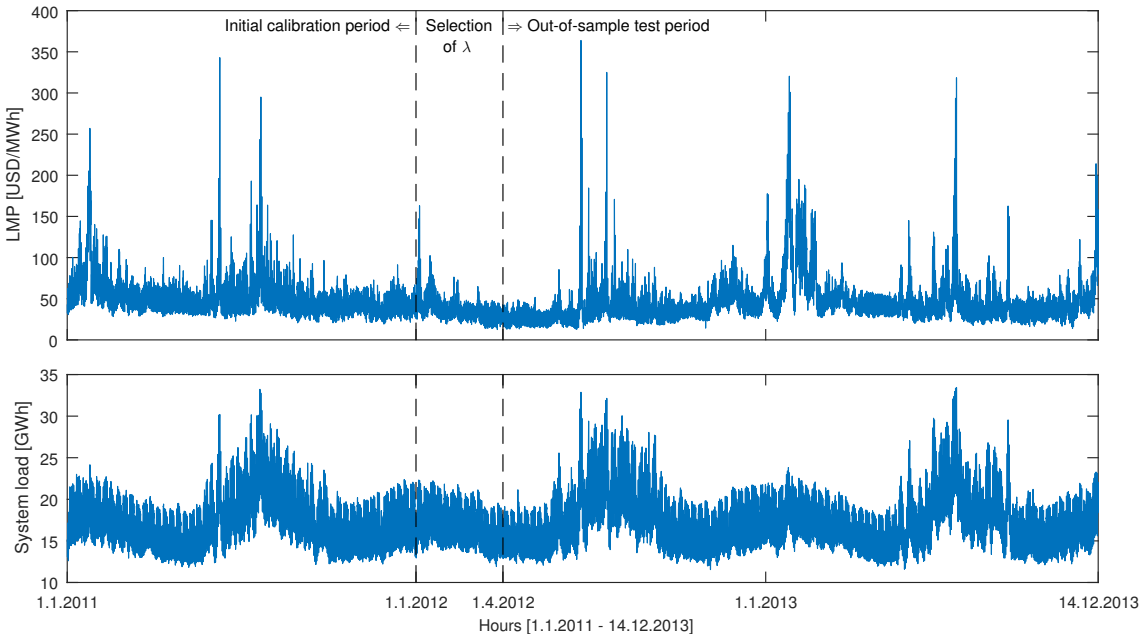

Figure 1. GEFCom2014 hourly locational marginal prices (LMP; top) and hourly day-ahead predictions of the system load (bottom) for the period 1 January 2011–14 December 2013. The day-ahead predictions of the zonal load are generally indistinguishable from those of the system load at this resolution; see Figure 8 in [28]. The vertical dashed lines mark the beginning of the 91-day period for selecting λ's (ridge regression, lasso, elastic nets) and the beginning of the 623-day long out-of-sample test period.

The second dataset comes from one of the major European power markets: Nord Pool (NP). It comprises hourly system prices, hourly consumption prognosis for four Nordic countries (Denmark, Finland, Norway and Sweden) and hourly wind prognosis for Denmark and covers the period 1 January 2013–29 March 2016; see Figure 2. The time series were constructed using data published by the Nordic power exchange Nord Pool (www.nordpoolspot.com) and preprocessed to account for missing values and changes to/from the daylight saving time, analogously as in [20] (Section 4.3.7). The missing data values (corresponding to the changes to the daylight saving/summer time; moreover, eight out of 28,392 hourly consumption figures were missing for Norway) were substituted by the arithmetic average of the neighboring values. The 'doubled' values (corresponding to the changes from the daylight saving/summer time) were substituted by the arithmetic average of the two values for the 'doubled' hour.

The third dataset comes from N2EX, the U.K. day-ahead power market operated by Nord Pool. It comprises hourly system prices for the period 1 January 2013–29 March 2016; see Figure 3. The time series was constructed using data published by Nord Pool (www.nordpoolspot.com) and, like the second dataset, preprocessed to account for changes to/from the daylight saving time. Note that the U.K. dataset includes only prices, as no day-ahead forecasts of fundamental variables were available to us. Hence, models calibrated to the U.K. data are 'pure price' models. To better see the effect of excluding fundamentals from forecasting models, we use the GEFCom2014 dataset twice, once with fundamentals (system and zonal load forecasts; to compare with the results for Nord Pool) and once without them.

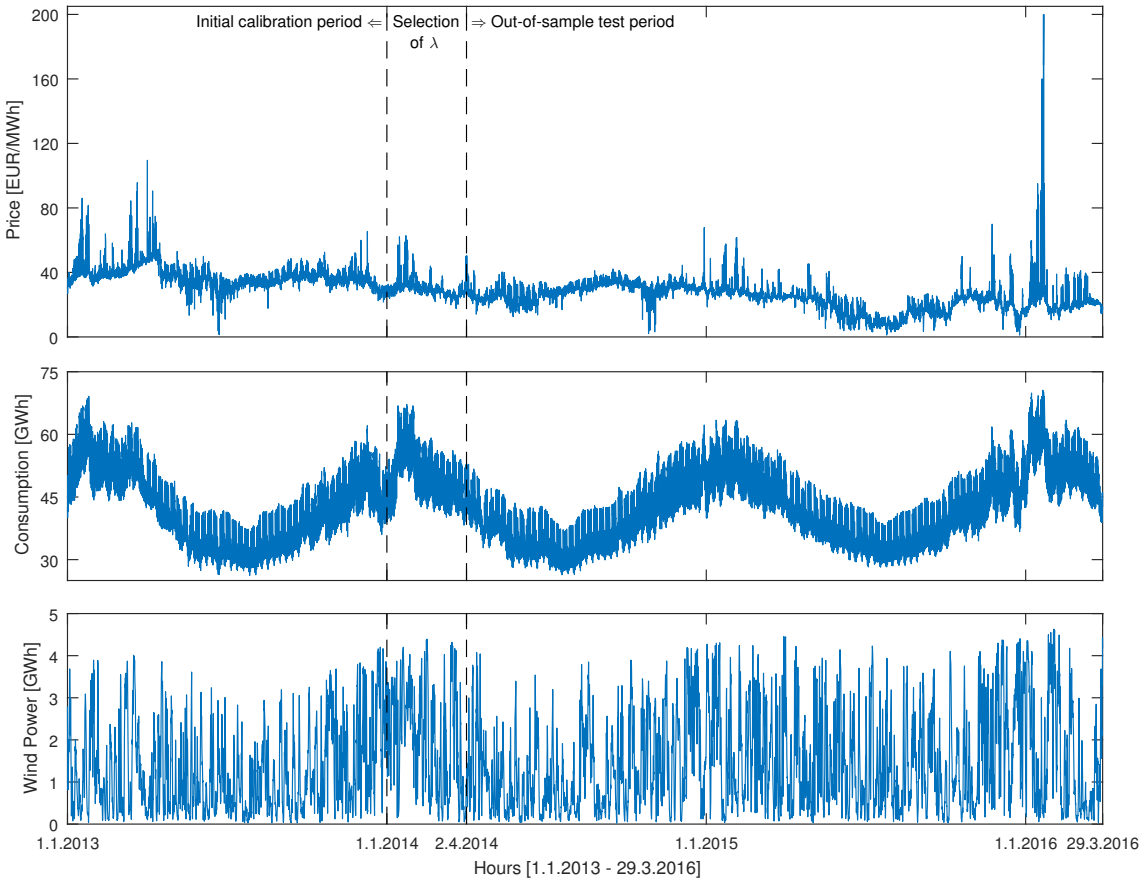

Figure 2. Nord Pool hourly system prices (top), hourly consumption prognosis (middle) and hourly wind power prognosis for Denmark (bottom) for the period 1 January 2013–29 March 2016. The vertical dashed lines mark the beginning of the 91-day period for selecting λ's (ridge regression, lasso, elastic net) and the beginning of the 728-day long out-of-sample test period.

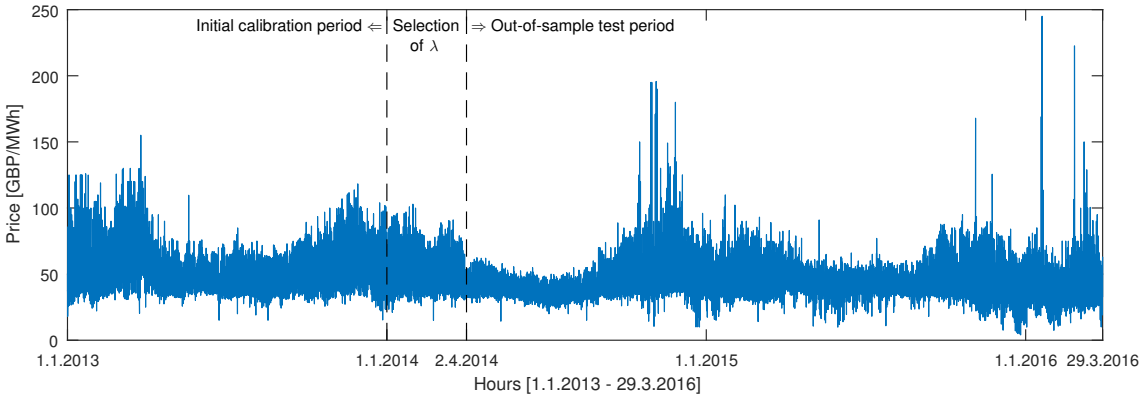

Figure 3. U.K. power market hourly system prices (Nord Pool's N2EX market) for the period 1 January 2013–29 March 2016. The vertical dashed lines mark the beginning of the 91-day period for selecting λ's (ridge regression, lasso, elastic net) and the beginning of the 728-day long out-of-sample test period.

3. Methodology

It should be noted that although we use here the terms short-term, spot and day-ahead interchangeably, the former two do not necessarily refer to the day-ahead market. Short-term EPF generally involves predicting 24 hourly (or 48 half-hourly) prices in the day-ahead market, cleared

typically at noon on the day before delivery, i.e., 12–36 h before delivery, the adjustment markets, cleared a few hours before delivery, and the balancing or real-time markets, cleared minutes before delivery [32]. The spot market, especially in the literature on European electricity markets, is often used as a synonym of the day-ahead market. However, in the U.S., the spot market is another name for the real-time market, while the day-ahead market is called the forward market [20,33]. Furthermore, some markets in Europe nowadays admit continuous trading for individual load periods up to a few hours before delivery. With the shifting of volume from the day-ahead to intra-day markets, also in Europe, the term spot is more and more often being used to refer to the real-time markets [1].

Throughout this article, we denote by $P_{d,h}$ the electricity price in the day-ahead market for day d and hour h. Like many studies in the EPF literature [1], we use the logarithmic transform to make the price series more symmetric (see Figure 4) and compare with the top panels in Figures 1–3. We can do this since all considered datasets are positive-valued. However, this is not a very restrictive property. If datasets with zero or negative values were considered, we could work with non-transformed prices. Furthermore, we center the log-prices by subtracting their in-sample mean prior to parameter estimation. We do this independently for each hour $h = 1, ..., 24$:

$$p_{d,h} = \log(P_{d,h}) - \frac{1}{T}\sum_{t=1}^{T}\log(P_{t,h}), \tag{1}$$

where T is the number of days in the calibration window; hence, the missing intercept ($\beta_{h,0} \equiv 0$) in our autoregressive models; for model parameterizations, see Sections 3.2–3.4.

For all three markets, the day-ahead forecasts of the hourly electricity price are determined within a rolling window scheme, using a 365-day calibration window. First, all considered models are calibrated to data from the initial calibration period (i.e., 1 January 2011–31 December 2011 for GEFCom2014 and 1 January 2013–31 December 2013 for Nord Pool and the U.K.), and forecasts for all 24 h of the next day (1 January) are determined. Then, the window is rolled forward by one day; the models are re-estimated, and forecasts for all 24 h of 2 January are computed. This procedure is repeated until the predictions for the 24 h of the last day in the sample (14 December 2013 for GEFCom2014 and 29 March 2016 for Nord Pool and the U.K.) are made.

For models requiring calibration of the regularization parameter (i.e., λ), we use a setup commonly considered in the machine learning literature. Namely, we divide our datasets into estimation (365 days), validation (91 days or 13 full weeks) and test periods (623 days for GEFCom2014, 728 days for Nord Pool and the U.K.; respectively 89 and 104 full weeks). For each of the five models—ridge regression, lasso and elastic nets with $\alpha = 0.25$, 0.50 and 0.75—34 different 'sub-models' with 34 values of λ spanning the regularization parameter space (see Sections 3.4.3 and 3.4.4 for details) are estimated in the 91-day validation period directly following the last day of the initial calibration period; see Figures 1–3. For all hours of the day, only one value of λ is chosen for each of the five models: the one that yields the smallest $\overline{\text{WMAE}}$ error during this 91-day period; for error definitions, see Section 4.1. This value of λ is later used for computing day-ahead price forecasts in the whole out-of-sample test period. To ensure that all models are evaluated using the same data, predictions of all models are compared only in the out-of-sample test periods: 1 April 2012–14 December 2013 (623 days) for GEFCom2014 and 2 April 2014–29 March 2016 (728 days) for Nord Pool and the U.K. Obviously, such a simple procedure for the selection of the regularization parameter may not be optimal. Generally, better performance is to be expected from shrinkage models when λ is recalibrated at every time step. Such an approach has been recently taken by Ziel [18], who used the Bayesian information criterion to select one out of 50 values of λ for every day and every hour in the 969-day-long out-of-sample test period. The downside of such an approach is, however, the increased computational time.

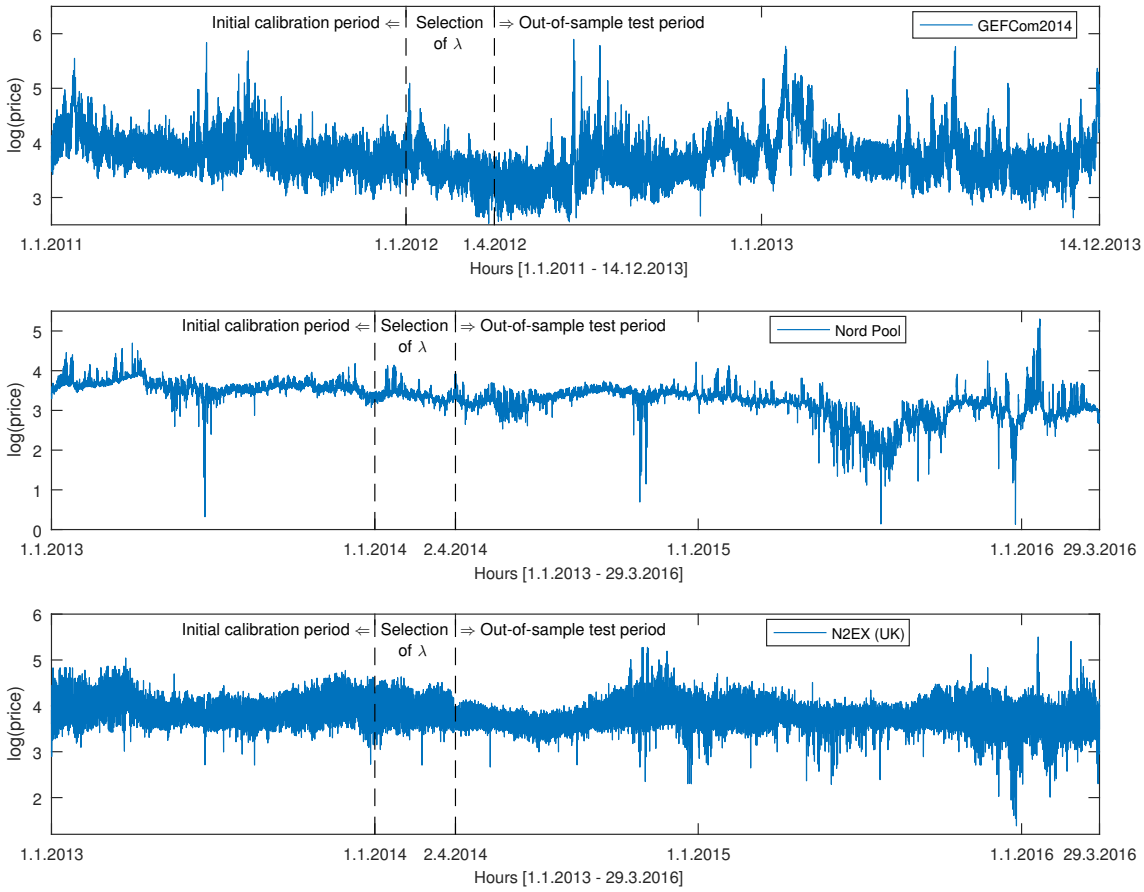

Figure 4. Global Energy Forecasting Competition 2014 (GEFCom2014) (top), Nord Pool (middle) and N2EX (U.K.; bottom) hourly log prices. As expected, the logarithmic transform makes the price series more symmetric. The vertical dashed lines mark the beginning of the 91-day period for selecting λ's (ridge regression, lasso, elastic nets) and the beginning of the out-of-sample test periods. Each day, the 365-day-long calibration window is rolled forward by 24 h; the models are re-estimated; and price forecasts for the 24 h of the next day are computed.

