The Role of Weather Predictions in Electricity Price Forecasting Beyond the Day-Ahead Horizon

Raffaele Sgarlato and Florian Ziel

***Abstract*—Forecasts of meteorology-driven factors, such as intermittent renewable generation, are commonly included in electricity price forecasting models. We show that meteorological forecasts can be used directly to improve price forecasts multiple days in advance. We introduce an autoregressive multivariate linear model with exogenous variables and LASSO for variable selection and regularization. We used variants of this model to forecast German wholesale prices up to ten days in advance and evaluate the benefit of adding meteorological forecasts, namely wind speed and direction, solar irradiation, cloud cover, and temperature forecasts of selected locations across Europe. The resulting regression coefficients are analyzed with regard to their spatial as well as temporal distribution and are put in context with underlying power market fundamentals. Wind speed in northern Germany emerges as a particularly strong explanatory variable. The benefit of adding meteorological forecasts is strongest when autoregressive effects are weak, yet the accuracy of the meteorological forecasts is sufficient for the model to identify patterns. Forecasts produced 2–4 days in advance exhibit an improvement in RMSE by 10–20%. Furthermore, the forecasting horizon is shown to impact the choice of the regularization penalty that tends to increase at longer forecasting horizons.**

***Index Terms*—Forecasting, meteorological factors, regression analysis, power system economics.**

I. Introduction

ELECTRICITY prices fluctuate substantially over time, because the possibility to economically store electricity is limited. This holds true, even when electricity is temporarily transformed into mechanical energy (e.g. by pumped-storage plants), or into chemical energy (e.g. by batteries). Short-term fluctuations of the wholesale electricity price are not only driven by variations in the demand, but also by the supply function changing over time. Intermittent renewable sources, whose generation depends on the weather conditions, are a major contributor.

In Europe, wholesale physical trading is carried out in the day-ahead market and in the intraday market. These markets have hourly or quarter-hourly resolution. At this timescale, the operations of supply units, storage units as well as demand-side flexibility can be subject to intertemporal constraints. For example, thermal generation units, generally, cannot adjust their output by an arbitrary amount due to technical constraints, storage units can supply electricity only if they had accumulated enough energy by that point in time, and industrial demand can be shed or shifted only if underlying processes are scheduled accordingly. These intertemporal dependencies let the bidding strategy depend on future prices. Therefore, price forecasts can support the decision-making of producers with regard to their bidding strategies as well as large consumers with regard to production schedule. Forecasts up to a few days ahead are decisive in day-to-day market operations [1]. Furthermore, operations might depend on prices beyond the next day, when the intertemporal constraints of a unit are particularly pronounced. This is the case for (i) relatively inflexible thermal generation units due to long start-up times and costs associated with cycling activities, (ii) storage units with a duration of more than a few hours, and (iii) large consumers whose (opportunity) costs of re-scheduling power intensive processes decrease for longer lead time. On one hand, this motivates financial trading in the form of future contracts, but also points at the value of forecasting horizons that go beyond the impending day-ahead auction.

In the short-term electricity price forecasting literature, weather effects are usually not incorporated directly. They are incorporated into forecasts of meteorology driven energy production and consumption components [1]–[3]. These components include, most notably, wind and solar power production as well as the electricity load. When forecasting day-ahead electricity prices, taking into account available forecasts of wind power production, solar power production, and load for the next day is an established practice [4]–[7]. This holds basically for all model structures, ranging from high-dimensional linear models to nonlinear ones like deep neural networks [8].

Matsumoto and Endo [9] propose and compare parsimonious models that forecast the weekly average spot price in Japan with the help of temperature forecasts.[1] We are not aware of other academic publications that integrate meteorology based forecasts to forecast electricity prices beyond the impending day-ahead auction. Not accounting for meteorology based forecasts is reasonable for mid- to long-term electricity price forecasting, which cover horizons of months and years [10]. In fact,

Manuscript received 21 November 2021; revised 27 March 2022; accepted 21 May 2022. Date of publication 3 June 2022; date of current version 24 April 2023. This work was supported by the National Science Center (NCN, Poland) through MAESTRO under Grant 2018/30/A/HS4/00444 (to FZ). Paper no. TPWRS-01780-2021. *(Corresponding author: Raffaele Sgarlato.)*

Raffaele Sgarlato is with the Hertie School, 10117 Berlin, Germany (e-mail: sgarlato@hertie-school.org).

Florian Ziel is with the University of Duisburg-Essen, 47057 Duisburg, Germany (e-mail: florian.ziel@uni-due.de).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPWRS.2022.3180119.

Digital Object Identifier 10.1109/TPWRS.2022.3180119

[1] In Japan, the contribution of photovoltaic and wind generation to the electricity mix is modest. In this regional context, prioritizing the integration of temperature rather than wind or solar forecasts is not surprising.

the predictive power of weather forecasts for those horizons is very limited, and basically matches the long-term periodically stationary behavior. Yet, weather forecasts with a horizon of few days (or even weeks) have a reasonable accuracy, which should be utilized in electricity price forecasting. Furthermore, Steinert and Ziel [11] elaborate that these forecasting horizons are highly relevant in practice and corresponding electricity futures markets are highly liquid. We regard the lack of publicly accessible forecasts of intermittent renewable generation and load that reach beyond the day-ahead auction as a key reason for the outlined literature gap. For instance, in Europe the TSOs publish forecasts of wind generation, solar generation and load with (quarter-)hourly granularity, but only at 6pm for the next day [12]. It is worth noting that electricity system forecasts based on weather predictions are highly relevant for a variety of use cases, and multiple business models that offer these products exist [13]. Hence, renewable generation forecasts with longer lead times are likely to exist too for in-house use or for sale in the private sector.

To overcome this limitation, we propose to utilize weather forecasts directly rather than the derived forecasts of production and consumption components. In general, weather data is rarely used directly in the electricity price forecasting literature. An example is given by Ludwig *et al.* [14] that analyze the performance of day-ahead price predictions using ARIMA, ARIMAX, LASSO, and random forest formulations. The authors use actual weather data of stations located in Germany as exogenous variables. The generated insights can be used and further developed by accounting for weather forecasts rather than weather actuals, and by increasing the geographical scope to account for explanatory variables (weather conditions in neighboring regions) that may correlate with German prices because of their impact on cross-border flows. Therefore, we contribute to the literature by:

i) Proposing an hourly electricity price forecasting model which covers horizons beyond the impending day-ahead auction.
ii) Incorporating Europe-wide weather forecasts (temperature, wind speed and direction, irradiation and cloud cover) in addition to established features such as lagged electricity prices, commodity prices, as well as seasonal and calendar effects.
iii) Considering a regularized high-dimensional forecasting model using LASSO to performs efficient automatic feature selection.
iv) Conducting a forecasting study for the German electricity market, including a detailed evaluation of the predictive performance with significance tests.
v) Interpreting the model components, especially the impact of weather forecasts with respect to their temporal and geographic distribution.

In the context of our analysis, we use a high-dimensional linear model with LASSO regularization because of the encouraging results shown by previous studies [14]–[17], the ability to perform feature selection, and the interpretability of results provided by the preservation of linearity. Furthermore, Lago *et al.* [8] illustrate that well-trained high-dimensional linear models and sophisticated deep neural networks achieve similar predictive accuracy, especially for European markets.

The general model structure is introduced in Section II, followed by a detailed description of data sets and implementation remarks in Section III. Section IV describes the setup of the experiment. Section V discusses the results and Section VI concludes.

II. Mathematical Formulation

In this section we outline the structure of the model (II-A), the composition of the feature space (II-B) and the determination of the regression coefficients (II-C).

A. Model Definition

Electricity day-ahead markets are organized in auctions that are cleared daily for delivery on the next day. This implies that a daily structure is embedded into the auction design, and that the price forecast should be produced before the auction closes. In Germany, the day-ahead auction closes at noon, whereas the last known price point refers to 11pm-midnight. Therefore, the point in time when data is collected and the time of the last known target variable do not coincide. We define a notation to reflect this daily structure, and differentiate, where necessary, between the time when the forecast is produced and the time of the last known price.

We denote with t the timestamp with hourly granularity, d_t the day, and s_t the hour of the day. The day when the forecast is produced can thus be indicated with d_t.[2] Further, we use h to indicate the forecasting horizon. We choose $h \in \{1, \dots, H\}$ with $H = 240$, because 10 days is the forecasting horizon of the used meteorological forecasts. Similarly to t, d_h and s_h are the corresponding day-hour decomposition for h. For example, for $h = 30 = 24 + 5 + 1$, we predict the price for the product at 5am-6am ($s_h = 5$) on the 2nd horizon day ($d_h = 2$).[3] When it is necessary to differentiate, we use t to represent the hour of the last known price and t' to represent the hour when the forecast is produced, which we chose to be 2 hours before the day-ahead auction closes, i.e. 10am. Because $d_t = d_{t'}$, this differentiation is mainly relevant when selecting the meteorological forecast in terms of publication time, see Section III-B. A graphical representation of the relationship between the electricity price forecasting horizon and the meteorological forecasting horizon is provided by Fig. 1.