Our choice of the model classes is guided by the existing literature on short-term EPF. Like in [12,18,25–27,30], the modeling is implemented separately across the hours, leading to 24 sets of parameters for each day the forecasting exercise is performed. As Ziel [18] notes, when we compare the forecasting performance of relatively simple models implemented separately across the hours and jointly for all hours (like in [9,34–36]), the latter generally performs better for the first half of the day, whereas the former are better in the second half of the day. At the same time, models implemented separately across the hours offer more flexibility by allowing for time-varying cross-hour dependency in a straightforward manner. Hence, our choice of the modeling framework.

In the remainder of this section, we first define the benchmarks: a simple similar-day technique and a collection of parsimonious autoregressive models. Since the latter are usually built on some prior knowledge of experts, like in [18], we refer to them as expert models. Then, we move on to describe the selection and shrinkage procedures used in this study.

3.1. The Naive Benchmark

The first benchmark, most likely introduced to the EPF literature in [34] and dubbed the naive method, belongs to the class of similar-day techniques (for a taxonomy of EPF approaches, see, e.g., [1]). It proceeds as follows: the electricity price forecast for hour h on Monday is set equal to the price for the same hour on Monday of the previous week, and the same rule applies for Saturdays

and Sundays; the electricity price forecast for hour h on Tuesday is set equal to the price for the same hour on Monday, and the same rule applies for Wednesdays, Thursdays and Fridays. As was argued in [34,35], forecasting procedures that are not calibrated carefully fail to outperform the naive method surprisingly often. We denote this benchmark by **Naive**.

3.2. Autoregressive Expert Benchmarks

The second benchmark is a parsimonious autoregressive structure originally proposed by Misiorek et al. [19] and later used in a number of EPF studies [18,20,21,23–27]. Within this model, the centered log-price on day d and hour h, i.e., $p_{d,h}$, is given by the following formula:

$$\begin{aligned} p_{d,h} = {} & \beta_{h,1} p_{d-1,h} + \beta_{h,2} p_{d-2,h} + \beta_{h,3} p_{d-7,h} + \beta_{h,4} p^{min}_{d-1} + \beta_{h,5} z_{d,h} \\ & + \beta_{h,6} D_{Sat} + \beta_{h,7} D_{Sun} + \beta_{h,8} D_{Mon} + \varepsilon_{d,h}, \end{aligned} \tag{2}$$

where the lagged log-prices $p_{d-1,h}$, $p_{d-2,h}$ and $p_{d-7,h}$ account for the autoregressive effects of the previous days (the same hour yesterday, two days ago and one week ago), while $p^{min}_{d-1} \equiv \min_{h=1,\ldots,24}\{p_{d-1,h}\}$ is the minimum of the previous day's 24 hourly log-prices. The exogenous variable $z_{d,h}$ refers to the logarithm of hourly system load or Nordic consumption for day d and hour h (actually, to forecasts made a day before, see Section 2). The three dummy variables—D_{Sat}, D_{Sun} and D_{Mon}—account for the weekly seasonality. Finally, the $\varepsilon_{d,h}$'s are assumed to be independent and identically distributed (i.i.d.) normal variables. We denote this autoregressive benchmark by **ARX1** to reflect the fact that the load (or consumption) forecast is used as the exogenous variable in Equation (2). The corresponding model with $\beta_{h,5} \equiv 0$, i.e., with no exogenous variable, is denoted by **AR1**. The **ARX1** and **AR1** models, as well as all autoregressive structures considered in Sections 3.2 and 3.3, are estimated in this study with least squares (LS), using MATLAB's regress.m function.

In what follows, we also consider two variants of Equation (2) that treat holidays as special days:

$$\begin{aligned} p_{d,h} = {} & \beta_{h,1} p_{d-1,h} + \beta_{h,2} p_{d-2,h} + \beta_{h,3} p_{d-7,h} + \beta_{h,4} p^{min}_{d-1} + \beta_{h,5} z_{d,h} \\ & + \beta_{h,6} D_{Sat} + \beta_{h,7} D_{Sun} + \beta_{h,8} D_{Mon} + \beta_{h,9} D_{Hol} + \varepsilon_{d,h}, \end{aligned} \tag{3}$$

and that additionally utilize the fact that prices for early morning hours depend more on the previous day's price at midnight, i.e., $p_{d-1,24}$, than on the price for the same hour, as recently noted in [18,29]:

$$\begin{aligned} p_{d,h} = {} & \beta_{h,1} p_{d-1,h} + \beta_{h,2} p_{d-2,h} + \beta_{h,3} p_{d-7,h} + \beta_{h,4} p^{min}_{d-1} + \beta_{h,5} z_{d,h} \\ & + \beta_{h,6} D_{Sat} + \beta_{h,7} D_{Sun} + \beta_{h,8} D_{Mon} + \beta_{h,9} D_{Hol} + \beta_{h,10} p_{d-1,24} + \varepsilon_{d,h}. \end{aligned} \tag{4}$$

We denote Models (3) and (4) by **ARX1h** and **ARX1hm**, respectively. Similarly, corresponding models with $\beta_{h,5} \equiv 0$ are denoted by **AR1h** and **AR1hm**. Note, that when forecasting the electricity price for the last load period of the day, i.e., $p_{d,24}$, models with suffix **hm** reduce to models with suffix **h** (this is true for all models considered in Section 3.2).

In Equations (3) and (4), D_{Hol} is a dummy variable for holidays. The holidays were identified using the Time and Date AS (www.timeanddate.com/holidays) web page: U.S. federal holidays (for GEFCom2014), national holidays in Norway (for Nord Pool) and public holidays, bank holidays and major observances in the U.K. (option 'Holidays and some observances').

The third benchmark is an extension of the **ARX1** model, which takes into account the experience gained during the GEFCom2014 competition that it may be beneficial to use different model structures for different days of the week, not only different parameter sets [29]. Hence, the multi-day ARX model (denoted later in the text by **mARX1**) is given by the following formula:

$$\begin{aligned} p_{d,h} = {} & \left(\sum_{i \in I} \beta_{h,1,i} D_i \right) p_{d-1,h} + \beta_{h,2} p_{d-2,h} + \beta_{h,3} p_{d-7,h} + \beta_{h,4} p^{min}_{d-1} + \beta_{h,5} z_{d,h} \\ & + \beta_{h,6} D_{Sat} + \beta_{h,7} D_{Sun} + \beta_{h,8} D_{Mon} + \beta_{h,11} D_{Mon} p_{d-3,h} + \varepsilon_{d,h}, \end{aligned} \tag{5}$$

where $I \equiv \{0, Sat, Sun, Mon\}$, $D_0 \equiv 1$ and the term $D_{Mon}p_{d-3,h}$ accounts for the autoregressive effect of Friday's prices on the prices for the same hour on Monday. Note that to some extent, this structure resembles periodic autoregressive moving average (PARMA) models, which have seen limited use in EPF [37,38]. Like for the **ARX1** model, also for **mARX1**, we consider two variants:

- **mARX1h**, which treats holidays as special days, i.e., with the $\beta_{h,9}D_{Hol}$ term in Equation (5),
- and **mARX1hm**, which additionally implements the dependence on the previous day's price at midnight, i.e., with the $\beta_{h,9}D_{Hol}$ and $\beta_{h,10}p_{d-1,24}$ terms in Equation (5).

The corresponding price only models, i.e., with $\beta_{h,5} \equiv 0$, are denoted by **mAR1**, **mAR1h** and **mAR1hm**.

Misiorek et al. [19] noted that the minimum of the previous day's 24 hourly prices was the best link between today's prices and those from the entire previous day. Their analysis, however, was limited to one small dataset (California CalPXprices, 3–9 April 2000) and only one simple function at a time (maximum, minimum, mean or median of the previous day's prices). To check if using more than one function leads to a better forecasting performance, we introduce a benchmark, which is an extension of the **ARX1** model that takes into account not only the minimum (p^{min}_{d-1}), but also the maximum (p^{max}_{d-1}) and the mean (p^{avg}_{d-1}) of the previous day's 24 hourly prices. Additionally, we include a second exogenous variable $(y_{d,h})$, which is taken as either the logarithm of the day-ahead zonal load forecast (GEFCom2014) or of the Danish wind power prognosis. The resulting **ARX2** model is given by the following formula:

$$\begin{aligned} p_{d,h} = {} & \beta_{h,1}p_{d-1,h} + \beta_{h,2}p_{d-2,h} + \beta_{h,3}p_{d-7,h} + \beta_{h,4}p^{min}_{d-1} + \beta_{h,5}z_{d,h} \\ & + \beta_{h,6}D_{Sat} + \beta_{h,7}D_{Sun} + \beta_{h,8}D_{Mon} \\ & + \beta_{h,11}p^{max}_{d-1} + \beta_{h,12}p^{avg}_{d-1} + \beta_{h,13}y_{d,h} + \varepsilon_{d,h}. \end{aligned} \tag{6}$$

Like for the **ARX1** and **mARX1** models, also for **ARX2**, we consider two variants:

- **ARX2h** with the $\beta_{h,9}D_{Hol}$ term in Equation (6),
- and **ARX2hm** with the $\beta_{h,9}D_{Hol}$ and $\beta_{h,10}p_{d-1,24}$ terms in Equation (6).

The corresponding price only models, i.e., with $\beta_{h,5}, \beta_{h,13} \equiv 0$, are denoted by **AR2**, **AR2h** and **AR2hm**.

3.3. Full Autoregressive Model

Finally, we define a much richer autoregressive model that includes as special cases all expert models discussed in Section 3.2 and call it the **full ARX** or **fARX** model. We consider all regressors that, in our opinion, posses a non-negligible predictive power. The **fARX** model is similar in spirit to the general autoregressive model defined by Equation (2) in [18]. However, there are some important differences between them. On one hand, **fARX** includes exogenous variables and a much richer seasonal structure. On the other, it does not look that far into the past and concentrates only on days $d-1$, $d-2$, $d-3$ and $d-7$. The **fARX** model is given by the following formula:

$$\begin{aligned} p_{d,h} = {} & \sum_{i=1}^{24} \left(\beta_{h,i}p_{d-1,i} + \beta_{h,i+24}p_{d-2,i} + \beta_{h,i+48}p_{d-3,i}\right) + \beta_{h,73}p_{d-7,h} \\ & + \sum_{j=1}^{3} \left(\beta_{h,j+73}p^{min}_{d-j} + \beta_{h,j+76}p^{max}_{d-j} + \beta_{h,j+79}p^{avg}_{d-j}\right) \\ & + \beta_{h,83}z_{d,h} + \beta_{h,84}z_{d-1,h} + \beta_{h,85}z_{d-7,h} + \beta_{h,86}y_{d,h} \\ & + \sum_{k=1}^{7} \beta_{h,86+k}D_k + \sum_{k=1}^{7} \beta_{h,93+k}D_k z_{d,h} + \sum_{k=1}^{7} \beta_{h,100+k}D_k p_{d-1,h} + \varepsilon_{d,h}, \end{aligned} \tag{7}$$

where $D_1 \equiv D_{Sat}, D_2 \equiv D_{Sun}, ..., D_7 \equiv D_{Fri}$ are dummies for the seven days of the week (we treat holidays as the eighth day of the week, hence $D_1 = ... = D_7 = 0$ for holidays). The price only variant,

fAR, is obtained by setting to zero all coefficients of the terms involving exogenous variables, i.e., $\beta_{h,i} \equiv 0$, for $i = 83, ..., 86, 94, ..., 100$.

Although we fit the **fARX** model to power market data and evaluate its forecasting performance, the main reason for including it in this study is to use it as the baseline model for the selection and shrinkage procedures discussed in Section 3.4. For this purpose, let us write the **fARX** model in a more compact form:

$$p_{d,h} = \sum_{i=1}^{n} \beta_{h,i} X_{d,h,i} + \varepsilon_{d,h}, \tag{8}$$

where $X_{d,h,i}$'s are the $n = 107$ regressors in Equation (7) and $\beta_{h,i}$'s are their coefficients.

3.4. Selection and Shrinkage Procedures

All autoregressive models considered in Sections 3.2 and 3.3 are estimated in this study with least squares (LS). However, there are many alternatives to using LS in multi-parameter models, in particular [39]:

- variable or subset selection, which involves identifying a subset of predictors that we believe to be influential, then fitting a model using LS on the reduced set of variables,
- shrinkage (also known as regularization), which fits the full model with all predictors using an algorithm that shrinks the estimated coefficients towards zero, which can significantly reduce their variance.

Depending on what type of shrinkage is performed, some of the coefficients may be shrunk to zero itself. As such, some shrinkage methods, like the lasso, de facto perform variable selection. It should be noted, however, that variable selection (or model sparsity) is beneficial for interpretability and faster simulation of model trajectories; for reducing the forecasting errors, only the shrinkage property is required.

3.4.1. Single-Step Elimination of Insignificant Predictors

This subset selection procedure is a simple alternative to stepwise regression discussed in Section 3.4.2 and has been used, for instance, in [5]. The idea is to fit the full regression model, in our case **fARX**, then in a single step, set to zero all statistically insignificant coefficients. We use MATLAB's regress.m function with the commonly-used 5% significance level. Setting to zero all coefficients in Equation (7) whose 95% confidence intervals (CI) include zero yields the **ssARX** model for a particular day and hour (the **ssAR** model is obtained analogously from **fAR**; see Section 3.3). This procedure can be conducted by imposing some additional constraints, for instance, leaving in the model all coefficients of the basic **ARX1** (or **AR1**) benchmark. This yields the **ssARX1** and **ssAR1** models. Of course, the most commonly-used significance level of 5% may not be optimal. We have additionally checked the performance of 90% and 97.5% CI. It turns out that the overall ranking of the **ssAR**-type models does not change much. However, **ssARX** and **ssAR** perform slightly better for the 90% CI, while **ssARX1** and **ssAR1** either for the 95% or the 97.5% CI.