Power prices (y) are represented via a linear combination of a feature vector ($\boldsymbol{x}$) weighted by a column vector of coefficients ($\boldsymbol{\beta}$), plus an error term (ε). The feature vector $\boldsymbol{x}$ contains information available at the time when the forecast is produced and relative to the corresponding auction. Thus, its information depends on the forecasting horizon. For instance, we can describe prices in three days as a function of the day of

[2] As we produce forecasts with a periodicity of S, where $S = 24$ for hourly day-ahead markets, s_t can be omitted.

[3] The timestamp notation h can be derived from the date-hour notation using $h = S(d_h - 1) + s_h$. Conversely, $s_h = (h - 1 \mod S) + 1$ and $d_h = (h - s_h)/S + 1$).

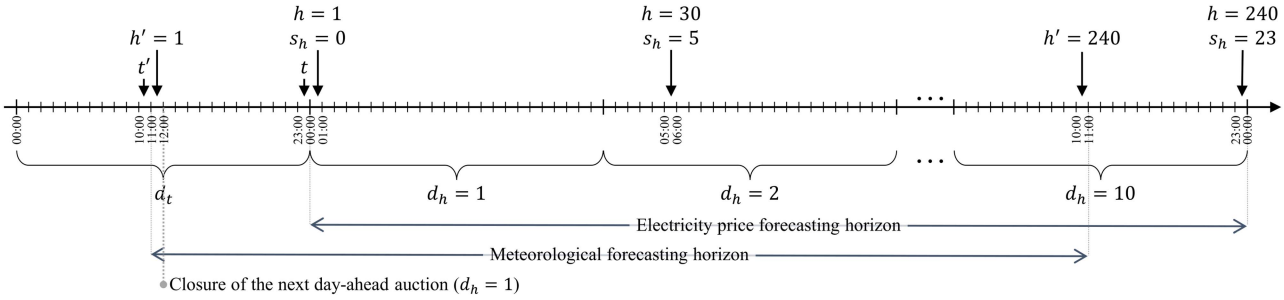

Fig. 1. This graphical representation summarizes how the proposed notation is used to capture: (i) the daily price structure, (ii) the relationship between the price forecasting horizon and the closure of the day-ahead auction, and (iii) the relationship between the price forecasting horizon and the meteorological forecasting horizon.

the week, but not of the result of the preceding auction that is yet unknown. We use the notation $\boldsymbol{x}_{d_t,d_h}$ for the feature vector to indicate its dependency on the day d_t and the forecasting horizon d_h. The feature space does not depend on s_h because of the market design, i.e. the auction being cleared at once for the whole day. Using this notation, the model can be formulated as:

$$y_{t+h} \overset{\text{def}}{=} y_{d_t+d_h,s_h} = \boldsymbol{x}_{d_t,d_h}^{\intercal}\boldsymbol{\beta}_{d_t,d_h,s_h} + \varepsilon_{d_t,d_h,s_h}. \quad (1)$$

The price $y_{d_t+d_h,s_t}$ corresponds to the feature vector $\boldsymbol{x}_{d_t,d_h}^{\intercal}$. Since the prices on the day $d_t + d_h$ are unknown on the day d_t, the training data can only reach to the price $y_{d_t+\cancel{d_h-d_h},s_h}$ corresponding to $x_{d_t-d_h,d_h}^{\intercal}$. We let the training data cover 12 weeks, rendering $T+1 = 84$ observations, and organize this data in a target vector and a training matrix:

$$\boldsymbol{y}_{d_t,s_h} = \begin{bmatrix} y_{d_t,s_h} \\ y_{d_t-1,s_h} \\ \vdots \\ y_{d_t-T,s_h} \end{bmatrix}, \boldsymbol{X}_{d_t,d_h} = \begin{bmatrix} \boldsymbol{x}_{d_t-d_h,d_h}^{\intercal} \\ \boldsymbol{x}_{d_t-d_h-1,d_h}^{\intercal} \\ \vdots \\ \boldsymbol{x}_{d_t-d_h-T,d_h}^{\intercal} \end{bmatrix}. \quad (2)$$

B. Feature Space

The feature space accounts for dummy variables to capture weekly patterns and holiday effects ($\boldsymbol{x}_{d_t,d_h}^{D}$), for weather forecasts for the selected location and weather attributes ($\boldsymbol{x}_{d_t,d_h}^{W}$), for lagged power prices to capture autoregressive effects ($\boldsymbol{x}_{d_t}^{P}$) and for commodities prices, namely gas, coal and EUA futures ($\boldsymbol{x}_{d_t}^{C}$). Lagged power prices and commodity futures depend only on d_t because the same data points are fed to the model regardless of the forecasting horizon d_h. Conversely, the dummy variables depend on d_h because $d_t + d_h$ is used to incorporate the effect of the day of the week and holiday patterns. Also the weather forecast vector depends on the horizon because d_h is used to determine which parts of the weather forecasts are fed to the model with hourly granularity (see formulation in Section III-B). The individual components of the feature space are described in detail in Section III and summarized by the vector:

$$\boldsymbol{x}_{d_t,d_h} = \begin{bmatrix} \boldsymbol{x}_{d_t,d_h}^{D} \\ \boldsymbol{x}_{d_t,d_h}^{W} \\ \boldsymbol{x}_{d_t}^{P} \\ \boldsymbol{x}_{d_t}^{C} \end{bmatrix}. \quad (3)$$

C. Regression Coefficients

The feature matrix and target vector are standardized. We call g the function that takes two arguments (matrices or column vectors), and returns the second argument, standardized according to the column-wise mean and variance of the first one. Furthermore, we define g^{-1} as the function reversing the standardization.[4] The standardized training matrix, the standardized feature vector, and the standardized target vector are denoted by an underscore (_) and defined as follows:

$$\underline{\boldsymbol{X}}_{d_t,d_h} \overset{\text{def}}{=} g(\boldsymbol{X}_{d_t,d_h}, \boldsymbol{X}_{d_t,d_h}), \quad (4)$$

$$\underline{\boldsymbol{x}}_{d_t,d_h} \overset{\text{def}}{=} g(\boldsymbol{X}_{d_t,d_h}, \boldsymbol{x}_{d_t,d_h}^{\intercal})^{\intercal}, \quad (5)$$

$$\underline{\boldsymbol{y}}_{d_t,d_h} \overset{\text{def}}{=} g(\boldsymbol{y}_{d_t,s_h}, \boldsymbol{y}_{d_t,s_h}). \quad (6)$$

The regression results in a coefficient vector:

$$\hat{\boldsymbol{\beta}}_{d_t,d_h,s_h}(\lambda) = \arg\min_{\beta} \left\| \underline{\boldsymbol{X}}_{d_t,d_h}\boldsymbol{\beta}_{d_t,d_h,s_h} - \underline{\boldsymbol{y}}_{d_t,s_h} \right\|_2^2 + \lambda \left\| \boldsymbol{\beta}_{d_t,d_h,s_h} \right\|_1, \quad (7)$$

that is a function of the regularization parameter $\lambda \geq 0$ characterizing the LASSO. This penalty modulates the shrinking effect towards zero on all parameters to reduce the estimation risk and the feature selection.

In a very limited number of forecasts, the solver [18] generates an unstable result characterized by an implausibly high price. This issue was solely affecting the Expanded model (see Section IV) in a few hours and is imputable to the high dimensionality of the feature space. Even if rare, the magnitude of such a forecast error would substantially distort the evaluation of the overall performance. Hence, standardized forecasted prices have been constrained to a predefined interval. This is not unprecedented. For example, Uniejewski *et al.* [19] used an interval of ± 3. We opted for a less stringent constraint (± 6), because it showed to sufficiently mitigate the effect of the unstable results on the evaluation:

$$\hat{\underline{y}}_{d_t,d_h,s_h}(\lambda) = \min(\max(\underline{x}_{d_t,d_h}^{\intercal}\hat{\boldsymbol{\beta}}_{d_t,d_h,s_h}(\lambda), -6), 6). \quad (8)$$

[4]For conciseness, $g(A) \overset{\text{def}}{=} g(A, A)\ \forall A$ could be introduced such that $\underline{\boldsymbol{X}}_{d_t,d_h} = g(\boldsymbol{X}_{d_t,d_h}, \boldsymbol{X}_{d_t,d_h}) = g(\boldsymbol{X}_{d_t,d_h})$.