3.4.2. Stepwise Regression

Although very fast, the single-step elimination may remove too many explanatory variables at once and lead to a poorly-performing subset of predictors. On the other hand, selecting the best subset from among all 2^n subsets of the n predictors is not computationally feasible for large n. Even if doable, it may lead to overfitting. For these reasons, stepwise methods, which explore a far more restricted set of models, are attractive alternatives to best subset selection [39]. In the context of EPF, they have been used, for instance, in [12,13,40].

There are two basic procedures: forward selection and backward elimination. Forward stepwise selection begins with a model containing no predictors and then iteratively adds variables to the model. At each step, the variable that gives the greatest additional improvement to the fit is added to the

model, and the procedure continues until all important predictors are in the model. We use MATLAB's stepwisefit.m function, which computes the p-value of an F-statistic at each time step to test models with and without a potential term. If a variable is not currently in the model, the null hypothesis is that it would have a zero coefficient if added to the model. If there is sufficient evidence to reject the null hypothesis, that variable may be added to the model (we use stepwisefit's default 5% significance level for adding variables; naturally, this could be further fine tuned as for the single-step elimination procedures). In a given step, the function adds the variable with the smallest p-value. We denote the resulting models by **fsARX** and **fsAR**.

Backward stepwise elimination (or selection) begins with the full model containing all n variables, i.e., **fARX** or **fAR**, and then iteratively removes the least useful predictor, one at a time. MATLAB's stepwisefit.m function computes the null hypothesis that a given variable has a zero coefficient. If there is insufficient evidence to reject the null hypothesis, the variable may be removed from the model (we use stepwisefit's default 10% significance level for removing variables). In a given step, the function removes the variable with the largest p-value. We denote the resulting models by **bsARX** and **bsAR**.

3.4.3. Ridge Regression

Ridge regression is a regularization method introduced in statistics by Hoerl and Kennard [41]. To our best knowledge, apart from a limited study of Barnes and Balda [16] in the context of evaluating the profitability of battery storage, the method has not been used for EPF. Ridge regression is very similar to least squares, except that the β_i's in (8) are not estimated by minimizing the residual sum of squares (RSS), but by RSS penalized by a quadratic shrinkage factor:

$$\hat{\boldsymbol{\beta}}^{ridge} = \underset{\beta_{h,i}}{\operatorname{argmin}} \left\{ \mathrm{RSS} + \lambda \sum_{i=1}^{n} \beta_{h,i}^2 \right\} \equiv \underset{\beta_{h,i}}{\operatorname{argmin}} \left\{ \sum_{d,h \in T} \left(p_{d,h} - \sum_{i=1}^{n} \beta_{h,i} X_{d,h,i} \right)^2 + \lambda \sum_{i=1}^{n} \beta_{h,i}^2 \right\}, \quad (9)$$

where T represents the calibration period and $\lambda \geq 0$ is a tuning or regularization parameter, to be determined separately. Note that for $\lambda = 0$, we get the standard LS estimator; for $\lambda \to \infty$, all $\beta_{h,i}$'s tend to zero; while for intermediate values of λ, we are balancing two ideas: minimizing the RSS and shrinking the coefficients towards zero (and each other).

Ridge regression produces a different set of coefficient estimates for each value of λ; hence, selecting a good value for λ is critical. Cross-validation provides a simple way to tackle this problem [39]. We choose a grid of λ values (here: 34 equally-spaced values spanning the range from 1–100; if $\lambda \in \{94, 97, 100\}$ was selected, we additionally checked another set of 34 equally-spaced values spanning the range from 101–200) and using MATLAB's ridge.m function (we scale the regressors) compute the prediction errors for each value of the tuning parameter in the 91-day validation period; see Section 2. We then select λ for which the $\overline{\mathrm{WMAE}}$ error (for the definition, see Section 4.1) is the smallest and use it for computing day-ahead price forecasts in the whole out-of-sample test period. The resulting model is denoted in the text by **RidgeX** or **Ridge** when the baseline model is **fAR**.

3.4.4. Lasso and Elastic Nets

Ridge regression has one unwanted feature when it comes to interpretation and model identification. Unlike stepwise regression, which will generally select models that involve just a subset of the variables, ridge regression will include all n predictors in the final model [39]. The quadratic shrinkage factor in Equation (9) will shrink all $\beta_{h,i}$'s towards zero, but it will not set any of them exactly to zero. In 1996, Tibshirani [42] proposed the least absolute shrinkage and selection operator (i.e., lasso or LASSO) that overcomes this disadvantage. It is the only shrinkage procedure that has been applied in EPF to a larger extent, however only in the last two years [7,9,18,25,43].

The lasso is a shrinkage method just like ridge regression. However, it uses a linear penalty factor instead of a quadratic one:

$$\hat{\boldsymbol{\beta}}^{lasso} = \underset{\beta_{h,i}}{\operatorname{argmin}} \left\{ \mathrm{RSS} + \lambda \sum_{i=1}^{n} |\beta_{h,i}| \right\}. \tag{10}$$

This subtle change makes the solutions nonlinear in $p_{d,h}$, and there is no closed form expression as in the case of ridge regression. Because of the nature of the shrinkage factor in Equation (10), making λ sufficiently large will cause some of the coefficients to be exactly zero [44]. Thus, the lasso de facto performs variable selection, just like the methods discussed in Sections 3.4.1 and 3.4.2. As in ridge regression, selecting a good value of λ for the lasso is critical. Here, we use MATLAB's lasso.m function and a grid of exponentially-decreasing λ's (the largest just sufficient to produce all $\beta_i = 0$; the function also automatically scales the regressors). We then select λ for which the $\overline{\mathrm{WMAE}}$ error (for the definition, see Section 4.1) in the 91-day validation period is the smallest. The resulting model is denoted in the text by **LassoX**, or **Lasso** when the baseline model is **fAR**.

The lasso does not handle highly-correlated variables very well. The coefficient paths tend to be erratic and can sometimes show wild behavior [44]. This is not a critical issue for forecasting, but for interpretation and model identification, this has more serious consequences. In 2005, Zou and Hastie [45] proposed the elastic net, a new regularization and variable selection method that can be seen as an extension of ridge regression and the lasso. It often outperforms the lasso, while exhibiting a similar sparsity of representation. The elastic net uses a mixture of linear and quadratic penalty factors:

$$\hat{\boldsymbol{\beta}}^{EN} = \underset{\beta_{h,i}}{\operatorname{argmin}} \left\{ \mathrm{RSS} + \lambda \left(\frac{1-\alpha}{2} \sum_{i=1}^{n} \beta_{h,i}^2 + \alpha \sum_{i=1}^{n} |\beta_{h,i}| \right) \right\}, \tag{11}$$

where $\alpha \in [0, 1]$. When $\alpha = 1$, the elastic net reduces to the lasso, and with $\alpha = 0$, it becomes ridge regression. The $\frac{1}{2}$ in the quadratic part of the elastic net penalty in Equation (11) leads to a more efficient and intuitive soft-thresholding operator in the optimization; the original formulation in [45] did not include the $\frac{1}{2}$ scaling. Note also that every elastic net problem can be rewritten as a lasso problem on augmented data. Hence, for fixed λ and α, the computational difficulty of the elastic net solution is similar to the lasso problem [44].

Compared to the lasso and ridge regression, the elastic net has an additional mixing parameter that has to be determined. It can be set on subjective grounds, as we do here, or optimized within a cross-validation scheme. We use MATLAB's lasso.m function (with a grid of exponentially-decreasing λ's; the function also automatically scales the regressors) and three values of the mixing parameter, $\alpha = 0.25$, 0.50 and 0.75. This yields six elastic net models:

- **EN25X**, **EN50X** and **EN75X** when the baseline model is **fARX**,
- and **EN25**, **EN50** and **EN75** when the baseline model is **fAR**,

that span the space between ridge regression (**RidgeX**, **Ridge**) and lasso models (**LassoX**, **Lasso**).

4. Empirical Results

We now present day-ahead forecasting results for the three considered datasets: GEFCom2014 hourly locational marginal prices, Nord Pool hourly system prices and U.K. hourly system prices. We use long, two-year out-of-sample test periods to make sure the obtained results are reliable (for the GEFCom2014 dataset, the test period is shorter: 623 days; see Figure 1). Recall from Section 2 that the models are re-estimated on a daily basis. Price forecasts $\hat{P}_{d+1,1}, \ldots, \hat{P}_{d+1,24}$ for all 24 h of the next day are determined at the same point in time, and the 365-day calibration window is rolled forward by one day.

4.1. Performance Evaluation in Terms of WMAE

Following [21,24,30,35], we compare the models in terms of the weekly-weighted mean absolute error (WMAE) loss function, which is a robust measure similar to MAPE, but with the absolute error normalized by the mean weekly price to avoid the adverse effect of negative and close to zero electricity spot prices. We evaluate the forecasting performance using weekly time intervals, each with $24 \times 7 = 168$ hourly observations. For each week $w = 1, ..., w_{max}$ in the out-of-sample test period, we calculate the error for each model as:

$$\text{WMAE}_w = \frac{1}{\bar{P}_{168}}\text{MAE}_w = \frac{1}{168 \cdot \bar{P}_{168}} \sum_{d=Mon}^{Sun} \sum_{h=1}^{24} \left|P_{d,h} - \hat{P}_{d,h}\right|, \tag{12}$$

where $P_{d,h}$ is the actual price for hour h (not the centered log-price $p_{d,h}$), $\hat{P}_{d,h}$ is the model predicted price for that hour, $\bar{P}_{168} = \frac{1}{168}\sum_{d=Mon}^{Sun}\sum_{h=1}^{24} P_{d,h}$ is the mean price for a given week and $w_{max} = 89$ for GEFCom2014 and 104 for Nord Pool and the U.K. Next, we aggregate these errors into one mean value over all weeks in the out-of-sample test period:

$$\overline{\text{WMAE}} = \frac{1}{w_{max}} \sum_{w=1}^{w_{max}} \text{WMAE}_w. \tag{13}$$

Note that we also analyzed the forecasts using the weekly root mean square error (see [1] (Section 3.3)), but the results were qualitatively the same and are omitted here due to space limitations.

In Table 1, we report $\overline{\text{WMAE}}$ errors for the three considered datasets and the 20 model types. We use the GEFCom2014 dataset twice: once we fit ARX-type models to the complete dataset with exogenous variables (system and zonal load; left part of the table) and once we fit AR-type models to the dataset without them (right part of the table). This allows us to compute the decrease in WMAE when exogenous variables are added to the model (the last column in Table 1). Several important conclusions can be drawn:

- All models beat the **Naive** benchmark and, except for the **fAR** model and the U.K. data, by a large margin. In particular, the improvement from using elastic nets can be as much as 5%! This indicates that they all are highly efficient forecasting tools.
- When we exclude single-step elimination without constraints (**ssAR/X**) and backward selection (**bsAR/X**) models, the selection and shrinkage methods generally outperform the expert benchmarks. In particular, the elastic net model with $\alpha = 0.75$ (i.e., closer in terms of α to the lasso than to ridge regression) beats every expert model, except **mAR1hm** for the U.K. data, where it is second best.
- The latter comment leads us to the next conclusion that adding the price for the last load period of the day, $p_{d-1,24}$, to the expert models improves their performance greatly. This fact has been recognized in the EPF literature only very recently [18,25,29] and apparently requires more attention. To see this, compare the models with suffix **m** to those without it. In particular, **mAR1hm** is the overall best performing model for the U.K. dataset and **ARX2hm** is the third best model for the Nord Pool dataset.
- Somewhat surprisingly, the full ARX model performs poorly. For the U.K. dataset, it is nearly as bad as the **Naive** benchmark. In all four cases (three datasets + GEFCom2014 without exogenous variables), it is worse than the overall best model and the best performing elastic net (**EN75/X**) by at least 1.4%. Given that a 1% improvement in MAPE translates into savings of ca. $1.5 million per year for a typical medium-size utility [2,3], this observation is of high practical value. Yet, from a statistical perspective, this finding is not that surprising. The **fARX** model has 107 parameters, which have to be calibrated to only 365 observations. Increasing the length of the calibration window should lead to a better performance of the full model.

Table 1. Mean values of the weekly-weighted mean absolute errors, i.e., $\overline{\text{WMAE}}$ defined by Equation (13), over all 89 weeks of the GEFCom2014 or all 104 weeks of the Nord Pool and U.K. out-of-sample test periods. $\overline{\text{WMAE}}$ errors are reported in percent, with standard deviation in parentheses. A heat map is used to indicate better (→ green) and worse (→ red) performing models. $\overline{\text{WMAE}}$ errors for the best performing model for each dataset are emphasized in bold. The last column presents the decrease in $\overline{\text{WMAE}}$ when exogenous variables are added to the model (**AR** → **ARX**; for the GEFCom2014 dataset). The bottom rows compare the performance across model classes.

	ARX-type			**AR-type**		**AR - ARX**
	GEFCom	**Nord Pool**		**GEFCom**	**N2EX (UK)**	**GEFCom**
Naive	14.708 (0.975)	11.141 (0.778)	**Naive**	14.708 (0.975)	9.767 (0.310)	0.000
Expert benchmarks						
ARX1	11.069 (0.639)	9.739 (0.614)	**AR1**	11.183 (0.701)	8.384 (0.253)	0.114
ARX1h	11.072 (0.639)	9.693 (0.616)	**AR1h**	11.181 (0.704)	8.389 (0.253)	0.109
ARX1hm	10.976 (0.617)	8.673 (0.516)	**AR1hm**	11.062 (0.657)	8.229 (0.247)	0.086
mARX1	11.102 (0.621)	9.482 (0.601)	**mAR1**	11.320 (0.696)	8.258 (0.253)	0.218
mARX1h	11.105 (0.622)	9.461 (0.602)	**mAR1h**	11.322 (0.699)	8.270 (0.254)	0.218
mARX1hm	10.974 (0.598)	8.461 (0.518)	**mAR1hm**	11.168 (0.644)	**8.098** (0.246)	0.195
ARX2	10.742 (0.575)	8.878 (0.546)	**AR2**	11.331 (0.700)	8.290 (0.253)	0.589
ARX2h	10.739 (0.575)	8.826 (0.546)	**AR2h**	11.333 (0.704)	8.288 (0.253)	0.594
ARX2hm	10.625 (0.565)	8.206 (0.485)	**AR2hm**	11.070 (0.656)	8.237 (0.249)	0.444
Full ARX model						
fARX	10.911 (0.507)	10.131 (0.708)	**fAR**	12.279 (0.602)	9.724 (0.334)	1.368
Selection and shrinkage methods						
ssARX	10.669 (0.577)	8.861 (0.537)	**ssAR**	12.061 (0.644)	9.344 (0.270)	1.393
ssARX1	9.894 (0.548)	8.409 (0.507)	**ssAR1**	11.343 (0.641)	8.395 (0.261)	1.449
fsARX	9.876 (0.502)	8.130 (0.502)	**fsAR**	11.193 (0.592)	8.563 (0.272)	1.317
bsARX	10.449 (0.502)	9.421 (0.599)	**bsAR**	11.968 (0.582)	9.252 (0.301)	1.519
RidgeX	9.777 (0.544)	8.972 (0.479)	**Ridge**	10.775 (0.653)	8.237 (0.260)	0.998
LassoX	9.476 (0.516)	8.419 (0.503)	**Lasso**	10.722 (0.609)	8.125 (0.253)	1.246
EN75X	9.475 (0.517)	**8.056** (0.489)	**EN75**	10.708 (0.610)	8.124 (0.253)	1.233
EN50X	**9.473** (0.518)	8.287 (0.496)	**EN50**	10.688 (0.611)	8.121 (0.253)	1.215
EN25X	9.474 (0.522)	8.529 (0.503)	**EN25**	**10.650** (0.613)	8.113 (0.253)	1.176
Comparisons						
Expert - Best	1.152	0.150	**Expert - Best**	0.412	0.000	
fARX - Best	1.438	2.075	**fAR - Best**	1.629	1.626	
Naive - Best	5.235	3.086	**Naive - Best**	4.058	1.670	

- Among the selection and shrinkage methods, the lasso and elastic nets tend to outperform single-step elimination (**ssAR/X/1**), stepwise regression (**fsAR/X**, **bsAR/X**) and even ridge regression (**Ridge/X**). Only for the Nord Pool dataset, the **fsARX** forward selection model is better than the lasso and two elastic nets.