Finally, the predicted price is obtained by reversing the standardization applied before the regression:

$$\hat{y}_{d_t,d_h,s_h}(\lambda) = g^{-1}(\boldsymbol{y}_{d_t,s_h}, \underline{\hat{y}}_{d_t,d_h,s_h}(\lambda)). \tag{9}$$

The model is tested for multiple values of $\lambda \in \Lambda$ and a number of consecutive days, starting with $d_t = 0$ representing the 1st of March 2021. The regularization penalty $\hat{\lambda}$ is then chosen for each horizon day d_h such the mean-squared error of the prices until d_t is lowest:

$$\hat{\lambda}_{d_t,d_h} = \arg\min_{\lambda\in\Lambda} \sum_{i=d_h}^{d_t} \sum_{s_h} (\hat{y}_{i-d_h,d_h,s_h}(\lambda) - y_{i,s_h})^2. \tag{10}$$

III. Data Sources and Implementation

A. Power Prices

We consider German day-ahead auction prices, available at ENTSO-E transparency [20]. Besides being the target vector of the model, they are also used as explanatory variables in the form of lagged prices. For convenience, they are organized in daily vectors $\boldsymbol{y}_{d_t}$. Furthermore, we define the vector $\boldsymbol{y}^f_{d_t}$ as containing values resulting from applying the generic function f to each of the daily vectors of the last 28 days. This allows to define the feature vector $x^P_{d_t}$ as comprising hourly lagged prices of the last 7 days (including d_t) and the average, the minimum, and the maximum daily prices of the last 28 days (including d_t) for a total of 252 variables.

$$\boldsymbol{y}_{d_t} = \begin{bmatrix} y_{d_t,0} \\ y_{d_t,1} \\ \vdots \\ y_{d_t,S-1} \end{bmatrix}, \boldsymbol{y}^f_{d_t} = \begin{bmatrix} f(\boldsymbol{y}_{d_t}) \\ f(\boldsymbol{y}_{d_t-1}) \\ \vdots \\ f(\boldsymbol{y}_{d_t-27}) \end{bmatrix} \tag{11}$$

$$\boldsymbol{x}^P_{d_t} = \begin{bmatrix} \boldsymbol{y}_{d_t} \\ \boldsymbol{y}_{d_t-1} \\ \vdots \\ \boldsymbol{y}_{d_t-6} \\ \boldsymbol{y}^{avg}_{d_t} \\ \boldsymbol{y}^{min}_{d_t} \\ \boldsymbol{y}^{max}_{d_t} \end{bmatrix} \tag{12}$$

B. Meteorological Forecast Data

Meteorological forecasts are provided by the *Deutscher Wetterdienst* for 5400 locations worldwide, each with an hourly forecast horizon of 10 days. From an area covering approximately continental Europe, we create $L = 100$ partitions with k-means clustering, and use the locations closest to the respective centroids as our 100 selected location (see Fig. 2). The selection is aimed at preserving computational tractability, and the clustering procedure at preserving the original density of the locations' distribution.

From the data provided by the *Deutscher Wetterdienst*, we use the *MOSMIX_S* data package, that is updated hourly and contains 40 weathers attributes [21]. From these attributes, we select those that are expected to be correlated with power prices because of their impact on the residual load and on cross-border flows. The latter motivates the selection not to be limited to locations within Germany. On the supply side, wind speed and wind direction are expected to be correlated with the generation from onshore and offshore wind units.[5] Global irradiance and effective could cover are expected to be correlated with the generation of PV units. On the demand side, temperature is expected to be correlated with the electricity load.[6]

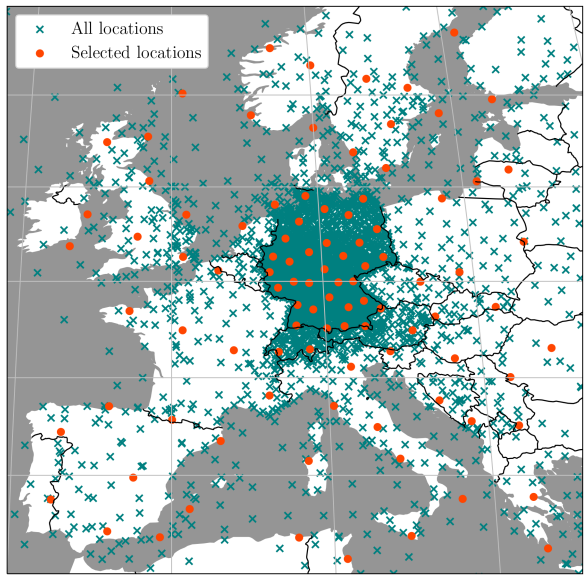

Fig. 2. A selection of one hundred representative locations is used to preserve computational tractability.

Let us denote with $w^{a,l}_{t',h'}$ the meteorological forecast produced at time t' with horizon $h' \in \{1,\dots,H\}$ for the weather attribute a and one of the hundred selected locations l. Note that whereas the horizons h' and h are characterized by the same domain $1,\dots,H$, t' precedes t by 13 hours, since we are using the weather forecast issued at 10am. This is taken into account when weather attributes are grouped in vectors $\boldsymbol{w}^{a,l}_{d_t,d^w_h}$ that depend on the horizon day $d^w_h \in \{0,1,\dots,10\}$ by introducing $s^{min}_{d^w_h}$ and $s^{max}_{d^w_h}$. For each forecasting horizon d_h, a vector $\boldsymbol{v}^{a,l}_{d_t,d_h}$ is constructed containing the hourly meteorological forecasts for that particular day $\boldsymbol{w}^{a,l}_{d_t,d_h}$ as well as the averaged meteorological forecast for all horizon days $\bar{\boldsymbol{w}}^{a,l}_{d_t,d^w_h}$. This aggregation is meant to reduce the dimensionality of the feature space whilst preserving intertemporal dependencies that span across multiple days. Finally, the meteorological covariates vector $\boldsymbol{x}^W_{d_t,d_h}$ is defined as the vectorized matrix $\boldsymbol{W}_{d_t,d_h}$ containing the compound vectors $\boldsymbol{v}^{a,l}_{d_t,d_h}$ for all meteorological attributes and all locations for a

[5] To allow for the wind direction, published in radiant, to be sensibly processed by the model, it has been split into a north-south and a west-east component by applying respectively the cosine and sine functions.

[6] The mentioned weather attributes correspond to the *Deutscher Wetterdienst* abbreviations: FF (wind speed), DD (wind direction), Rad1h (global irradiance), Neff (effective could cover), TTT (temperature)

total of 19800 elements.

$$\boldsymbol{w}^{a,l}_{d_t,d^w_h} = \begin{bmatrix} w^{a,l}_{d_t,d^w_h,s^{min}_{d^w_h}} \\ \vdots \\ w^{a,l}_{d_t,d^w_h,s^{max}_{d^w_h}} \end{bmatrix} \tag{13}$$

$$s^{min}_{d^w_h} = \begin{cases} 11 & \text{if } d^w_h = 0 \\ 0 & \text{otherwise} \end{cases}, s^{max}_{d^w_h} = \begin{cases} 10 & \text{if } d^w_h = 10 \\ 23 & \text{otherwise} \end{cases} \tag{14}$$

$$\boldsymbol{v}^{a,l}_{d_t,d_h} = \begin{bmatrix} \boldsymbol{w}^{a,l}_{d_t,d_h} \\ \bar{w}^{a,l}_{d_t,0} \\ \bar{w}^{a,l}_{d_t,1} \\ \vdots \\ \bar{w}^{a,l}_{d_t,10} \end{bmatrix}, \boldsymbol{W}_{d_t,d_h} = \begin{bmatrix} \boldsymbol{v}^{(a,l)}_{d_t,d_h} \end{bmatrix} \tag{15}$$

$$\boldsymbol{x}^W_{d_t,d_h} = vec(\boldsymbol{W}_{d_t,d_h}) \tag{16}$$

C. Commodity Prices

Commodity prices are downloaded from the Quandl database CHRIS [22] and include gas NCG, EUA and coal ARA futures prices. Because these futures contracts commit traders to buy or sell these commodities at a predefined price when the contract expires, their price does account for the market expectation on the future price development. During this work, the database CHRIS has been partially discontinued. The selected gas and EUA futures prices have been last updated on the 16th of April 2021, which poses a limit on the duration of experiment outlined in Section IV. This motivates our experiment period ending on this date to allow for a consistent evaluation throughout the entire duration of the experiment.

The definition of the vector representing commodity prices is straight forward, as it only contains three values corresponding to the latest settlement prices known by d_t for the selected commodities. This vector is used as is for the entire forecasting horizon, thus it does not depend on d_h:

$$\boldsymbol{x}^C_{d_t} = \begin{bmatrix} \mathrm{Coal}_{d_t} \\ \mathrm{Gas}_{d_t} \\ \mathrm{EUA}_{d_t} \end{bmatrix}. \tag{17}$$

D. Dummy Variables

To capture seasonal patterns with weekly periodicity, we implement weekday dummy variables ($\mathrm{Weekday}^n_i$) and cumulated weekday dummy variables ($\mathrm{WeekdayC}^n_i$), where n represents one of the weekdays and i the day to be forecasted. Furthermore, a holiday dummy variable ($\mathrm{Holiday}_i$) is used to account for holiday effects.[7] These dummies count 15 elements in total; they are defined in (18)–(20) and added to the feature space according to (21) and (23).

$$\mathrm{Weekday}^n_i = \begin{cases} 1 & \text{if i is the } n^{\text{th}} \text{day of the week} \\ 0 & \text{otherwise} \end{cases} \tag{18}$$

$$\mathrm{WeekdayC}^n_i = \sum_{j=1}^{n} \mathrm{Weekday}^n_j \tag{19}$$

$$\mathrm{Holiday}_i = \begin{cases} 1 & \text{if i is a Sunday or a Holiday} \\ 0.5 & \text{if i is a Saturday} \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

$$\mathbf{Weekday}_i = \begin{bmatrix} \mathrm{Weekday}^1_i \\ \mathrm{Weekday}^2_i \\ \vdots \\ \mathrm{Weekday}^7_i \end{bmatrix} \tag{21}$$

$$\mathbf{WeekdayC}_i = \begin{bmatrix} \mathrm{WeekdayC}^1_i \\ \mathrm{WeekdayC}^2_i \\ \vdots \\ \mathrm{WeekdayC}^6_i \end{bmatrix} \tag{22}$$

$$\boldsymbol{x}^D_{d_t,d_h} = \begin{bmatrix} \mathbf{Weekday}_{d_t+d_h} \\ \mathbf{WeekdayC}_{d_t+d_h} \\ \mathrm{Holiday}_{d_t+d_h} \end{bmatrix} \tag{23}$$

IV. Experiment Setup

Not only do weather conditions affect power prices, the relationship between weather conditions and renewable generation is also not linear. Wind power is known to have a strong cubic relation to the wind speed [23]. To evaluate the benefits of using weather forecasts and explore the impact of nonlinear relationships, we perform the experiment using three different variants of the feature matrices, thus resulting in three model variants:

- Baseline – Excludes weather attributes $\boldsymbol{x}^W_{d_t,d_h}$ from the feature matrix to recreate a consistent benchmark model.
- Linear – Accounts for weather attributes, as in (3). The only transformations are the standardization and the decomposition of the wind direction in a north-south and a west-east component.
- Expanded – Adds to the feature matrix shown in (3) and used by the Linear variant squared and cubed weather data $\boldsymbol{x}^W_{d_t,d_h}$.

On each day from the 1st of March 2021 to the 16th of April 2021 the models are (re-)estimated and used to generate hourly forecasts for the following 10 days.[8] Because each weather forecast is stored on the *Deutscher Wetterdienst* server only for a few days, the 1st of March 2021 was the point in time when enough weather forecasts had been regularly downloaded to cover the model estimation window (see Section II-A). Longer

[7] Note that the only bank holidays within the experiment period are the Eastern ones. Therefore, we decided to (partially) flag also Saturdays and Sundays as holidays.

[8] In total, 72000 models are trained daily. This corresponds to the combination of 3 model variants, 240 hours composing the forecasting horizon (due to the multivariate model structure), and 100 values of λ.

estimation windows (up to a few years) are beneficial to capture trends, whereas shorter windows (down to a few weeks) allow for a faster adaptation of the model [24]–[26]. Although electricity price forecasting models tend to use longer estimation windows [7], [11], [14], [15], [27]–[30], this option is not viable in the context of this experiment because: (i) such a long estimation window paired with the exceptionally large number of features used in this study would make the model training computationally intractable,[9] and (ii) not enough historic weather forecasts had been stored at the time of writing. One could also hypothesize, that adopting a longer estimation window is more likely to benefit highly dimensional model variants Linear and Expanded, therefore, further highlighting the benefit of adding weather forecast data. The estimation window covers 84 days (12 weeks). For example, the forecast generated on the 1st of March 2021 is trained on the 84 daily observations ranging from the 6th of December 2020 to the 28th of February 2021. The observation on the 6th of December 2020 is described by the corresponding feature vector as summarized by (3) (including the weather forecast published at 10am on that day and lagged prices that reach back to 8th of November), whose components are described in detail in Section III. The experiment reaches until the 16th of April 2021 because commodity prices were no longer provided by the chosen data source (see Section III-C), totaling 47 forecasts for each λ and model variant. Eventually, the accuracy achieved by these three variations are compared and the significance of the results is quantified with the Diebold-Mariano test. The latter step is particularly important due to the limited number of forecasts.

V. Results

Whereas the models are re-estimated using a rolling window, the selection of the λ_{d_t,d_h} is based on an expanding window, as described by (10). For the sake of simplicity, the following analyses will refer to the best performing regularization parameter reached at the end of the experiment horizon λ_{d_h}. This is motivated by the observation that the selection appears to converge towards a more stable value after an initial burn-in period, as shown in Fig. 3. Part of this stabilizing behavior is due to the nature of the expanding window. As will become apparent in the following analysis, adding weather forecasts when the price forecasting horizon exceeds 5 days does not materialize in accuracy improvements plausibly because of the diminishing precision of the weather forecast. But even for $d_h \leq 5$ noticeable fluctuations $\hat{\lambda}_{d_h}$ remain, especially for the variants Linear and Expanded. Part of these fluctuations are caused by the model adapting when the exceptionally low prices observed around Eastern (4th of March 2021) become part of the training set. Note that, as discussed in Section II-A and described by (2), the point in time when the results of an auction become part of the training matrix depends on the d_h. This is the reason for models trained on longer horizons exhibiting a stronger adaptation delay. The delayed adaptation is particularly apparent in the case of the Linear variant, where all horizon days d_h increase sequentially. The Expanded variant seems to be characterized by more noise, indicating that longer evaluation periods could be beneficial.

[9]By comparison, other proposed parameter rich models contain approximately 400 explanatory variables [29], whereas the weather data alone proposed in (16) counts 19800 elements.

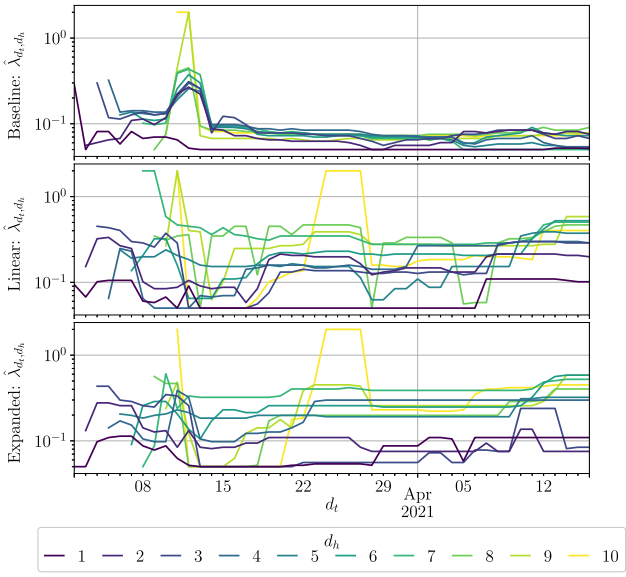

Fig. 3. The best performing regularization parameter ($\hat{\lambda}_{d_t,d_h}$) changes over time due to the expanding window evaluation, but tends to stabilize towards the end of the experiment period (d_t = 16th of April 2021) for most of the forecasting horizons (d_h).

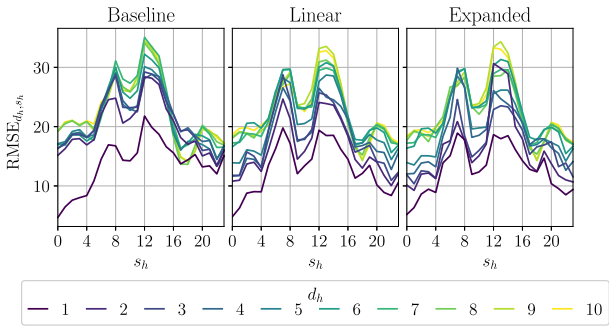

Fig. 4. The hourly RMSE (RMSE_{d_h,s_h}) exhibits a strong daily pattern; error peaks around 8am and noon are explained by the higher variance of historic prices in these hours.

The RMSE_{d_h,s_h} is characterized by a pronounced daily pattern, see Fig. 4. All three models exhibit a similar error distribution across the day, peaking in the morning (7am and 8am) and in the afternoon. This is explained by the standard deviation of historic prices being highest in these hours of the day (see Appendix C, Fig. 13). For the sake of simplicity, we do not focus further on the daily distribution of the error, but refer to the daily error RMSE_{d_h}, where the mean squared error is calculated for all hours of the day.

A. Accuracy

Fig. 5 shows the RMSE_{d_h} for all three variants, and the accuracy improvements relative to the Baseline variant quantified

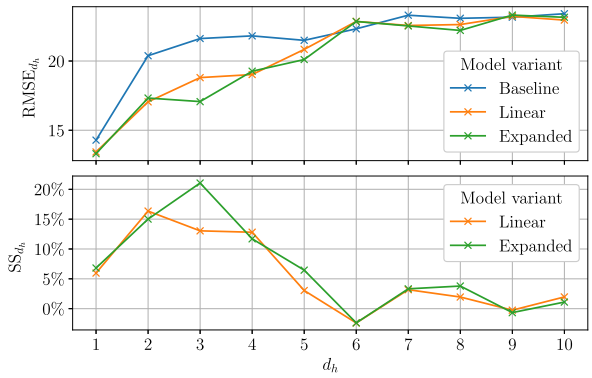

Fig. 5. Adding meteorological forecasts is particularly beneficial when forecasting prices between 2 and 4 days before delivery, whereas the benefit of accounting for non-linear weather effects is ambiguous.