4.2. Diebold–Mariano Tests

In order to formally investigate the advantages from using selection and shrinkage methods, we apply the Diebold–Mariano (DM; see [31]) test for significant differences in the forecasting performance. Since predictions for all 24 h of the next day are made at the same time using the same information set, forecast errors for a particular day will typically exhibit high serial correlation. Therefore, like [24,30,46], we conduct the DM tests for each of the 24 load periods separately, using absolute error losses of the model forecast:

$$L(\varepsilon_t) = |\varepsilon_t| = |P_{d,h} - \hat{P}_{d,h}|. \tag{14}$$

For each pair of models and for each hour independently, we calculate the loss differential series:

$$d_t = L(\varepsilon_t^{model_X}) - L(\varepsilon_t^{model_Y}). \tag{15}$$

We perform two one-sided DM tests at the 5% significance level: (i) a test with the null hypothesis H_0:$E(d_t) \leq 0$, i.e., the outperformance of the forecasts of $model_Y$ by those of $model_X$; and (ii) the complementary test with the reverse null H_0^R:$E(d_t) \geq 0$, i.e., the outperformance of the forecasts of $model_X$ by those of $model_Y$. Note that, like in [24,30,46], we assume here forecasts for consecutive days, hence loss differentials are not serially correlated. For the better performing models, this is a generally valid assumption.

In Figures 5 and 6, we summarize the DM results for all test cases (three datasets + GEFCom2014 without exogenous variables). Namely, we sum the number of significant differences in forecasting performance across the 24 h and use a heat map to indicate the number of hours for which the forecasts of a model on the X-axis are significantly better than those of a model on the Y-axis. Two extreme cases—(i) the forecasts of a model on the X-axis are significantly better for all 24 h of the day and (ii) the forecasts of a model on the X-axis are not significantly better for any hour—are indicated by white and black squares, respectively. Naturally, the diagonal (white crosses on black squares) should be ignored, as it concerns the same model on both axes. Columns with many non-black squares (the more green or white the better) indicate that the forecasts of a model on the X-axis are significantly better than the forecasts of many of its competitors. Conversely, rows with many non-black squares mean that the forecasts of a model on the Y-axis are significantly worse than the forecasts of many of its competitors. For instance, for the GEFCom2014 dataset and ARX-type models displayed in the left panel of Figure 5, the white row for the **Naive** benchmark indicates that the forecasts of this simple model are significantly worse than the forecasts of all of its competitors for all 24 h, while the black column for the **Naive** benchmark means that not a single competitor produces significantly worse forecasts than **Naive**, even for a single hour of the day.

The obtained DM-test results support our observations from Section 4.1 on WMAE errors. Again, we can conclude that applying the lasso or one of the elastic nets improves forecasting accuracy. Especially for the GEFCom2014 dataset (both for ARX- and AR-type models), these variable selection schemes lead to models that yield significantly better forecasts than those of the expert models (see the white columns for **Lasso/X**, **EN75/X**, **EN50/X** and **EN25/X** in the left panels of Figures 5 and 6), while their predictions are never outperformed by any of the competitors (see the black rows for these four models). For the Nord Pool and U.K. datasets, the results are not that clear cut, but still there are many more green or white squares in the columns than in the rows corresponding to these four selection schemes.

Again, **EN75/X** stands out as the best performing model overall. For the GEFCom2014 test cases, it always leads to significantly better forecasts than any of the expert benchmarks. For the Nord Pool dataset, its forecasts are significantly better for 10–23 h of the day and significantly worse for at most 2 h (only for models with suffix **m**: **mARX1hm** and **ARX2hm**, 2 h, and **ARX1hm**, 1 h). Finally, for the U.K. dataset, the results are the least convincing. **EN75/X** yields significantly better forecasts for 4–12 h of the day and significantly worse for at most 2 h (only for **mAR**-type models: **mAR1** and **mAR1h**, 2 h, and **mAR1hm**, 1 h).

Now, let us look in detail at the performance for each hour of the day. In Figure 7, we provide a graphical representation of the DM test statistic for four models and all considered datasets. The models include: the best overall **EN75/X** model and three benchmarks (**Naive**, **mARX1hm/mAR1hm** and **fAR/X**). For the GEFCom2014 dataset, **EN75/X** clearly beats all benchmarks across all hours. The situation for the remaining two datasets would be nearly the same if it was not for the early morning hours (Hours 6 and 7 for Nord Pool and Hour 8 for the U.K.), when the expert benchmarks yield significantly better predictions. This is somewhat surprising, since the morning peak comes a bit later in both markets. Perhaps looking at variables selected by the elastic net algorithm will provide more insight.

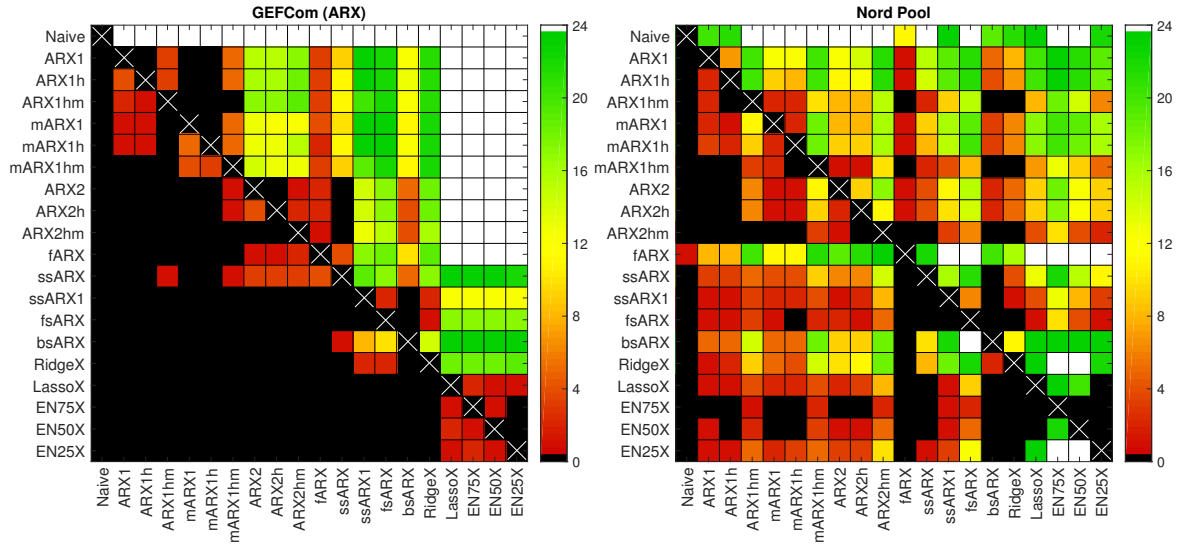

Figure 5. Results for conducted one-sided Diebold–Mariano tests at the 5% level for autoregressive model structures with exogenous variables (ARX)-type models and two datasets: GEFCom2014 (left panel) and Nord Pool (right panel). We sum the number of significant differences in forecasting performance across the 24 h and use a heat map to indicate the number of hours for which the forecasts of a model on the X-axis are significantly better than those of a model on the Y-axis. A white square indicates that forecasts of a model on the X-axis are better for all 24 h, while a black square that they are not better for a single hour.

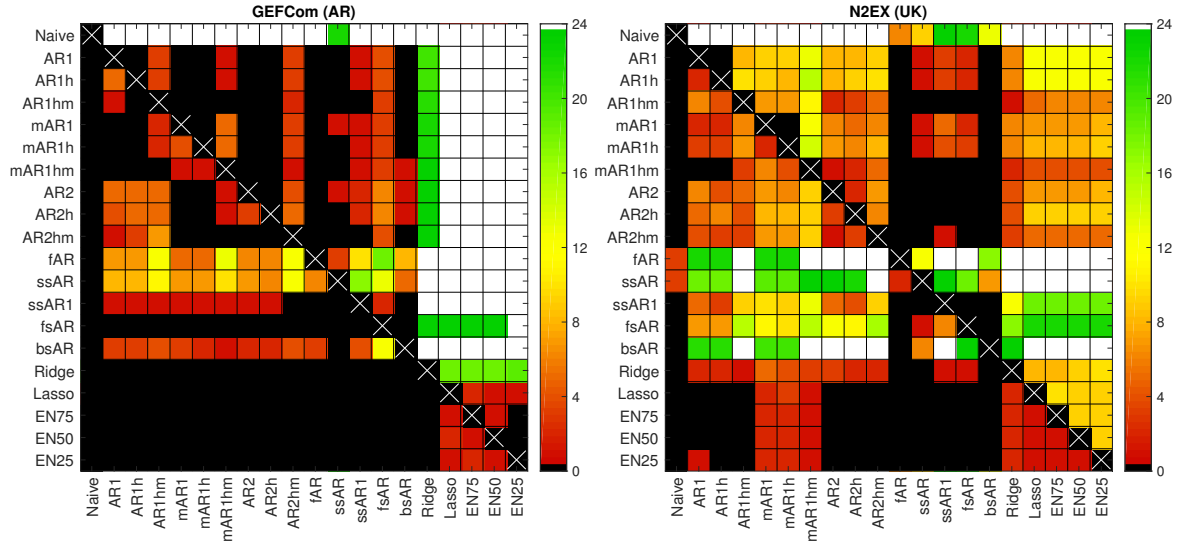

Figure 6. Results for conducted one-sided Diebold–Mariano tests at the 5% level for AR-type models and two datasets: GEFCom2014 (left panel) and N2EX (U.K.; right panel). We sum the number of significant differences in forecasting performance across the 24 h and use a heat map to indicate the number of hours for which the forecasts of a model on the X-axis are significantly better than those of a model on the Y-axis. A white square indicates that forecasts of a model on the X-axis are better for all 24 h, while a black square that they are not better for a single hour.

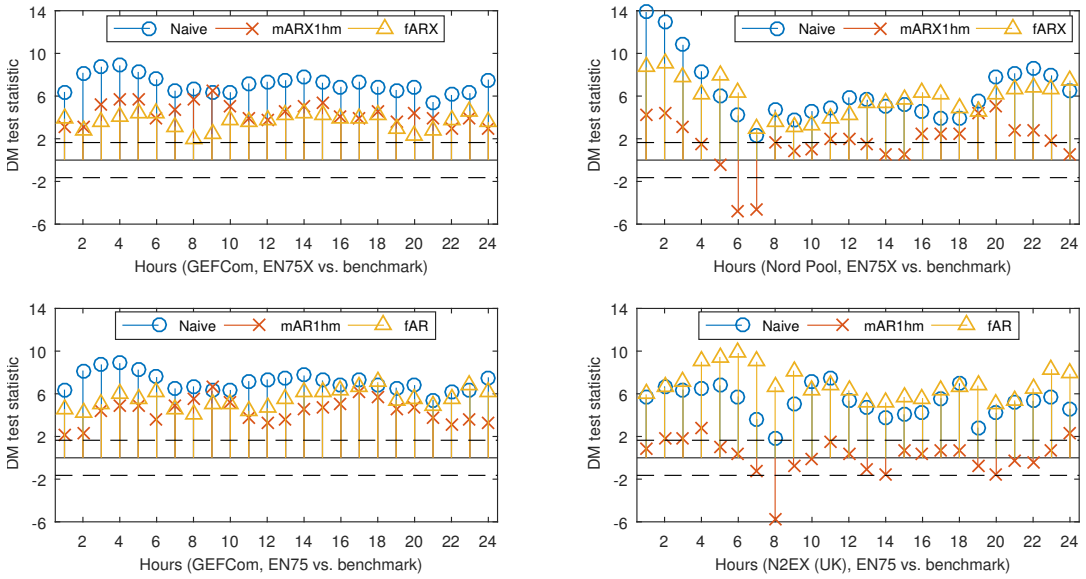

Figure 7. Results for the conducted one-sided Diebold–Mariano tests at the 5% significance level for selected ARX-type models and the GEFCom2014 and Nord Pool datasets (top panels) and selected AR-type models and the GEFCom2014 and N2EX (U.K.) datasets (bottom panels). The tests were conducted separately for each of the 24 h. The figures report the value of the test statistic for each test, as well as two thresholds (dashed lines in the plots). The lower one refers to null hypothesis H_0:$E(d_t) \leq 0$, i.e., the outperformance of the forecasts of **EN75/X** by those of a given benchmark (**Naive**, **mARX1hm/mAR1hm**, **fAR/X**). The upper threshold refers to the complementary test with the reverse null, i.e., H_0:$E(d_t) \geq 0$ or the outperformance of the forecasts of a given benchmark by those of **EN75/X**. Only points lying below (or above) the dashed threshold lines are significant at the 5% level. See also Figures 5 and 6.