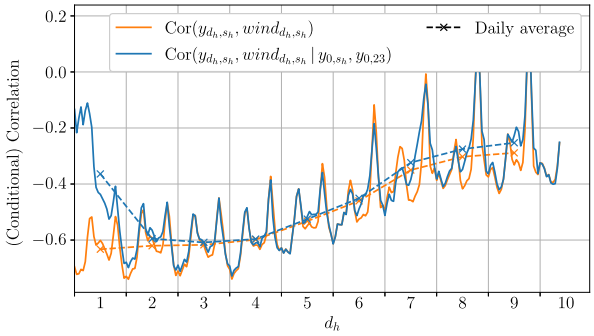

Fig. 6. The correlation between prices and forecasted wind speed is almost constant for the first 4 days of the forecasting horizon. When the last known price $y_{0,23}$ and the price of the last known auction at the corresponding hour y_{0,s_h} are accounted for. The correlation conditional on the lagged prices $\mathrm{Cor}(y_{d_h,s_h}, \mathrm{wind}_{d_h,s_h} \,|\, y_{0,s_h}, y_{0,23})$ shows that the contribution of wind speed is partially captured by lagged prices when forecasting up to 1 d in advance.

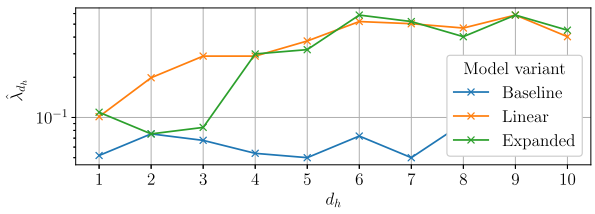

Fig. 7. The best performing regularization parameter ($\hat{\lambda}_{d_h}$) tends to increase at longer forecasting horizons (d_h) especially for the models containing meteorological data.

by the skill score:

$$\mathrm{SS}_{d_h} = 1 - \frac{\mathrm{RMSE}_{d_h}}{\mathrm{RMSE}_{d_h}^{\mathrm{Baseline}}}. \tag{24}$$

All model variants exhibit a similarly low error on the first day of the forecasting horizon. This is caused by the autoregressive effects, which are more pronounced at shorter horizons. Hence, the benefit of adding meteorological data when forecasting the next day appears perceptible with a skill score of about 5%, but moderate relative to the following horizon days. In fact, we observe that the Linear and the Expanded variants outperform

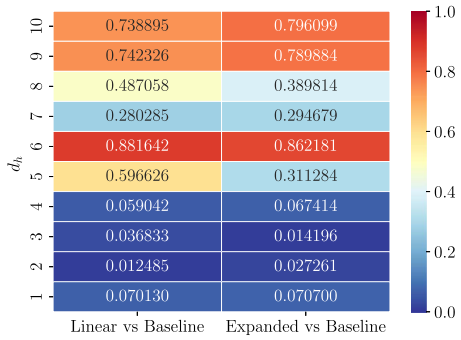

Fig. 8. The p-value derived from the Diebold-Mariano test shows that improvements observed on the 2nd and 3rd day of the forecast horizon are statistically significant.

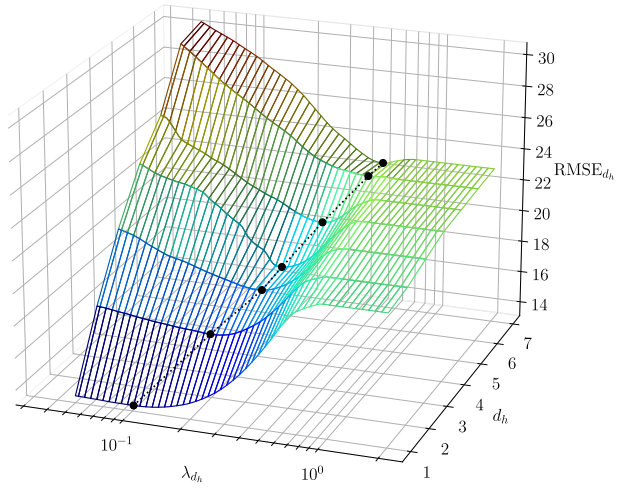

Fig. 9. The lowest error is used to select the regularization penalty $\hat{\lambda}$ (black dots). Especially when forecasting multiple days in advance, the RMSE is particularly sensitive to suboptimal regularization penalties $\lambda \in \Lambda$, and susceptible to overfitting.

the reference variant (Baseline) with a skill score of between 10% and 20% when forecasting 2-4 days in advance. On the 5th horizon day, the improvements are less pronounced, and from the 6th horizon day onward, adding meteorological forecasts results in a less accurate forecast. It is worth noting that, exception made for the 3rd horizon day, adding squared and cubed model meteorological data (Expanded) does not seem to substantially improve the accuracy over the simpler linear model (Linear).

The autocorrelation of prices and the correlation between prices and the forecasted weather explains why weather forecasts are particularly beneficial when forecasting prices 2-4 days in advance. Wind speed forecasted at the location close to Bad Oldesloe, Germany (wind_{d_h,s_h}) exhibited a pronounced predictive power (see section V-B, Fig. 12). Therefore, wind speed at this location has been used as representative weather forecast element to evaluate the correlation with prices. As summarized by Fig. 6, the correlation between prices and the forecasted wind speed $\mathrm{Cor}(y_{d_h,s_h}, \mathrm{wind}_{d_h,s_h})$ is pronounced at the beginning of the forecasting horizon and starts decaying substantially only at forecasting horizons that exceed 4 days. This indicates that wind

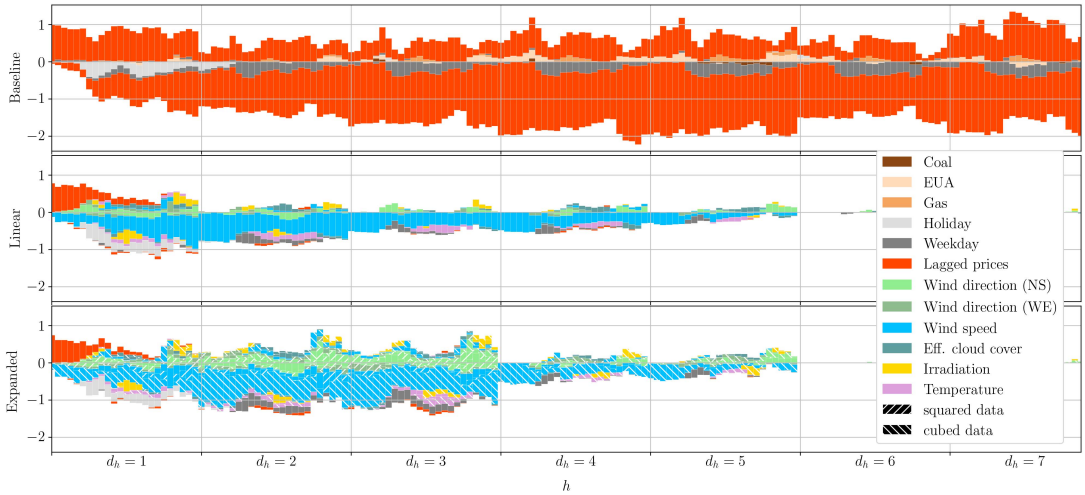

Fig. 10. Coefficients of feature groups over time are aggregated by positive and negative values to highlight the relationship with the electricity price; weather attributes decrease over time due to the deteriorating accuracy of the meteorological forecast.

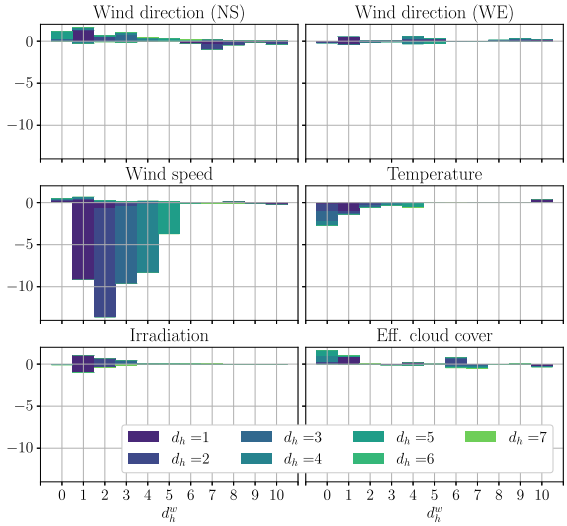

Fig. 11. The sum of negative and positive meteorological coefficients grouped by the horizon of the meteorological forecast for different price forecast horizons (Linear variant) shows that only wind speed and irradiation exhibit a pronounced alignment between the meteorological and the price forecasting horizon.

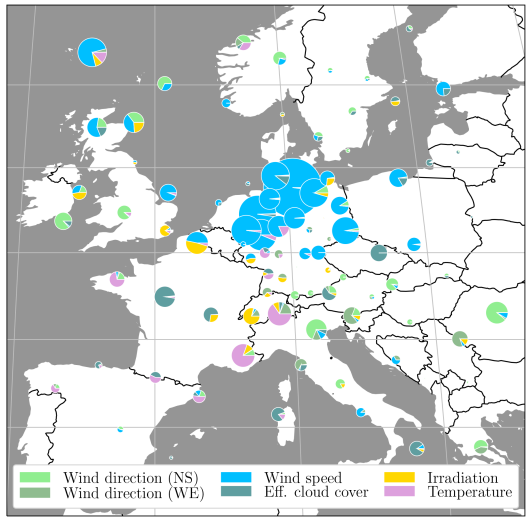

Fig. 12. The geographical distribution of meteorological coefficients (Linear variant) exhibits forecasted wind speed in northern Germany are the most important predictor; the size of the pie charts is proportional to the sum of the absolute coefficients.

speed preserves its predictive power over several days. Wind speed itself is also characterized by autocorrelation. Therefore, when the forecasting horizon is short, part of the information contained in the forecasted wind speed is already contained in the lagged prices. This is captured by the correlation between prices and forecasted wind speed, conditional on the lagged prices $\mathrm{Cor}(y_{d_h,s_h}, \mathrm{wind}_{d_h,s_h} \mid y_{0,s_h}, y_{0,23})$. In fact, the conditional correlation is much weaker only in the first few hours of the forecasting horizon. Therefore, the conditional correlation is highest for horizons between 2 and 4 days, which is in line with the observed accuracy gains observed in Fig. 5.

The negative impact of meteorological data when the horizon exceeds 5 days is explained by Fig. 7. It shows that the models Linear and Expanded require a more stringent regularization to avoid overfitting, thus penalizing also other features like the dummy ones. The downsides of the stronger regularization prevails in this case over the benefits of adding meteorological forecast.

The Diebold-Mariano test estimates the significance of the observed model predictive performance differences. When we compare the errors generated by the Linear and Expanded variants against the Baseline variant in Fig. 5, we observe a similar pattern. This pattern is reflected in Fig. 8 that shows the p-value associated to the Diebold-Mariano test (for further details, refer to Appendix B). Both comparisons exhibit a p-value of 0.071 or lower when the horizon does not exceed 4 days, and in particular for horizon days $d_h = 2$ and $d_h = 3$ the p-value does not exceed 0.002. After $d_h = 4$ the p-value increases quickly.

Fig. 9 highlights the sensitivity of the forecasting error when suboptimal regularization penalties are used by focusing on the Linear model, and showing the $\mathrm{RMSE}_{d_h}(\lambda)$ as a function of $\lambda \in \Lambda$ and $d_h \in \{1, \ldots, 7\}$. For each horizon day d_h, the

minimum reached by the surface corresponds to the expected error shown in Fig. 5, and the regularization penalty $\hat{\lambda}_{d_h}$ shown in Fig. 7. When λ is lower than $\hat{\lambda}_{d_h}$, the poor accuracy is due to overfitting. Within the tested range $\lambda \in \Lambda$, the loss in accuracy due to overfitting is more pronounced the longer the forecast horizon. Conversely, when λ is higher, the model is underfitted. Because of the scaling described in (5)–(8), when λ is sufficiently high to cause all coefficients to be zero, the forecasted price is simply the average of the past prices at that hour of the day. This is why part of the surface is characterized by a constant error.

B. Validation of the Regression Coefficients β

As discussed in Section II-B, and in more detail in Section III, we account for a variety of features. Hereinafter, we discuss the coefficients associated with these features, and highlight their temporal and geographical distribution. The regression coefficients highlight interactions that potentially depend on the period of the year. Interactions that could cause a different distribution of coefficients depending on the period of the year are, for example, the correlation between temperature and electricity demand [31] and electricity generation from PV units. Hence, it is worth highlighting that these coefficients refer to the training of the last Linear forecast generated on the 16th of April. This causes the coefficients to reflect the training period from the 22nd of January to the 15th for April 2021 (i.e. 84 days preceding d_t).

Fig. 10 shows the sum of all negative and all positive coefficients associated with a feature[10] as a function of the forecasting horizon. For example, the sum of negative coefficients associated with the meteorological attribute *Wind speed* and the 2nd day for the forecasting horizon,[11] are represented by the negative light blue bars in $d_h = 2$ ($\forall s_h \in \{0, \ldots, 23\}$). All three models are shown:

- The $\mathrm{Baseline}$ variant does by definition not include meteorological data. We see that especially coefficients related to the lagged electricity prices are dominant. A positive correlation prevails in the beginning (the higher the lagged prices, the higher the forecasted price). After approximately 6 hours, also a negative correlation emerges. It is expected to observe negative coefficients since they enable the representation of trends and seasonalities. Nevertheless, the magnitude is surprising and possibly caused by the interplay between the true underlying autoregressive process and the shrinkage of weakly correlated coefficients caused by the LASSO regularization.[12] Holiday and weekday dummies exhibit a negative correlation. This is in line with intuition, since holidays and weekends are typically characterized by lower power prices. Commodity prices play a minor role.[13]
- The Linear variant exhibits a relatively smooth reduction of coefficients over the consecutive horizon days. Autoregressive effects play a significant role only during the first horizon day, with a predominantly positive correlation. Wind speed, especially after the first horizon day, appears to be the major price driver. Like wind speed, also the temperature exhibits a negative relationship, since lower temperatures in colder periods are correlated with a higher electricity demand, especially in northern European regions [31]. It is worth nothing that irradiation has a positive relationship during nighttime. This suggests that higher PV generation at daytime tends to cause higher prices at nighttime.[14] This effect could be explained by the costs and constraints faced by inflexible thermal units. In fact, a low residual load at daytime (caused by high PV generation) causes the load gradient between daytime and nighttime to be smaller. Hence, fewer inflexible units are required to adjust their output and are, thus, willing to bid below marginal costs in order to run-though at nighttime.
- The $\mathrm{Expanded}$ variant shows more pronounced coefficients, in particular on the 2nd and 3rd horizon day. Nevertheless, observed patterns are similar to the Linear variant. It is interesting to note that mostly cubed wind speed coefficients are selected, plausibly, to resemble the polynomial shape of wind turbine power curves.

Fig. 1 shows the time span covered by the meteorological forecast in relation to the horizon of the electricity price forecast. On one hand, time interdependencies and meteorological forecast uncertainty suggest that the price at a particular hour does not only correlate with the meteorological forecast for that specific hour, but also with the forecast for hours before and after that. On the other hand, the additional explanatory power of hourly meteorological data that becomes available several days before the target price is likely to be limited. This consideration is reflected in the way meteorological data is aggregated to form the feature vector, see (15). The question about the benefit of adding forecasted meteorological data that precedes or succeeds the day of the price to be forecasted can be analyzed by exposing the distribution of meteorological coefficients as a function of the meteorological and price forecasting horizon. Fig. 11 shows positive and negative coefficients for each horizon of the price forecast d_h, grouped by the horizon of the individual meteorological attribute d_h^w. We see that for the most important meteorological attribute, wind speed, there is a strong correspondence between d_h and d_h^w. This means that the price on the day $d_t + d_h$ is represented by the model, mostly, using the meteorological forecast for that particular day $d_t + d_h^w$.

[10]For the sake of simplicity, dummy variables are grouped under the labels *Holiday* and *Weekday*.

[11]The sum does not differentiate between locations and horizons of the meteorological forecast (d_h^w, see Section III-B).

[12]E.g. imagine an autoregressive process described by many small positive coefficients and fewer larger negative coefficients. In such a configuration, a L1 regularization would tend to shrink and eliminate predominantly positive coefficients.

[13]It is worth noting that gas prices appear positively correlated predominantly in the daytime, whereas EUA prices predominantly in the nighttime. In the context of the German market, this makes sense as lignite-fired plants (that are particularly sensitive to EUA price) are more likely to be price setting during the night when demand is lower, whereas gas-fired plants are more frequently prices setting during the day.

[14]The possibility that the effect is caused by confounding variables cannot be ruled out. Reverse causation is unlikely because only a fraction of PV units react to price signals when prices are close to zero or lower.

This observation holds also true for the irradiation, whereas the remaining meteorological attributes seem to be characterized by more noise.

Besides the time dimension explored so far, the used meteorological features have also a geographical dimension represented by the locations. Fig. 12 represents this geographical distribution by summing the absolute coefficients of the Linear variant of each location and meteorological attribute across all hours of the forecasting horizon:

- *Wind speed* coefficients are concentrated in northern Germany, which is in line with the actual distribution of onshore wind farms. The proximity to the coast is also consistent with the location of offshore wind farms. Within Germany, other coefficients play a minor role. Outside Germany, wind speed is less pronounced.
- *Wind direction* coefficients appear predominantly outside Germany and might correlate with the probability distribution of future weather conditions.
- *Temperature* coefficients within Germany are close to the population-dense and industry-rich regions around the Rhine and Neckar rivers. On the other hand, the most pronounced temperature coefficients are in southern France and at the Italian-Swiss border, potentially pointing at the relationship between power prices in Germany and electric heating in these neighboring regions [32].
- *Cloud cover* and *Irradiation* are not concentrated in one particular region. Especially in the latter case, this might be caused by collinearity across different locations.

In general, it is important to highlight that the regression model per se only points at correlations. Considering the fundamentals governing the day-ahead power market, some coefficients seem to plausibly represent a causal relationship, other correlations are more ambiguous or clearly the result of model noise (such as wind speed coefficients close to the Faroe Islands).

VI. Conclusion

In this article, we showed the potential of meteorological forecasts to improve the accuracy of power price forecasting models. While forecasts for the next day can take advantage of autoregressive effects, this is less true for longer forecasting horizons. Especially between the 2nd and 4th horizon day, forecasts featuring meteorological data showed to be more accurate, as evidenced in the improvement of the RMSE by 10–20%.