4.3. Variable Selection

In Tables 2 and 3, we provide the number of days in the out-of-sample test period for which a given $\beta_{h,i}$ was selected for the best performing elastic net model, i.e., **EN75/X**. The maximum number of days is 623 ($= 7 \times 89$ weeks) for GEFCom2014 and 728 ($= 7 \times 104$ weeks) for Nord Pool and N2EX (U.K.). A heat map is used to indicate more ($\rightarrow$ green) and less ($\rightarrow$ red) commonly-selected $\beta_{h,i}$'s. The $\beta_{h,i}$'s are numbered as in Equation (7). Note that $\beta_{h,83}, ..., \beta_{h,86}, \beta_{h,94}, ..., \beta_{h,100} \equiv 0$ in the **EN75** model; see Table 3. Several interesting conclusions can be drawn:

- There is no single variable that is always used, regardless of the dataset, hour of the day or the day in the out-of-sample test period. The closest to 'perfection' is the day-ahead load forecast for the predicted hour, i.e., $z_{d,h}$ (see Row 83 in Table 2). Surprisingly, this dependence on the load forecast is stronger than the autoregressive effect (see the next bullet point). This may be a hint that the load-price relationship should be given more attention and that functionals of load-related (or other fundamental) variables should be included in EPF models, like in [10].
- As expected, the price 24 h ago, i.e., $p_{d-1,h}$, is an influential variable; see the diagonals in Rows 1–24 in both Tables. However, it is not only the same hour a day earlier, but also the neighboring hours. The diagonal is less visible around mid-day, and for Nord Pool, it almost disappears except for the late night hours. The latter may be to some extent due to the importance of wind in this market and the explanatory power of the day-ahead wind prognosis for the predicted hour.
- As recently observed in [18,29], the price for Hour 24, i.e., $p_{d-1,24}$, is an influential variable. Somewhat surprisingly, sometime between 7–9 a.m. and 9–11 p.m., Hour 22, i.e., $p_{d-1,22}$, becomes more important. What is more surprising, these late night hours are generally more often selected than the same hour a day ago, i.e., $p_{d-1,h}$. These observations require more thorough studies. Nevertheless, our limited results suggest that these late hour variables should be taken into account when constructing expert models.
- Clearly, the least important variables for all markets are the daily average prices over the last three days, i.e., p^{avg}_{d-j} for $j = 1, 2, 3$, which are almost never selected. There are some exceptions, though, for the GEFCom2014 dataset and the **EN75** model; see Table 3. Of the two other aggregated variables, p^{max}_{d-j} is slightly more influential than p^{min}_{d-j}, which contradicts the observations of Misiorek et al. [19] and may suggest its use in expert models instead of the minimum.
- If prices from days $(d-2)$ or $(d-3)$ are ever selected, it is only for hours around midnight the day before (i.e., $p_{d-2,23}$, $p_{d-2,24}$, $p_{d-3,1}$) or similar hours (i.e., the diagonals in Rows 25–48 and 49–72). On the other hand, the same hour one week ago, i.e., $p_{d-7,h}$, has a high explanatory power (see Row 73 for all datasets), which justifies its use in expert models [18–23,30].
- Finally, the weekly dummies (Rows 87–93), the dummy-linked load forecasts (Rows 94–100 in Table 2 only) and the dummy-linked last day's prices (Rows 101–107) are generally selected for the **EN75/X** model. This may be an indication that the weekly seasonality requires better modeling than offered by typically-used expert models.

Table 2. The number of days in the out-of-sample test period for which a given $\beta_{h,i}$ (see Equation (7)) was selected for the **EN75X** model. The maximum number of days is 623 (= 7×89 weeks) for GEFCom2014 and 728 (= 7×104 weeks) for Nord Pool. A heat map is used to indicate more ($\rightarrow$ green) and less ($\rightarrow$ red) commonly-selected $\beta_{h,i}$'s.

GEFCom2014	Hour																						
$\beta_{h,i}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	600	621	314	116	126	522	623	495	135	316	402	498	490	452	489	456	469	297	411	343	312	452	562	545
2	610	623	471	390	191	352	44	21	11	2	102	78	26	52	13	0	0	0	44	30	112	133	53	51
3	38	453	580	465	271	164	206	75	106	69	155	158	14	191	154	267	298	171	76	164	21	0	0	0
4	205	197	390	550	369	184	2	50	10	0	0	0	0	0	0	9	18	0	0	1	0	0	7	5
5	289	239	396	460	621	282	101	180	41	201	293	154	81	17	1	38	0	62	210	0	128	26	6	0
6	307	467	596	618	604	623	566	365	370	184	184	115	284	316	336	337	257	336	138	147	176	51	233	452
7	8	80	205	158	329	197	623	623	595	613	399	264	80	16	40	15	51	216	317	507	286	326	311	130
8	119	85	164	205	233	275	406	623	619	535	422	151	41	58	38	3	90	204	181	103	251	98	19	5
9	87	0	0	0	6	91	43	67	318	314	73	61	61	56	28	131	30	0	0	0	43	0	14	45
10	269	3	3	0	2	9	261	112	100	170	58	15	64	170	251	304	315	296	268	280	266	278	277	177
11	17	16	17	0	0	0	73	64	242	426	501	556	245	139	7	14	7	123	150	87	1	3	11	0
12	12	17	10	0	0	4	0	0	0	0	177	284	263	168	187	162	74	18	0	15	1	0	2	23
13	61	88	0	78	9	0	0	13	0	117	269	305	324	304	158	93	45	54	53	59	101	83	53	62
14	13	64	0	0	0	14	35	106	5	0	18	51	206	239	182	109	90	5	78	160	183	122	334	19
15	16	0	26	131	244	132	139	198	70	27	40	75	128	280	346	344	356	338	339	121	152	144	20	0
16	347	372	335	443	340	328	234	216	88	26	0	2	119	112	144	206	154	599	188	136	92	64	106	178
17	32	54	20	36	76	468	580	528	566	447	298	212	202	145	257	213	360	121	311	313	344	361	138	146
18	352	367	268	253	143	129	1	132	71	130	104	30	48	63	322	439	488	623	329	16	66	37	211	186
19	354	475	565	477	421	325	385	576	325	365	360	470	491	575	547	541	323	495	623	369	173	358	416	283
20	118	239	169	85	72	387	623	452	298	298	362	388	511	455	538	457	407	258	303	623	332	209	362	272
21	615	561	326	189	247	381	283	623	623	623	608	623	623	623	623	623	623	618	479	556	623	623	368	220
22	326	324	426	385	387	329	353	600	623	623	623	623	623	623	623	623	623	623	623	614	619	623	623	623
23	177	93	227	246	279	238	510	555	574	445	436	340	241	242	265	414	304	359	585	320	372	400	623	571
24	623	623	623	623	623	623	579	594	623	623	623	577	536	407	348	324	279	501	490	493	596	618	623	623
25	15	217	287	386	445	201	17	14	16	22	55	54	70	63	103	112	59	47	19	66	101	29	30	19
26	0	20	8	31	35	92	50	81	131	144	239	111	115	100	51	50	47	57	80	3	53	34	12	66
27	0	0	18	1	44	35	6	125	153	166	76	55	218	224	271	297	337	111	202	83	21	36	24	21
28	8	13	0	0	17	0	0	28	0	49	86	75	95	48	14	12	2	74	141	92	52	53	0	0
29	141	233	214	153	176	52	16	0	0	0	0	0	70	82	127	94	106	126	26	78	0	3	14	45
30	185	30	61	141	189	121	116	13	186	146	173	128	2	9	0	0	0	0	146	39	22	237	17	29
31	59	152	259	246	46	238	168	28	0	94	73	1	225	288	293	409	369	413	357	163	28	63	19	178
32	0	0	0	3	5	0	109	27	23	71	133	216	342	302	256	448	224	363	306	157	56	61	62	18
33	3	238	218	215	225	257	181	178	199	169	64	157	155	226	204	60	104	24	87	205	260	183	155	238
34	53	136	74	8	0	162	74	0	33	34	232	296	227	192	136	42	42	27	15	22	1	29	107	40
35	208	47	106	53	124	206	264	218	64	1	52	0	0	0	43	53	133	238	235	177	45	42	193	268
36	0	0	11	17	0	0	0	0	62	22	4	3	47	142	248	196	101	122	4	16	19	2	82	13
37	0	0	1	3	0	32	22	0	0	30	88	103	131	404	287	333	90	32	4	0	5	32	0	0
38	6	2	79	59	249	259	206	16	54	48	137	72	192	164	136	96	6	0	0	4	0	0	4	12
39	234	126	26	31	126	195	1	0	4	5	7	16	334	202	393	355	23	97	87	5	48	3	0	59
40	137	103	28	118	18	49	0	27	4	0	160	87	286	197	77	233	150	94	165	3	49	0	0	157
41	179	6	38	58	0	3	10	8	124	4	0	0	15	1	0	1	276	61	83	104	55	0	0	31
42	237	242	29	6	49	0	11	2	3	34	0	5	4	45	17	57	22	470	162	69	23	48	57	200
43	31	172	166	234	228	112	56	58	149	97	205	382	550	610	566	536	447	280	189	175	517	115	112	58
44	537	358	380	211	229	89	2	139	257	105	97	222	313	330	444	413	450	490	319	181	238	380	223	272
45	569	556	512	471	467	375	269	209	143	325	353	333	355	372	361	329	351	377	403	339	5	347	387	333
46	230	378	144	92	227	188	160	122	88	77	166	149	190	214	194	243	282	372	396	559	351	183	89	272
47	132	204	126	114	183	211	294	372	295	303	289	280	483	605	483	523	379	290	267	251	286	272	215	339
48	476	504	480	562	569	572	604	602	492	581	584	583	548	500	491	519	323	361	471	474	484	597	590	395
49	145	86	214	102	105	104	0	0	0	2	0	0	4	115	204	271	245	262	278	295	153	45	2	8
50	350	221	14	51	112	65	104	288	483	344	429	324	279	261	218	212	240	201	133	293	239	263	269	304
51	185	256	305	309	274	120	288	175	28	108	42	45	97	82	43	23	98	144	204	0	5	226	118	25
52	0	178	155	19	230	32	93	525	217	295	280	309	408	347	373	301	392	410	224	86	220	267	49	134
53	0	0	0	4	0	0	0	22	68	89	101	51	6	3	5	4	0	0	0	101	1	0	37	0
54	161	231	71	283	247	329	161	40	25	68	156	151	89	47	60	64	126	103	42	46	53	78	0	210
55	225	25	0	0	1	70	294	181	139	95	227	228	245	254	310	205	204	168	178	343	2	0	166	59
56	77	115	164	22	211	26	51	0	58	25	106	44	11	98	243	283	68	461	360	366	328	411	369	271
57	383	135	106	153	58	71	52	24	0	13	52	101	231	196	75	170	187	16	6	0	0	0	102	114
58	62	148	67	40	48	0	145	0	2	6	1	0	1	6	40	9	6	54	135	182	186	58	1	44
59	0	0	2	5	49	2	8	30	49	64	86	158	180	170	75	25	36	8	19	0	57	16	6	129
60	0	0	23	22	32	0	72	25	141	40	30	13	24	48	34	37	35	194	56	100	0	168	71	125
61	291	180	184	129	151	120	65	35	126	194	9	40	0	0	0	4	170	178	66	188	31	30	151	132
62	61	127	42	4	48	4	103	112	5	19	151	199	301	258	129	149	52	0	53	46	60	101	24	0
63	237	86	136	0	32	0	41	10	0	10	102	179	211	340	445	450	130	18	3	0	50	48	64	107
64	0	120	7	3	64	4	82	152	83	0	0	0	104	67	275	199	68	30	89	3	93	88	0	36
65	0	1	0	0	31	17	123	0	267	314	42	0	0	0	1	0	137	9	71	60	20	15	101	34
66	323	48	41	68	61	46	0	31	99	155	203	359	496	527	529	483	141	184	142	348	97	121	143	17
67	76	113	16	4	45	85	33	17	25	121	298	347	135	85	50	86	28	219	125	205	274	168	0	240
68	32	53	40	74	85	84	57	71	157	158	8	0	0	36	35	55	75	36	87	623	152	46	49	13
69	63	75	21	138	0	37	22	26	0	0	98	149	141	196	279	254	114	199	110	6	133	0	68	106
70	6	4	3	4	1	98	0	0	0	0	8	0	23	21	44	18	46	1	0	0	0	3	0	8
71	20	50	64	196	43	220	333	432	430	150	157	49	61	110	200	117	197	245	95	229	104	54	94	158
72	14	79	136	18	0	0	0	0	1	4	23	42	53	80	65	23	8	8	6	0	0	0	23	201
73	590	503	517	529	597	623	619	621	618	611	615	608	610	621	622	614	422	621	584	405	619	540	614	615
74	115	94	194	203	138	135	0	120	0	20	40	61	58	5	23	24	4	54	262	46	91	153	190	216
75	0	0	0	0	0	57	105	183	0	6	97	133	90	175	118	139	68	121	93	8	4	38	22	75
76	154	202	394	406	283	228	212	72	81	35	29	38	75	36	77	103	131	112	90	1	82	79	233	104
77	117	0	4	0	0	23	35	0	0	88	155	375	370	327	351	295	376	339	322	282	230	91	39	1
78	246	78	347	308	311	97	334	77	138	298	351	426	355	303	197	221	229	160	288	370	297	312	327	318
79	79	28	230	189	79	299	116	116	45	2	21	6	31	21	13	5	0	5	14	2	121	41	136	228
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
83	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623
84	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623	623
85	262	447	623	623	623	623	427	326	430	483	482	607	623	623	623	623	623	623	623	543	623	600	596	595
86	264	373	332	301	345	405	557	469	357	147	227	224	240	259	287	406	328	171	304	255	286	368	308	357
87	623	623	623	623	623	623	366	294	399	433	550	623	620	570	559	538	595	528	461	442	565	623	623	623
88	607	623	623	611	353	611	623	623	623	623	623	599	518	467	447	590	489	477	427	349	324	295	403	603
89	618	379	169	161	194	253	432	523	536	613	562	414	442	401	518	427	599	375	439	393	623	623	607	623
90	604	619	609	623	600	385	196	151	145	176	418	197	233	283	268	320	457	353	342	427	316	406	438	304
91	317	371	511	523	382	431	623	606	324	232	272	294	394	337	230	283	441	324	394	437	434	442	350	363
92	275	398	412	464	439	572	601	595	616	623	561	602	603	490	523	489	586	470	421	499	441	451	445	435
93	560	109	184	279	373	532	623	623	407	420	402	395	357	364	249	325	174	513	588	623	576	438	407	322
94	501	623	513	489	464	494	450	468	494	517	415	314	418	455	417	419	374	425	461	451	478	299	294	321
95	412	419	491	440	348	349	230	276	321	346	294	253	270	277	298	299	251	250	290	378	439	447	500	455
96	535	506	579	453	542	549	446	321	226	386	312	612	616	613	611	622	623	586	458	596	623	579	209	418
97	342	239	90	324	262	162	571	398	384	398	336	474	461	417	548	574	623	542	463	433	483	492	194	255
98	415	417	600	466	576	404	476	355	259	323	584	560	507	329	423	437	608	595	564	566	544	504	503	524
99	463	619	613	618	591	522	403	504	527	510	396	469	486	477	496	622	621	477	616	597	477	483	573	621
100	150	163	201	307	373	172	529	580	479	276	286	600	620	617	614	614	623	621	565	623	599	364	477	316
101	467	572	175	296	332	622	562	479	491	569	591	510	411	414	462	436	461	510	433	492	522	506	399	495
102	487	362	462	517	238	276	420	522	394	410	401	474	462	490	546	493	525	485	451	515	298	417	355	476
103	436	181	207	324	234	417	314	492	433	538	504	419	454	343	328	293	360	529	555	550	452	261	505	463
104	496	483	531	415	338	417	398	500	370	545	388	484	586	599	597	611	438	359	534	519	463	576	576	603
105	421	461	579	621	510	623	554	389	535	393	485	482	580	593	576	543	401	519	529	546	544	527	508	608
106	582	601	608	600	601	604	466	335	569	527	446	360	521	427	373	547	567	604	613	579	590	586	573	127
107	149	354	566	394	288	418	491	512	412	563	358	463	561	519	522	369	399	176	579	619	418	436	458	531