We compared the accuracy of a purely linear model with a model containing also squared and cubed meteorological data. Although we found that this additional data is used by the model to capture underlying nonlinearities (such is the case for wind power curves), the forecasting accuracy of the extended model improved in one case only ($d_h = 3$). No general conclusion can thus be drawn on whether adding squared and cubed meteorological data in this particular setup is advantageous.

The importance of choosing an adequate regularization parameter has been examined on the basis of the RMSE's sensitivity to the regularization parameter. The examination highlighted how suboptimal configurations can quickly lead to underfitting or severe overfitting. We observed that the choice of the best performing regularization is a function of the forecasting horizon and that, in general, a longer forecasting horizon benefits from a more stringent regularization.

Finally, we observed that coefficients were mostly plausible when taking into consideration the fundamentals of the German day-ahead power market. Examples are the negative correlation between forecasted wind speed and power prices, as well as the concentration of wind speed coefficients in northern Germany.

The presented model variants performed well under particular circumstances, such as the forecasting horizon and the calibration of the regularization. Hence, using a forecast combination model, as advocated by Bunn [33], appears to be a promising continuation of this work. This includes the combination of models that are estimated using different calibration windows, as proposed by previous studies [24], [25]. In terms of data acquisition, two challenges allowed forecasts to be produced only from the 1st of March to the 16th of April: the limited access to historic weather forecasts and the discontinued provision of commodity futures by the selected source. Nevertheless, evaluating the behavior of coefficients over a longer period could generate insights about the explanatory power of features in different periods of the year.

Finally, the presented model was fitted only on power prices. Nevertheless, further domain knowledge could be encoded in a forecasting model by fitting also on other data, such as renewable generation, net cross-border flows and electricity load. This could be achieved by complementing the loss function of a single-stage model, or by adopting a two-stage approach, as proposed by Jonsson *et al.* [34]. Accounting for nonlinearities by applying informed feature transformations (e.g. map wind speed using a representative power curve, map the residual load forecast to an estimated merit-order curve) might lead to further improvements.

Appendix A
Notation

Sets and paramenters

$t \in \mathbb{Z}$	Integer number ($\mathbb{Z}$) representing the hourly timestamp, positioned such that 1 denotes the hour of the first forecasted price by the first forecast (1st of March, midnight-1am).
$h \in \{1, \ldots, H\}$	Hourly offset representing the horizon of the electricity price forecasts where $H = 240$.
$s_i \in \{0, \ldots, 23\}$	The hour of the day of a timestamp or of a time offset $i \in \{t, h\}$.
$d_i \in \mathbb{Z}$	The day of a timestamp or of a time offset $i \in \{t, h\}$.
$d_h^w \in \{0, \ldots, 10\}$	Special case of d_i where the horizon of the meteorological forecast is partitioned in segments that overlap with d_h, exception made for $d_h^w = 0$.
$a \in \mathbb{A}$	Meteorological attributes: wind speed and direction, solar irradiation, cloud cover, and temperature.
$l \in \mathbb{L}$	Selected weather locations across Europe, where $\mathbb{L} = 1, \ldots, L$ and $L = 100$.
$\lambda \in \Lambda$	Regularization penalty where Λ contains 100 test values for λ ranging from 0.05 to

2.00, spaced evenly on a logarithmic scale in base 10.

$s^{min}_{d^w_h} \in \{0, 11\}$ Helper set used to align the price forecast horizon with the horizon of the meteorological forecast.

$s^{max}_{d^w_h} \in \{10, 23\}$ Helper set used to align the price forecast horizon with the horizon of the meteorological forecast.

Scalars

$\hat{\lambda}_{d_t,d_h}$ Best performing λ calculated on an expanding window.

$\hat{\lambda}_{d_h}$ Best performing λ reached at the end of the experiment.

y_{d_t,s_t} Electricity price on day d_t and hour s_t.

$\hat{y}_{d_t,d_h,s_h}(\lambda)$ Electricity price (day $d_t + d_h$ and hour s_h) forecasted on day d_t, as a function of λ.

$\underline{\hat{y}}_{d_t,d_h,s_h}(\lambda)$ Standardized electricity price (day $d_t + d_h$ and hour s_h) forecasted on day d_t, as a function of λ.

$w^{a,l}_{d_t,d^w_h,s_{d^w_h}}$ Meteorological forecast published on day d_t for location l, attribute a, and horizon d^w_h and $s_{d^w_h}$.

$\bar{w}^{a,l}_{d_t,d^w_h}$ Average of the meteorological attribute a and location l on the horizon day d^w_h.

Coal_{d_t} Coal future price.

Gas_{d_t} Gas future price.

EUA_{d_t} EUA future price.

$\text{Weekday}^n_{d_t+d_h}$ Weekday dummy variable.

$\text{WeekdayC}^n_{d_t+d_h}$ Cumulated weekday dummy variable.

$\text{Holiday}_{d_t+d_h}$ Holyday dummy variable.

RMSE_{d_h,s_h} Root-mean-square error for each horizon hour.

RMSE_{d_h} Root-mean-square error calculated for the entire horizon day.

Vectors and Matrices

$\boldsymbol{x}_{d_t,d_h}$ Feature vector.

$\underline{\boldsymbol{x}}_{d_t,d_h}$ Standardized feature vector.

$\boldsymbol{x}^W_{d_t,d_h}$ Weather data vector.

$\boldsymbol{x}^D_{d_t,d_h}$ Dummy data vector.

$\boldsymbol{x}^P_{d_t}$ Lagged electricity prices vector.

$\boldsymbol{x}^C_{d_t}$ Commodity futures prices vector.

$\boldsymbol{y}_{d_t,s_h}$ Training prices.

$\underline{\boldsymbol{y}}_{d_t,s_h}$ Standardized training prices.

$\boldsymbol{y}_{d_t}$ Vector of all 24 prices on day d_t.

$\boldsymbol{y}^f_{d_t}$ Vector containing daily price aggregates for the 4 weeks preceeding d_t, where f is the function used for the aggregation (min, max, avg).

$\hat{\boldsymbol{\beta}}_{d_t,d_h,s_h}(\lambda)$ Regression coefficients as a function of λ.

$\boldsymbol{X}_{d_t,d_h}$ Training data i.e. feature matrix.

$\underline{\boldsymbol{X}}_{d_t,d_h}$ Standardized feature matrix.

$\boldsymbol{w}^{a,l}_{d_t,d^w_h}$ Meteorological forecast published on d_t for the horizon day d^w_h.

$\boldsymbol{v}^{a,l}_{d_t,d_h}$ Compound vector containing hourly meteorological forecasts as well as daily aggregates.

$\boldsymbol{W}_{d_t,d_h}$ Matrix of the compound vector $\boldsymbol{v}^{a,l}_{d_t,d_h}\ \forall a, l$.

$\mathbf{Weekday}_{d_t+d_h}$ Weekday dummy vector.

$\mathbf{WeekdayC}_{d_t+d_h}$ Cumulated weekday dummy vector.

APPENDIX B
DIEBOLD-MARIANO TEST

$L^{\text{Linear}}_{d_t,d_h,s_h}$ and $L^{\text{Baseline}}_{d_t,d_h,s_h}$ denote the squared error produces by the variants Linear and Baseline. The Diebold-Mariano test focuses on the estimation of the scaled expected difference between the errors produced by the variants $Linear$ and $Baseline$. Hence, the p-value (p) corresponding to the hypothesis that the Linear model is not more accurate than Baseline is a function of the cumulative normal distribution Φ. The estimation is analogous when other model pairs are tested.

$$\Delta^{\text{Linear}}_{d_t,d_h,s_h} = L^{\text{Linear}}_{d_t,d_h,s_h} - L^{\text{Baseline}}_{d_t,d_h,s_h} \tag{25}$$

$$\overline{\boldsymbol{\Delta}}^{\text{Linear}}_{d_h} = \frac{1}{(E+1)\cdot S}\sum_{d_t,s_h}\Delta^{\text{Linear}}_{d_t,d_h,s_h} \tag{26}$$

$$\sigma(\overline{\boldsymbol{\Delta}}^{\text{Linear}}_{d_h,s_h}) = \frac{1}{\sqrt{(E+1)\cdot S}}\cdot\frac{1}{(E+1)\cdot S-1}\cdot \cdot\sum_{d_t}\left|\Delta^{\text{Linear}}_{d_t,d_h,s_h} - \overline{\boldsymbol{\Delta}}^{\text{Linear}}_{d_h,s_h}\right| \tag{27}$$

$$\text{test}^{\text{Linear}}_{d_h} = \frac{\overline{\boldsymbol{\Delta}}^{\text{Linear}}_{d_h}}{\sigma(\overline{\boldsymbol{\Delta}}^{\text{Linear}}_{d_h})} \tag{28}$$

$$p^{\text{Linear}}_{d_h} = \Phi(1 - \text{test}^{\text{Linear}}_{d_h}) \tag{29}$$

APPENDIX C
HISTORIC PRICES

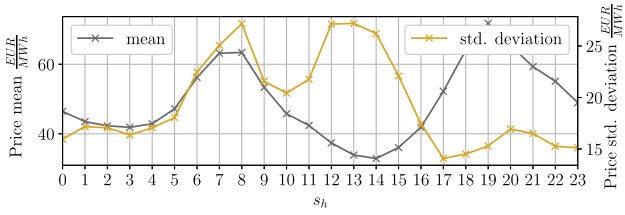

Fig. 13. Mean and standard deviation of historic German day-ahead power prices by hour of the day. Statistic spans from the 2nd of April 2021 ($d_t = 1$ and $d_h = 1$) to the 26th of April ($d_t = 47$ and $d_h = 10$).