Nord Pool	Hour																						
$\beta_{h,i}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	492	393	338	284	355	146	295	420	385	358	346	351	354	304	451	392	321	388	397	362	407	348	381	594
2	153	115	79	139	100	114	15	10	45	41	39	30	34	28	60	72	46	93	19	12	20	81	35	129
3	190	363	384	348	313	366	4	48	66	6	56	35	39	64	53	27	103	125	71	141	238	285	223	497
4	19	57	73	40	39	0	17	73	110	80	89	22	7	8	6	6	9	0	5	24	21	54	16	36
5	315	260	384	411	461	409	387	305	230	206	140	74	43	132	127	148	112	95	63	18	138	189	165	340
6	323	364	359	484	495	365	421	311	194	250	347	345	396	486	344	381	262	376	335	210	319	299	309	442
7	85	442	438	383	445	550	282	273	270	272	256	350	301	229	335	342	380	386	243	116	131	190	32	162
8	11	0	0	24	17	298	2	300	254	220	95	32	6	61	35	17	91	56	51	27	38	171	239	245
9	129	176	165	135	207	437	251	113	171	298	179	82	102	124	175	247	173	100	308	381	423	381	395	252
10	177	45	56	81	56	109	160	105	141	121	61	144	226	220	259	244	329	333	212	194	158	154	173	7
11	14	57	44	44	191	41	2	357	428	384	355	294	83	141	0	31	146	183	176	78	3	5	7	0
12	228	244	259	314	325	95	282	81	50	6	55	121	251	279	269	195	11	21	4	73	59	28	0	0
13	36	163	158	155	144	12	10	25	20	0	23	87	76	68	92	102	99	42	38	82	105	15	0	0
14	143	33	2	15	32	8	57	20	56	17	30	26	113	110	121	65	31	50	9	0	0	0	13	17
15	157	185	171	153	92	1	16	162	266	193	107	6	1	4	1	13	8	110	4	188	37	4	31	55
16	73	73	72	141	266	208	143	218	322	251	216	125	0	19	31	3	151	329	398	248	2	0	5	84
17	189	319	357	402	287	93	306	370	174	240	459	431	443	436	414	419	365	161	142	189	183	90	216	440
18	214	176	211	369	390	483	442	471	404	489	429	415	436	414	422	411	554	639	659	411	275	199	238	307
19	244	277	295	400	458	212	119	38	110	165	148	54	51	48	103	122	48	104	109	121	50	91	60	38
20	194	275	310	429	456	44	70	180	231	257	226	191	233	196	202	216	252	231	275	260	35	27	31	6
21	88	126	163	162	166	179	339	112	65	55	51	68	64	82	70	72	53	17	48	391	459	217	33	7
22	194	437	625	695	696	654	651	728	728	728	728	728	728	724	721	728	707	657	682	677	683	692	674	375
23	15	86	218	359	279	300	721	553	217	128	341	391	533	553	596	432	360	405	408	374	291	247	656	600
24	728	728	728	728	728	728	542	204	127	179	252	216	173	142	117	154	130	129	188	126	157	105	452	728
25	359	374	648	686	551	498	454	454	195	188	189	84	166	363	395	317	303	262	151	187	415	446	496	679
26	124	222	222	271	274	231	200	0	77	34	71	97	15	7	36	43	11	48	37	20	82	53	94	135
27	0	15	74	11	34	124	24	81	143	114	181	93	15	17	54	128	39	8	91	104	117	155	132	251
28	179	157	48	129	140	156	148	197	78	20	0	0	1	0	5	1	3	40	41	57	0	3	0	0
29	466	337	300	272	155	229	77	209	134	107	121	235	236	229	222	200	119	38	50	189	98	34	94	223
30	183	79	267	393	490	441	361	412	318	304	300	508	523	518	489	336	295	309	293	282	289	315	479	337
31	210	327	263	247	263	465	278	87	42	136	110	149	197	194	137	154	170	165	4	2	62	43	99	266
32	347	408	356	364	366	160	105	250	202	67	13	9	57	144	224	225	253	180	270	152	150	63	65	64
33	20	23	81	102	129	54	147	459	467	405	318	226	189	215	196	173	150	112	132	103	60	15	23	10
34	8	121	196	314	378	197	27	75	82	87	85	46	22	34	60	44	44	62	50	33	19	16	29	96
35	7	155	229	227	203	13	114	203	140	190	169	96	17	148	208	318	319	286	278	176	0	6	7	1
36	11	218	293	256	195	98	43	370	347	130	95	99	118	107	105	121	122	186	91	63	0	98	51	3
37	87	149	96	211	153	26	179	412	422	421	244	225	237	225	250	245	206	247	256	139	59	55	62	10
38	35	0	0	4	11	36	257	38	5	37	50	42	136	143	172	213	170	174	235	277	273	192	116	116
39	11	1	0	20	156	116	65	1	3	1	0	0	0	6	44	49	108	5	62	91	152	202	274	283
40	186	6	32	39	66	14	45	109	158	109	43	14	7	10	5	4	115	37	28	22	14	7	16	28
41	340	85	17	28	14	99	141	377	387	388	311	283	215	287	293	295	356	391	235	115	94	399	300	188
42	4	45	30	39	36	125	74	29	69	46	82	38	47	66	68	134	68	103	12	49	198	351	130	37
43	55	62	64	16	7	60	255	375	491	232	138	261	244	185	232	95	82	110	53	14	9	0	40	236
44	46	8	225	458	525	339	349	257	247	409	491	263	85	41	19	12	1	103	304	325	179	13	0	105
45	24	31	14	15	23	15	58	154	120	90	201	323	330	179	172	291	250	213	280	438	617	61	41	8
46	309	219	164	52	44	74	56	99	19	37	66	30	0	53	68	76	62	61	60	22	49	60	165	71
47	464	637	638	675	670	629	381	302	133	21	5	19	91	132	136	140	149	89	28	8	0	32	156	226
48	216	288	403	425	446	505	461	244	367	318	391	444	453	438	527	610	628	642	400	327	303	396	485	467
49	530	610	416	436	483	462	488	341	373	372	300	382	360	380	333	432	401	384	422	449	658	659	680	573
50	6	0	6	34	125	137	124	81	293	249	256	153	187	111	65	74	33	21	3	145	125	270	242	132
51	10	50	30	14	18	105	66	134	117	124	83	129	184	143	117	211	162	269	276	144	130	134	132	10
52	0	6	46	52	46	112	154	231	89	93	107	104	104	104	109	110	107	146	154	73	7	10	14	86
53	51	16	74	88	374	336	136	96	62	48	44	126	119	112	101	119	96	89	64	190	133	209	102	187
54	131	258	363	302	221	387	51	385	373	296	316	288	213	139	154	94	45	120	155	229	127	166	70	107
55	249	368	365	375	413	72	53	288	215	171	361	501	420	378	285	222	132	71	47	27	47	92	89	84
56	25	15	28	119	256	286	203	232	161	197	172	305	321	358	312	481	648	499	472	253	381	358	344	291
57	8	58	34	75	36	64	390	144	131	116	107	103	113	316	373	426	266	116	296	360	253	178	171	204
58	6	54	95	54	41	154	236	51	66	62	11	26	63	149	199	172	153	190	184	117	123	85	72	105
59	76	86	103	154	56	54	128	186	112	129	114	147	63	69	81	15	25	102	100	9	98	65	78	107
60	171	265	387	427	357	35	56	161	161	64	148	104	0	1	1	0	0	22	0	4	15	3	33	36
61	67	204	247	291	459	331	137	77	115	102	40	92	75	71	70	57	68	71	3	0	0	3	113	34
62	1	2	29	53	76	37	409	240	183	269	235	208	191	185	188	167	165	163	8	73	2	4	83	21
63	0	15	27	10	17	30	24	101	102	38	23	2	5	6	34	80	141	122	91	28	30	118	248	4
64	56	195	242	257	283	125	101	145	124	47	50	21	0	0	3	68	62	7	44	14	1	49	61	14
65	72	77	63	105	114	192	231	322	421	454	393	319	282	275	232	232	236	268	199	208	70	0	171	269
66	45	48	121	48	59	199	183	21	26	71	148	226	103	59	62	72	41	61	136	253	212	116	173	135
67	22	42	80	60	34	0	24	2	102	53	258	243	276	301	282	258	172	184	124	203	49	97	18	74
68	129	318	339	341	388	275	259	124	55	93	64	71	123	97	147	74	92	141	182	255	245	115	107	61
69	1	3	4	3	8	68	68	115	326	206	192	196	259	263	284	274	249	217	196	257	303	128	75	118
70	102	191	142	89	66	221	192	155	67	38	6	14	7	55	84	141	129	147	193	151	239	496	423	297
71	474	539	541	367	447	208	26	6	135	444	362	272	210	160	158	159	205	247	260	121	132	139	384	211
72	43	63	45	55	92	133	394	312	192	33	34	22	23	60	60	63	26	26	58	74	95	142	148	390
73	522	482	369	457	585	437	665	728	648	626	673	653	643	690	691	639	642	632	651	632	707	704	600	471
74	399	554	626	528	558	338	265	144	28	1	35	35	36	46	26	160	245	223	288	284	180	112	112	56
75	385	391	388	387	385	121	28	76	134	59	45	16	6	13	24	233	243	268	281	93	130	39	163	129
76	452	468	431	429	329	275	489	254	148	128	85	61	88	34	5	89	20	46	108	62	115	106	90	310
77	128	220	296	303	334	288	314	283	498	409	356	304	196	94	70	176	266	170	206	148	126	155	185	350
78	221	284	290	378	443	394	160	171	143	173	239	227	252	201	238	268	390	277	73	90	13	76	183	293
79	390	372	459	515	470	428	174	343	401	360	327	234	56	9	37	101	118	135	273	240	228	194	209	237
80	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
81	0	0	1	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
83	523	711	719	726	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	725	724	728
84	56	78	71	83	113	200	623	728	728	728	726	654	644	656	357	250	310	352	306	246	220	71	123	350
85	274	432	585	658	588	486	577	599	621	627	616	634	594	614	591	587	672	724	718	655	555	530	530	529
86	480	671	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728	728
87	728	720	715	699	378	695	262	393	448	539	700	706	718	728	728	728	724	718	523	438	728	728	728	616
88	354	491	587	573	641	525	571	655	419	584	568	600	587	595	610	509	499	451	408	451	605	593	639	680
89	591	633	599	523	542	728	728	678	716	728	728	728	728	728	728	728	728	728	728	728	728	728	639	722
90	645	483	578	683	667	672	693	702	727	728	728	728	728	728	728	727	721	726	692	610	617	621	680	578
91	502	198	230	313	403	697	691	596	639	628	581	510	501	470	357	366	381	402	486	612	508	305	302	252
92	383	524	527	690	698	714	728	716	695	717	683	669	706	680	690	718	683	642	583	318	398	442	436	476
93	479	493	557	612	525	720	698	713	695	641	590	497	558	602	516	548	652	567	504	726	728	728	728	576
94	355	388	495	681	596	408	323	603	648	651	314	389	655	687	635	421	323	136	229	722	728	728	728	392
95	573	590	585	622	585	549	711	543	538	446	381	282	249	272	553	680	525	470	267	350	212	408	259	579
96	397	500	597	516	557	412	695	554	504	417	528	529	503	500	470	561	585	658	442	278	462	330	328	484
97	370	339	414	652	636	576	591	353	350	331	417	448	567	598	534	442	414	444	552	651	578	550	527	608
98	606	543	645	636	644	527	519	607	649	637	561	427	422	332	326	279	367	360	471	432	386	471	336	339
99	527	597	628	516	394	562	395	630	593	627	654	712	660	640	649	651	593	426	521	395	313	508	566	450
100	494	532	424	464	456	266	536	393	532	556	524	521	401	379	378	392	533	476	513	460	475	689	726	627
101	679	649	552	671	702	728	728	728	728	728	719	715	721	713	715	712	682	645	592	361	453	370	409	595
102	523	449	474	519	572	706	622	649	594	726	713	702	688	709	718	714	711	497	543	563	497	563	502	628
103	665	503	557	522	659	334	591	489	504	665	620	462	406	362	387	398	629	494	452	641	374	415	479	463
104	514	449	458	396	375	546	441	663	715	683	653	621	597	539	509	547	601	700	712	551	425	323	307	342
105	194	612	652	587	618	674	682	496	586	460	514	465	486	512	489	550	656	728	498	598	684	668	702	716
106	512	675	684	657	563	604	681	696	692	514	665	648	525	358	359	322	456	518	508	556	645	668	601	666
107	363	691	666	627	719	728	728	579	560	540	459	529	678	689	728	728	654	668	657	676	715	728	728	728

Table 3. The number of days in the out-of-sample test period for which a given $\beta_{h,i}$ (see Equation (7)) was selected for the **EN75** model. The maximum number of days is 623 (=7 × 89 weeks) for GEFCom2014 and 728 (=7 × 104 weeks) for N2EX (U.K.). A heat map is used to indicate more (→ green) and less (→ red) commonly-selected $\beta_{h,i}$'s. Note that $\beta_{h,83}, \ldots, \beta_{h,86}, \beta_{h,94}, \ldots, \beta_{h,100} \equiv 0$ in the **EN75** model.

GEFCom2014

$\beta_{h,i}$	Hour																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	454	393	0	0	0	81	371	156	126	260	321	389	394	394	376	367	334	250	317	255	257	345	422	448
2	573	623	383	180	59	169	47	0	0	7	129	120	21	55	44	40	24	58	15	10	1	150	36	117
3	31	437	493	245	36	20	191	0	7	5	160	115	50	144	192	218	232	219	72	177	103	0	0	0
4	332	259	511	623	396	176	0	184	35	0	0	20	0	0	19	9	20	30	2	1	0	0	8	68
5	303	285	362	406	617	279	161	207	42	129	184	10	0	0	3	7	5	45	67	6	74	0	0	0
6	249	492	575	599	623	623	546	334	358	209	145	164	291	333	374	375	371	282	51	171	93	63	0	1
7	30	54	51	41	169	17	623	589	537	551	57	86	279	468	618	534	458	322	214	131	265	239	246	258
8	0	0	0	0	59	141	281	576	505	287	0	51	215	178	170	160	107	25	0	22	30	31	147	0
9	8	0	0	0	0	0	3	0	27	73	60	38	236	264	298	357	383	308	242	237	25	208	73	0
10	33	0	0	0	0	0	52	0	38	81	46	184	316	416	496	561	559	426	473	493	457	455	472	182
11	0	0	0	0	0	0	0	0	64	57	120	0	0	0	0	0	0	0	0	0	0	0	0	2
12	0	18	0	0	0	0	0	0	0	0	135	134	137	100	44	90	93	0	0	0	0	0	0	0
13	0	48	0	12	0	0	0	0	0	22	143	289	211	148	107	52	75	105	92	58	14	50	35	7
14	56	60	0	0	0	1	10	93	20	0	17	236	355	305	183	21	11	64	73	147	169	60	57	0
15	0	3	0	36	74	118	144	216	172	62	0	141	304	446	462	399	224	332	350	140	63	85	0	0
16	213	204	129	334	262	515	481	541	462	320	157	197	243	298	373	411	88	453	423	361	209	166	207	0
17	149	152	106	0	0	63	182	18	354	292	137	209	257	362	486	568	623	185	53	80	97	87	35	240
18	523	579	565	485	379	261	3	282	34	13	23	43	24	66	207	239	426	623	280	1	21	22	246	336
19	420	495	623	528	571	617	623	623	528	438	321	374	351	377	267	247	163	301	623	285	214	252	269	185
20	103	191	198	136	116	498	623	623	585	585	588	422	549	476	505	414	316	164	202	623	286	3	263	26
21	587	364	185	45	183	358	362	610	623	623	623	623	617	575	564	478	443	371	159	422	623	613	105	4
22	326	363	438	465	456	370	558	623	623	623	623	623	623	623	623	623	623	623	623	622	622	623	623	623
23	153	112	221	195	177	183	549	574	580	560	512	417	392	420	512	589	568	542	623	518	491	497	623	606
24	623	623	623	623	623	623	573	581	623	623	613	578	571	534	519	421	404	513	564	622	607	623	623	623
25	4	109	157	256	286	35	16	6	8	14	21	17	4	4	84	4	1	22	1	25	0	0	0	26
26	0	2	2	10	3	13	3	0	0	0	2	12	26	37	114	68	36	78	21	38	0	0	0	108
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	10	4	0	66	0	0	0	0
28	0	0	0	0	0	0	0	13	0	0	13	2	41	42	142	97	66	168	119	70	50	44	0	0
29	78	105	140	97	99	3	0	0	0	0	0	0	0	0	0	0	48	36	0	0	0	0	0	0
30	0	0	0	0	13	61	65	0	76	0	52	315	313	306	261	252	119	42	129	25	60	153	45	0
31	5	147	191	208	59	149	51	0	0	49	1	30	1	12	22	78	67	135	53	1	4	0	28	2
32	0	0	0	0	0	0	52	4	27	13	91	105	118	96	91	93	84	51	61	110	83	47	83	9
33	0	95	53	88	62	176	44	2	52	47	27	75	34	17	31	20	71	9	3	60	35	86	0	4
34	19	25	5	0	0	133	8	0	12	41	21	5	2	22	41	12	19	19	0	13	9	2	90	1
35	54	0	87	52	102	196	253	222	65	44	2	1	0	22	58	78	116	118	205	211	181	45	228	271
36	0	0	0	0	0	0	0	10	38	0	0	0	0	0	6	12	14	13	23	85	145	139	86	36
37	0	0	0	2	7	6	0	0	0	0	0	0	0	8	0	0	0	3	0	0	0	0	0	0
38	14	29	102	121	250	218	211	47	37	4	0	5	38	122	155	171	1	0	0	8	78	0	22	14
39	252	208	223	177	295	232	228	227	9	104	2	18	256	334	339	236	44	233	186	296	61	47	170	88
40	388	302	266	416	426	513	365	309	217	257	77	0	28	23	8	21	0	152	130	15	97	48	41	315
41	79	0	0	0	28	142	145	375	499	506	21	0	0	0	0	0	44	0	0	143	90	48	0	0
42	57	0	0	0	0	0	0	0	17	22	93	96	112	249	118	171	16	278	79	34	0	0	0	150
43	0	88	0	0	0	0	0	0	9	24	136	478	540	572	550	549	519	125	135	271	448	112	106	83
44	513	373	321	277	280	101	0	0	17	3	15	287	359	359	407	395	428	563	280	0	155	518	237	382
45	587	563	558	470	431	350	209	215	138	334	387	369	342	422	427	417	420	441	621	510	278	427	438	540
46	186	208	56	0	6	105	119	73	36	25	42	1	1	0	0	0	0	0	8	421	131	0	0	23
47	55	200	69	53	163	217	319	470	292	302	292	239	153	150	73	91	9	61	158	5	268	213	0	224
48	398	459	451	509	516	501	558	507	333	526	570	567	533	281	232	192	131	150	304	326	329	336	265	229
49	0	0	0	0	26	27	0	0	0	0	0	0	22	50	146	120	111	79	5	26	18	4	60	18
50	173	163	17	46	70	53	9	210	380	280	351	288	258	250	255	221	354	232	165	299	222	260	279	299
51	204	233	305	308	250	67	180	35	13	115	30	151	214	123	176	153	253	169	235	0	20	125	90	47
52	31	306	182	92	225	62	125	518	224	294	249	90	156	128	143	114	233	180	220	130	195	272	128	9
53	0	0	0	0	0	0	0	1	0	0	4	0	0	0	0	6	0	0	0	20	0	0	0	0
54	83	196	94	240	224	390	141	60	4	0	25	172	171	261	307	320	452	175	38	55	58	146	40	22
55	243	79	135	45	72	114	511	221	189	137	67	60	162	235	252	270	339	125	80	172	15	0	53	34
56	189	283	201	188	354	5	37	6	32	77	144	0	0	0	0	0	0	84	50	23	128	17	13	89
57	214	81	93	108	93	17	17	0	0	0	17	13	152	182	111	252	355	3	0	0	0	0	0	22
58	33	61	43	38	64	3	4	39	51	55	15	0	0	0	0	0	0	9	75	66	85	55	2	0
59	0	0	4	0	0	0	0	3	4	9	8	24	25	2	3	0	0	0	0	0	15	0	0	0
60	0	0	1	0	0	0	3	34	50	2	0	0	0	1	3	15	80	189	4	35	0	95	87	118
61	197	239	271	190	178	201	45	0	31	0	0	0	0	0	0	0	93	391	232	68	16	57	316	197
62	20	50	4	0	0	0	0	0	0	0	60	236	350	330	242	191	77	0	60	55	66	60	44	0
63	123	119	113	231	93	160	79	39	38	30	5	137	185	269	257	309	5	0	52	24	48	5	86	95
64	29	147	78	30	54	301	109	92	29	17	0	0	52	116	259	285	100	0	130	2	21	0	0	14
65	23	3	17	46	23	319	584	456	594	564	329	0	0	0	0	0	66	0	4	6	38	0	0	40
66	47	0	0	9	10	0	0	0	0	0	90	233	355	467	517	403	79	200	57	72	14	50	95	16
67	1	2	0	0	2	0	0	0	0	0	141	205	27	24	27	45	25	99	211	0	110	39	0	40
68	8	12	0	0	3	7	0	3	6	0	0	0	0	0	19	25	0	12	0	457	0	85	23	52
69	59	20	11	52	0	4	0	0	0	0	68	62	27	70	128	161	7	103	49	0	30	0	0	127
70	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	7	79	43	3	0	1	4	0	7
71	0	0	0	0	14	21	171	145	123	0	0	0	0	10	42	115	15	43	22	20	4	0	0	0
72	0	0	0	0	0	0	0	0	0	0	0	0	70	269	317	349	454	379	300	153	22	22	168	358
73	617	540	500	433	536	623	623	623	605	607	613	602	613	623	623	623	623	623	611	614	588	560	622	623
74	53	89	175	257	156	53	0	28	0	0	0	0	0	0	0	0	0	48	125	55	80	43	7	43
75	0	0	0	68	0	57	38	36	0	0	0	29	24	31	23	0	11	30	43	0	0	0	0	0
76	145	213	368	422	257	247	214	131	192	55	45	47	54	0	9	57	42	84	51	0	81	81	83	34
77	157	0	32	0	4	0	0	0	0	0	0	0	0	5	5	30	4	0	0	18	0	0	60	0
78	176	234	303	231	235	96	314	111	259	341	438	413	340	229	179	225	214	162	280	326	316	361	349	234
79	0	164	200	206	165	223	21	5	0	5	41	32	58	36	11	0	73	5	82	169	285	139	72	162
80	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
81	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
82	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
83																								
84																								
85																								
86																								
87	623	623	623	623	581	554	623	623	623	623	608	570	553	563	593	605	622	623	591	607	337	162	309	585
88	623	623	623	515	323	623	623	623	623	623	623	623	623	623	623	623	623	623	584	267	248	219	344	98
89	592	406	257	116	452	623	623	623	623	623	623	623	601	538	518	544	473	623	603	623	525	161	52	331
90	343	523	524	602	607	235	77	342	121	33	74	101	120	162	164	237	238	108	284	320	263	399	421	622
91	492	463	493	435	228	421	575	622	403	443	527	608	570	542	600	621	536	608	533	301	218	117	128	290
92	370	428	347	324	290	623	590	603	620	623	623	623	623	623	623	623	623	623	623	614	592	554	542	444
93	566	63	151	257	253	364	518	522	547	551	568	541	531	533	508	522	482	383	258	623	615	568	352	418
94																								
95																								
96																								
97																								
98																								
99																								
100																								
101	276	405	254	215	396	623	622	490	396	507	535	403	474	468	501	549	438	418	303	269	455	350	439	218
102	556	402	392	280	61	298	268	333	116	204	279	401	411	496	604	599	609	544	339	572	353	320	163	342
103	355	211	100	107	241	382	373	408	554	523	435	406	511	427	424	517	518	285	387	323	457	198	541	465
104	443	259	289	257	226	280	438	486	289	400	368	514	597	593	590	591	586	492	286	326	177	296	423	607
105	432	507	623	623	592	623	378	380	570	499	432	455	448	443	443	423	591	464	479	526	519	576	597	400
106	595	501	558	568	551	601	203	237	527	436	228	271	279	379	314	205	388	458	483	581	594	576	539	316
107	325	121	375	153	231	360	418	571	418	270	528	471	539	394	444	512	548	270	408	422	607	455	170	249

N2EX (UK)

$\beta_{h,i}$	Hour																							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	728	291	287	382	374	370	393	115	366	175	137	11	71	84	143	267	67	181	0	142	23	7	0	272
2	144	724	627	0	67	27	35	105	167	453	194	172	358	439	514	509	156	1	372	353	295	198	104	135
3	0	282	382	356	116	94	121	249	113	53	113	103	35	30	48	0	0	0	131	100	4	7	121	0
4	0	106	226	616	566	418	88	164	101	123	77	34	214	64	53	51	68	0	85	248	92	86	105	148
5	66	53	85	183	182	96	96	161	45	0	51	85	8	1	0	0	0	323	4	11	214	0	40	0
6	215	115	170	185	298	550	256	89	2	1	0	1	118	16	51	2	0	44	186	194	142	45	91	36
7	165	57	149	456	226	474	273	285	554	591	443	296	242	45	161	0	0	136	380	80	68	81	0	63
8	430	636	567	558	520	521	698	509	391	220	319	61	58	82	180	11	110	0	106	101	0	0	84	179
9	0	93	389	425	456	435	225	401	547	446	377	75	0	352	200	0	0	39	66	58	15	0	26	224
10	399	68	0	0	0	0	5	0	171	728	728	481	76	333	272	0	0	65	145	68	267	178	115	339
11	0	0	2	202	326	323	224	120	8	0	117	181	0	0	0	0	4	101	405	43	286	440	364	301
12	78	47	2	61	249	144	429	389	70	92	0	495	66	175	260	34	0	234	444	114	208	560	728	384
13	413	69	84	34	158	453	595	289	378	448	153	491	728	696	529	198	54	18	138	59	118	1	41	228
14	20	129	394	478	400	376	368	327	5	52	42	331	161	653	529	157	304	346	181	280	375	67	140	384
15	146	276	358	330	119	295	202	167	34	14	112	270	353	450	718	269	125	132	239	91	122	82	137	297
16	166	210	39	6	87	95	439	306	724	606	569	236	78	135	386	698	634	0	271	0	87	0	205	320
17	303	278	256	216	40	226	224	394	171	50	281	404	183	142	380	591	728	127	433	28	53	120	1	240
18	117	432	376	212	97	98	274	87	96	235	305	177	192	385	545	564	598	728	88	692	79	55	1	348
19	316	439	353	126	40	153	331	660	130	0	0	3	41	399	627	721	383	569	728	70	305	6	443	30
20	0	668	703	728	588	569	701	534	208	608	369	110	125	137	20	181	82	397	182	728	86	210	271	65
21	59	0	33	61	270	465	443	571	646	640	314	255	65	75	161	276	96	69	33	495	728	27	22	300
22	66	14	221	456	586	688	728	690	660	658	680	660	594	503	224	426	378	490	501	137	76	728	558	116
23	717	384	463	508	599	518	128	366	214	632	641	668	720	695	511	196	431	591	58	266	203	0	728	436
24	728	728	728	728	728	728	728	711	520	573	510	586	392	14	23	5	53	178	470	481	439	373	169	726
25	577	0	71	19	77	223	43	223	355	356	387	303	81	88	312	143	31	50	100	424	115	15	317	340
26	6	408	361	256	141	221	91	0	11	0	0	11	7	0	129	93	335	248	13	28	0	11	13	0
27	3	21	38	161	12	102	221	295	114	360	214	14	0	81	84	13	160	101	152	126	2	268	281	138
28	27	181	139	140	14	102	427	289	131	215	98	11	112	245	34	6	3	0	230	32	188	38	23	12
29	105	211	317	310	425	223	219	135	266	89	10	47	0	45	87	138	0	0	488	126	17	0	11	121
30	15	103	0	25	255	442	279	0	0	146	4	0	71	229	355	176	9	189	1	356	57	47	264	216
31	0	57	74	430	313	36	377	70	0	66	5	0	1	171	82	43	198	251	61	38	5	101	71	2
32	8	78	56	12	17	159	94	17	212	167	25	0	0	2	0	0	40	87	6	0	143	108	0	4
33	7	319	310	283	294	157	83	37	109	180	57	0	1	17	0	43	0	1	56	228	66	1	0	0
34	145	0	9	6	60	343	267	2	56	0	33	0	11	0	0	4	0	8	152	394	203	105	473	473
35	310	105	14	11	117	10	137	393	279	1	132	69	165	74	117	67	0	1	26	243	25	1	159	17
36	83	33	108	22	55	315	195	95	2	182	183	266	287	309	244	376	303	279	168	13	381	339	292	227
37	17	26	62	73	130	183	166	81	205	191	260	371	400	473	428	217	294	102	188	40	1	286	278	190
38	0	25	183	251	158	172	28	29	45	173	299	459	504	70	61	2	252	67	36	1	49	220	240	93
39	259	107	26	202	230	200	132	8	3	13	0	15	175	339	245	79	150	114	70	18	1	0	412	118
40	0	348	111	5	25	5	19	55	110	1	0	1	31	222	180	285	479	92	63	533	47	147	79	235
41	141	0	5	87	98	167	95	67	12	117	0	353	387	151	158	32	313	79	103	48	38	86	208	139
42	50	70	2	31	83	15	17	24	21	59	0	65	129	161	484	231	76	338	114	104	50	92	124	198
43	126	0	0	0	0	25	5	0	2	0	148	55	10	23	78	194	214	451	678	84	374	7	0	0
44	312	106	123	135	186	143	0	24	139	0	11	69	34	75	77	0	100	161	224	354	264	159	332	116
45	145	165	215	195	196	153	238	366	355	328	28	92	25	24	162	139	62	30	388	262	728	32	63	11
46	24	146	56	48	114	93	0	63	12	235	93	60	61	39	10	52	5	15	289	415	24	728	146	24
47	250	56	23	43	86	19	1	11	4	0	19	178	88	14	170	109	125	0	24	122	281	67	451	23
48	0	268	514	532	431	165	147	122	38	198	363	73	174	57	0	95	0	185	122	375	225	616	39	80
49	198	67	82	120	178	203	207	279	172	237	122	243	44	160	290	426	454	517	181	50	256	0	58	433
50	0	114	71	127	175	165	294	1	16	16	126	113	90	24	205	16	0	71	0	92	16	72	0	25
51	15	273	361	231	223	290	233	119	36	0	59	106	170	118	26	0	49	10	82	200	115	135	33	11
52	66	110	377	219	219	207	336	389	481	331	0	26	69	107	122	9	0	8	56	134	184	87	0	41
53	40	140	75	287	444	433	340	318	264	136	310	246	22	40	123	122	58	14	363	361	93	263	171	110
54	159	51	58	0	0	85	55	213	44	3	74	2	10	178	224	44	0	34	23	2	8	1	0	94
55	0	0	0	30	8	1	115	251	142	138	137	0	0	0	1	41	19	86	98	0	10	47	85	5
56	2	61	97	101	89	340	332	247	128	0	108	170	329	219	110	58	40	57	0	48	58	30	0	0
57	10	181	73	73	0	205	59	27	132	1	324	80	90	2	0	23	0	16	278	205	58	86	1	0
58	155	113	85	25	26	126	348	215	279	393	310	314	172	39	0	0	0	31	0	77	21	0	17	3
59	29	8	0	6	110	36	73	224	290	395	242	147	3	0	0	8	0	0	162	89	126	5	138	87
60	235	68	84	14	181	186	195	384	395	218	16	142	131	38	49	330	480	301	488	111	29	516	506	169
61	0	9	146	0	7	88	65	6	10	43	196	447	409	189	320	316	416	56	11	166	150	231	151	200
62	318	12	36	21	102	323	91	7	56	0	1	194	58	218	57	19	78	61	197	92	93	153	34	344
63	411	110	319	230	343	225	217	204	62	77	43	61	257	482	473	297	179	399	77	42	1	163	110	126
64	52	104	411	28	88	320	293	128	99	205	179	27	0	183	499	495	12	1	1	60	34	95	0	1
65	3	0	0	3	19	289	254	188	50	7	34	150	151	239	402	251	130	16	63	27	228	216	26	0
66	114	0	0	427	557	534	489	556	400	334	407	273	4	191	188	144	0	308	189	267	0	0	20	175
67	23	611	110	21	2	28	101	1	0	34	142	156	87	55	4	64	42	150	538	88	62	333	8	137
68	129	16	20	0	184	330	235	58	45	146	44	73	153	70	160	9	2	143	9	728	0	521	320	326
69	300	108	173	18	32	52	3	70	210	162	221	9	9	12	60	26	0	87	158	212	728	42	0	1
70	0	123	1	44	3	33	151	2	15	52	88	73	111	85	18	0	0	5	94	141	478	728	283	89
71	31	0	7	5	47	151	109	22	273	660	482	492	108	15	0	53	276	252	274	130	13	107	572	210
72	358	24	4	119	10	137	2	42	0	0	108	81	178	207	219	181	231	329	119	29	28	168	404	482
73	440	509	396	441	379	289	712	728	728	728	728	720	464	554	543	666	636	728	728	728	728	723	728	475
74	15	97	122	241	553	368	393	252	208	21	27	0	1	105	294	2	9	339	74	171	88	254	58	31
75	0	10	0	66	111	191	249	16	33	0	94	0	21	161	170	188	107	129	400	26	0	130	0	176
76	8	0	96	0	5	206	188	141	165	0	1	0	107	323	163	121	0	0	0	75	124	130	3	0
77	55	230	233	150	36	109	196	436	620	400	113	0	123	100	101	379	410	23	200	646	76	68	102	185
78	115	52	35	129	144	124	0	1	7	81	0	107	478	214	91	91	166	36	156	17	15	1	52	261
79	39	169	9	0	98	232	160	155	197	276	19	79	121	72	130	288	41	193	449	3	24	57	137	276
80	50	114	80	60	1	38	0	76	6	0	130	120	5	0	0	82	170	0	0	0	0	0	0	21
81	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0
82	0	6	0	0	0	0	9	0	0	108	10	0	0	0	0	0	0	0	0	0	0	0	0	0
83																								
84																								
85																								
86																								
87	728	728	728	643	590	239	673	673	728	727	95	728	728	619	728	728	728	728	122	689	575	213	140	447
88	728	728	728	492	554	358	708	728	728	728	728	728	35	728	641	0	329	102	69	336	520	728	519	136
89	458	310	302	336	203	122	712	728	728	715	711	536	354	175	97	693	718	728	728	436	187	1	181	133
90	61	9	0	91	110	115	0	137	700	495	0	0	518	130	0	118	201	312	101	97	81	8	0	68
91	0	181	62	120	440	207	89	497	355	468	237	243	1	58	14	81	81	48	167	232	125	144	0	0
92	48	170	181	73	135	257	489	415	346	394	85	371	256	24	30	110	600	685	148	211	151	272	13	83
93	0	293	343	353	156	277	312	302	377	728	726	561	403	14	424	708	643	728	728	728	728	728	706	271
94																								
95																								
96																								
97																								
98																								
99																								
100																								
101	7	620	696	222	290	285	446	497	461	257	283	388	343	445	236	281	445	597	608	347	33	499	469	17
102	47	202	201	286	468	439	185	250	170	34	1	163	240	96	138	51	121	214	448	347	252	73	207	26
103	355	613	494	449	571	478	387	727	322	557	391	378	131	121	116	75	4	20	446	710	259	104	20	315
104	236	174	6	493	523	658	600	724	640	728	271	131	117	325	228	7	275	351	495	649	494	496	524	425
105	271	375	407	248	199	446	433	708	381	569	523	569	671	296	174	165	300	156	435	649	451	722	496	559
106	303	293	194	129	384	284	304	618	721	728	159	248	533	177	94	40	3	136	556	672	180	80	86	117
107	120	68	139	101	268	487	484	327	579	193	151	242	411	278	152	316	372	504	728	236	297	728	230	286