ACKNOWLEDGMENT

The authors would like to thank Lion Hirth, Anselm Eicke, Oliver Ruhnau, Silvana Tiedemann for the insightful feedback, as well as Simone Zappalà for the inspiring discussions that shaped this work.

REFERENCES

[1] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, Oct. 2014.

[2] N. V. Karakatsani and D. W. Bunn, "Forecasting electricity prices: The impact of fundamentals and time-varying coefficients," *Int. J. Forecasting*, vol. 24, no. 4, pp. 764–785, Oct. 2008.

[3] F. Petropoulos *et al.*, "Forecasting: Theory and practice," *Int. J. Forecasting*, 2022. to be published, doi: 10.1016/j.ijforecast.2021.11.001.

[4] F. Ziel, R. Steinert, and S. Husmann, "Efficient modeling and forecasting of electricity spot prices," *Energy Econ.*, vol. 47, pp. 98–111, 2015.

[5] A. Gianfreda and D. Bunn, "A stochastic latent moment model for electricity price formation," *Oper. Res.*, vol. 66, no. 5, pp. 1189–1203, 2018.

[6] K. Maciejowska, B. Uniejewski, and T. Serafin, "PCA forecast averaging-predicting day-ahead and intraday electricity prices," *Energies*, vol. 13, no. 14, Jan. 2020, Art. no. 3530.

[7] G. Marcjasz, B. Uniejewski, and R. Weron, "Beating the Naïve-combining LASSO with Naïve Intraday electricity price forecasts," *Energies*, vol. 13, no. 7, Jan. 2020, Art. no. 1667.

[8] J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Appl. Energy*, vol. 293, 2021, Art. no. 116983.

[9] T. Matsumoto and M. Endo, "One-week-ahead electricity price forecasting using weather forecasts, and its application to arbitrage in the forward market: An empirical study of the Japan electric power exchange," *J. Energy Markets*, vol. 14, no. 3, pp. 1–26, Sep. 2021.

[10] F. Ziel and R. Steinert, "Probabilistic mid-and long-term electricity price forecasting," *Renewable Sustain. Energy Rev.*, vol. 94, pp. 251–266, 2018.

[11] R. Steinert and F. Ziel, "Short- to mid-term day-ahead electricity price forecasting using futures," *Energy J.*, vol. 40, no. 1, pp. 105–127, Jan. 2019.

[12] L. Hirth, J. Mühlenpfordt, and M. Bulkeley, "The entso-e transparency platform-A review of Europe's most ambitious electricity data platform," *Appl. Energy*, vol. 225, pp. 1054–1067, 2018.

[13] C. Sweeney, R. J. Bessa, J. Browell, and P. Pinson, "The future of forecasting for renewable energy," *Wiley Interdiscipl. Reviews: Energy Environ.*, vol. 9, no. 2, 2020, Art. no. e365.

[14] N. Ludwig, S. Feuerriegel, and D. Neumann, "Putting Big Data analytics to work: Feature selection for forecasting electricity prices using the lasso and random forests," *J. Decis. Syst.*, vol. 24, no. 1, pp. 19–36, 2015.

[15] F. Ziel and R. Weron, "Day-ahead electricity price forecasting with high-dimensional structures: Univariate vs. multivariate modeling frameworks," *Energy Econ.*, vol. 70, pp. 396–420, Feb. 2018.

[16] A. Brusaferri, L. Fagiano, M. Matteucci, and A. Vitali, "Day ahead electricity price forecast by narx model with lasso based features selection," in *Proc. IEEE 17th Int. Conf. Ind. Informat.*, 2019, pp. 1051–1056.

[17] C. Kath and F. Ziel, "Conformal prediction interval estimation and applications to day-ahead and intraday power markets," *Int. J. Forecasting*, vol. 37, no. 2, pp. 777–799, 2021.

[18] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.

[19] B. Uniejewski, R. Weron, and F. Ziel, "Variance stabilizing transformations for electricity spot price forecasting," *IEEE Tran. Power Syst.*, vol. 33, no. 2, pp. 2219–2229, Mar. 2018.

[20] ENTSO-E, "ENTSO-E Transparency Platform," Accessed: Apr. 26, 2021. [Online]. Available: https://transparency.entsoe.eu/

[21] Deutscher Wetterdienst, "Wetter und Klima - Deutscher Wetterdienst - our services - model output statistics-MIX (MOSMIX)," Accessed: Apr. 16, 2021. [Online]. Available: https://www.dwd.de/EN/ourservices/met_application_mosmix/met_application_mosmix.html

[22] Quandl, "Quandl,"Accessed: Apr. 16, 2021. [Online]. Available: https://www.quandl.com

[23] F. Ziel, C. Croonenbroeck, and D. Ambach, "Forecasting wind power-modeling periodic and non-linear effects under conditional heteroscedasticity," *Appl. Energy*, vol. 177, pp. 285–297, 2016.

[24] K. Hubicka, G. Marcjasz, and R. Weron, "A note on averaging day-ahead electricity price forecasts across calibration windows," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 321–323, Jan. 2019.

[25] G. Marcjasz, T. Serafin, and R. Weron, "Selection of calibration windows for day-ahead electricity price forecasting," *Energies*, vol. 11, Sep. 2018, Art. no. 2364.

[26] C. Fezzi and L. Mosetti, "Size matters: Estimation sample length and electricity price forecasting accuracy," *Energy J.*, vol. 41, no. 4, pp. 231–254, 2020.

[27] B. Uniejewski, J. Nowotarski, and R. Weron, "Automated variable selection and shrinkage for day-ahead electricity price forecasting," *Energies*, vol. 9, no. 8, Aug. 2016, Art. no. 621.

[28] F. Ziel, "Forecasting electricity spot prices using lasso: On capturing the autoregressive intraday structure," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4977–4987, Nov. 2016.

[29] B. Uniejewski and R. Weron, "Efficient forecasting of electricity spot prices with expert and LASSO models," *Energies*, vol. 11, no. 8, Aug. 2018, Art. no. 2039.

[30] B. Uniejewski, G. Marcjasz, and R. Weron, "Understanding intraday electricity markets: Variable selection and very short-term price forecasting using LASSO," *Int. J. Forecasting*, vol. 35, no. 4, pp. 1533–1547, 2019.

[31] T. Gallo Cassarino, E. Sharp, and M. Barrett, "The impact of social and weather drivers on the historical electricity demand in Europe," *Appl. Energy*, vol. 229, pp. 176–185, Nov. 2018.

[32] F. Lombardi, M. V. Rocco, L. Belussi, L. Danza, C. Magni, and E. Colombo, "Weather-induced variability of country-scale space heating demand under different refurbishment scenarios for residential buildings," *Energy*, vol. 239, Jan. 2022, Art. no. 122152.

[33] D. Bunn, "Forecasting loads and prices in competitive power markets," *Proc. IEEE*, vol. 88, no. 2, pp. 163–169, Feb. 2000.

[34] T. Jonsson, P. Pinson, H. A. Nielsen, H. Madsen, and T. S. Nielsen, "Forecasting electricity spot prices accounting for wind power predictions," *IEEE Trans. Sustain. Energy*, vol. 4, no. 1, pp. 210–218, Jan. 2013.

Raffaele Sgarlato received the bachelor's degree from the University of Catania, Catania, Italy, in 2012, and the master's degree in industrial engineering and management from the Technical University of Berlin, Berlin, Germany, in 2016. He is currently a Ph.D. candidate and Research Associate for the Sustainable Energy Transitions Laboratory project (SENTINEL) with the Hertie School, Berlin, Germany. During his studies, he was with Energy Brainpool, MVV Energie and Hella New Zealand. His research interests include energy economics and, in particular, power system modeling. Raffaele worked for three years with Aurora Energy Research in Berlin, most recently as technical lead.

Florian Ziel received the M.Sc. degree in statistics from University College Dublin, Dublin, Ireland, in 2012, the Diploma in mathematics from the Dresden University of Technology, Dresden, Germany, in 2013 and the Ph.D. degree in forecasting energy markets from the European-University Viadrina, Frankfurt, Germany, in 2016. He is currently an Assistant Professor of environmental economics with the House of Energy Markets and Finance, University of Duisburg-Essen, Duisburg, Germany. He is the first author of various peer-reviewed journal articles, most notably in top-tier IEEE TRANSACTIONS ON POWER SYSTEMS, *Applied Energy*, *Energy Economics*, *Renewable and Sustainable Energy Reviews* and *International Journal of Forecasting*. His research focuses on data analytics with application to energy markets and systems.