5. Conclusions

A key point in electricity price forecasting (EPF) is the appropriate choice of explanatory variables. The typical approach has been to select predictors in an ad hoc fashion, sometimes using expert knowledge, but very rarely based on formal selection or shrinkage procedures. However, is this the right approach? Can the application of automated selection and shrinkage procedures to large sets of explanatory variables lead to better forecasts than those of the commonly-used expert models?

Conducting an empirical study involving state-of-the-art parsimonious autoregressive structures as benchmarks, datasets from three major power markets and five classes of automated selection and shrinkage procedures (single-step elimination, stepwise regression, ridge regression, lasso and elastic nets), we have addressed these important questions. To this end, we have compared the predictive performance of 20 types of models over three two-year-long out-of-sample test periods in terms of the robust weekly-weighted mean absolute error (WMAE) and tested the statistical significance of the results using the Diebold–Mariano [31] test.

We have shown that two classes of selection and shrinkage procedures—the lasso and elastic nets—lead to on average better performance than any of the considered expert benchmarks. On the other hand, single-step elimination, stepwise regression and ridge regression are not recommended for EPF as they do not yield significant accuracy gains compared to well-structured parsimonious autoregressive models. The lasso has been recently shown to perform well in EPF [9,18], but it is the more flexible elastic net that stands out as the best performing model overall. Given that both are automated procedures that do not require advanced expert knowledge or supervision, our results may have far reaching consequences for the practice of electricity price forecasting.

We have also looked at variables selected by the elastic net algorithm to gain insights for constructing efficient parsimonious models. In particular, we have confirmed the high explanatory power of the load forecasts for the target hour, of last day's prices for the same or neighboring hours and of the price for the same hour a week earlier. Somewhat surprisingly, we have found that not only the last available data point (price for Hour 24), but also prices for Hours 21–23 of the previous day should be considered when building expert models.

Acknowledgments: The study was partially supported by the National Science Center (NCN, Poland) through Grants 2015/17/B/HS4/00334 (to Bartosz Uniejewski and Rafał Weron) and 2013/11/N/HS4/03649 (to Jakub Nowotarski).

Author Contributions: All authors conceived of and designed the forecasting study. Bartosz Uniejewski performed the numerical experiments. Jakub Nowotarski and Rafał Weron supervised the experiments. All authors analyzed the results and contributed to writing the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Weron, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. *Int. J. Forecast.* **2014**, *30*, 1030–1081.
2. Zareipour, H.; Canizares, C.A.; Bhattacharya, K. Economic impact of electricity market price forecasting errors: A demand-side analysis. *IEEE Trans. Power Syst.* **2010**, *25*, 254–262.
3. Hong, T. Crystal Ball Lessons in Predictive Analytics. *EnergyBiz Mag.* **2015**, 35–37.
4. Amjady, N.; Keynia, F. Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm. *IEEE Trans. Power Syst.* **2009**, *24*, 306–318.
5. Gianfreda, A.; Grossi, L. Forecasting Italian electricity zonal prices with exogenous variables. *Energy Econ.* **2012**, *34*, 2228–2239.
6. Maciejowska, K. Fundamental and speculative shocks, what drives electricity prices? In Proceedings of the 11th International Conference on the European Energy Market (EEM14), Kraków, Poland, 28–30 May 2014; doi:10.1109/EEM.2014.6861289.
7. Ludwig, N.; Feuerriegel, S.; Neumann, D. Putting Big Data analytics to work: Feature selection for forecasting electricity prices using the LASSO and random forests. *J. Decis. Syst.* **2015**, *24*, 19–36.

8. Monteiro, C.; Fernandez-Jimenez, L.A.; Ramirez-Rosado, I.J. Explanatory information analysis for day-ahead price forecasting in the Iberian electricity market. *Energies* **2015**, *8*, 10464–10486.
9. Ziel, F.; Steinert, R.; Husmann, S. Efficient modeling and forecasting of electricity spot prices. *Energy Econ.* **2015**, *47*, 89–111.
10. Dudek, G. Multilayer perceptron for GEFCom2014 probabilistic electricity price forecasting. *Int. J. Forecast.* **2016**, *32*, 1057–1060.
11. Keles, D.; Scelle, J.; Paraschiv, F.; Fichtner, W. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Appl. Energy* **2016**, *162*, 218–230.
12. Karakatsani, N.; Bunn, D. Forecasting electricity prices: The impact of fundamentals and time-varying coefficients. *Int. J. Forecast.* **2008**, *24*, 764–785.
13. Misiorek, A. Short-term forecasting of electricity prices: Do we need a different model for each hour? *Medium Econom. Toepass.* **2008**, *16*, 8–13.
14. Amjady, N.; Keynia, F. Electricity market price spike analysis by a hybrid data model and feature selection technique. *Electr. Power Syst. Res.* **2010**, *80*, 318–327.
15. Voronin, S.; Partanen, J. Price forecasting in the day-ahead energy market by an iterative method with separate normal price and price spike frameworks. *Energies* **2013**, *6*, 5897–5920.
16. Barnes, A.K.; Balda, J.C. Sizing and economic assessment of energy storage with real-time pricing and ancillary services. In Proceedings of the 2013 4th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), Rogers, AR, USA, 8–11 July 2013; doi:10.1109/PEDG.2013.6785651.
17. González, C.; Mira-McWilliams, J.; Juárez, I. Important variable assessment and electricity price forecasting based on regression tree models: Classification and regression trees, bagging and random forests. *IET Gener. Transm. Distrib.* **2015**, *9*, 1120–1128.
18. Ziel, F. Forecasting Electricity Spot Prices Using LASSO: On Capturing the Autoregressive Intraday Structure. *IEEE Trans. Power Syst.* **2016**, doi:10.1109/TPWRS.2016.2521545.
19. Misiorek, A.; Trück, S.; Weron, R. Point and interval forecasting of spot electricity prices: Linear vs. non-linear time series models. *Stud. Nonlinear Dyn. Econom.* **2006**, *10*, doi:10.2202/1558-3708.1362.
20. Weron, R. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*; John Wiley & Sons: Chichester, UK, 2006.
21. Weron, R.; Misiorek, A. Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *Int. J. Forecast.* **2008**, *24*, 744–763.
22. Serinaldi, F. Distributional modeling and short-term forecasting of electricity prices by Generalized Additive Models for Location, Scale and Shape. *Energy Econ.* **2011**, *33*, 1216–1226.
23. Kristiansen, T. Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy* **2012**, *49*, 328–332.
24. Nowotarski, J.; Raviv, E.; Trück, S.; Weron, R. An empirical comparison of alternate schemes for combining electricity spot price forecasts. *Energy Econ.* **2014**, *46*, 395–412.
25. Gaillard, P.; Goude, Y.; Nedellec, R. Additive models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting. *Int. J. Forecast.* **2016**, *32*, 1038–1050.
26. Maciejowska, K.; Nowotarski, J.; Weron, R. Probabilistic forecasting of electricity spot prices using factor quantile regression averaging. *Int. J. Forecast.* **2016**, *32*, 957–965.
27. Nowotarski, J.; Weron, R. To combine or not to combine? Recent trends in electricity price forecasting. *ARGO* **2016**, *9*, 7–14.
28. Hong, T.; Pinson, P.; Fan, S.; Zareipour, H.; Troccoli, A.; Hyndman, R.J. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. *Int. J. Forecast.* **2016**, *32*, 896–913.
29. Maciejowska, K.; Nowotarski, J. A hybrid model for GEFCom2014 probabilistic electricity price forecasting. *Int. J. Forecast.* **2016**, *32*, 1051–1056.
30. Nowotarski, J.; Weron, R. On the importance of the long-term seasonal component in day-ahead electricity price forecasting. *Energy Econ.* **2016**, *57*, 228–235.
31. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263.
32. Garcia-Martos, C.; Conejo, A. Price forecasting techniques in power systems. In *Wiley Encyclopedia of Electrical and Electronics Engineering*; Wiley: Chichester, UK, 2013; pp. 1–23, doi:10.1002/047134608X.W8188.
33. Burger, M.; Graeber, B.; Schindlmayr, G. *Managing Energy Risk: An Integrated View on Power and Other Energy Markets*; Wiley: Chichester, UK, 2007.

34. Nogales, F.J.; Contreras, J.; Conejo, A.J.; Espinola, R. Forecasting next-day electricity prices by time series models. *IEEE Trans. Power Syst.* **2002**, *17*, 342–348.
35. Conejo, A.J.; Contreras, J.; Espínola, R.; Plazas, M.A. Forecasting electricity prices for a day-ahead pool-based electric energy market. *Int. J. Forecast.* **2005**, *21*, 435–462.
36. Paraschiv, F.; Fleten, S.E.; Schürle, M. A spot-forward model for electricity prices with regime shifts. *Energy Econ.* **2015**, *47*, 142–153.
37. Broszkiewicz-Suwaj, E.; Makagon, A.; Weron, R.; Wyłomańska, A. On detecting and modeling periodic correlation in financial data. *Physica A* **2004**, *336*, 196–205.
38. Bosco, B.; Parisio, L.; Pelagatti, M. Deregulated wholesale electricity prices in Italy: An empirical analysis. *Int. Adv. Econ. Res.* **2007**, *13*, 415–432.
39. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning with Applications in R*; Springer: New York, NY, USA, 2013.
40. Bessec, M.; Fouquau, J.; Meritet, S. Forecasting electricity spot prices using time-series models with a double temporal segmentation. *Appl. Econ.* **2016**, *48*, 361–378.
41. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67.
42. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. Royal Stat. Soc. B* **1996**, *58*, 267–288.
43. Ziel, F. Iteratively reweighted adaptive lasso for conditional heteroscedastic time series with applications to AR-ARCH type processes. *Comput. Stat. Data Anal.* **2016**, *100*, 773–793.
44. Hastie, T.; Tibshirani, R.; Wainwright, M. *Statistical Learning with Sparsity: The Lasso and Generalizations*; CRC Press: Philadelphia, PA, USA, 2015.
45. Zou, H.; Hastie, T. Regularization and Variable Selection via the Elastic Nets. *J. Royal Stat. Soc. B* **2015**, *67*, 301–320.
46. Bordignon, S.; Bunn, D.W.; Lisi, F.; Nan, F. Combining day-ahead forecasts for British electricity prices. *Energy Econ.* **2013**, *35*, 88–103.