energies

MDPI

Article

Electricity Day-Ahead Market Conditions and Their Effect on the Different Supervised Algorithms for Market Price Forecasting †

Stylianos Loizidis [1,‡], Georgios Konstantinidis [1,‡], Spyros Theocharides [1,*,‡], Andreas Kyprianou [2,*,‡] and George E. Georghiou [1,*,‡]

1 PV Technology Laboratory, FOSS Research Centre for Sustainable Energy, Department of Electrical and Computer Engineering, University of Cyprus, Nicosia 2109, Cyprus; loizidis.stylianos@ucy.ac.cy (S.L.); konstantinidis.georgios@ucy.ac.cy (G.K.)

2 PV Technology Laboratory, FOSS Research Centre for Sustainable Energy, Department of Mechanical and Manufacturing Engineering, University of Cyprus, Nicosia 2109, Cyprus

* Correspondence: theocharidis.spyros@ucy.ac.cy (S.T.); kyprianou.andreas.1@ucy.ac.cy (A.K.); geg@ucy.ac.cy (G.E.G.)

† This paper is an extended version of the paper published in 2022, 2nd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED), Thessaloniki, Greece, 17–19 October 2022; pp. 1–6. 978-1-6654-6107-8/22/$31.00 © 2022 IEEE.

‡ These authors contributed equally to this work.

Abstract: Participants in deregulated electricity markets face risks from price volatility due to various factors, including fuel prices, renewable energy production, electricity demand, and crises such as COVID-19 and energy-related issues. Price forecasting is used to mitigate risk in markets trading goods which have high price volatility. Forecasting in electricity markets is difficult and challenging as volatility is attributed to many unpredictable factors. This work studies and reports the performance both in terms of forecasting error and of computational time of forecasting algorithms that are based on Extreme Learning Machine, Artificial Neural Network, XGBoost and random forest. All these machine learning techniques are combined with the Bootstrap technique of creating new samples from the available ones in order to improve the forecasting errors. In order to assess the performance of these methodologies, the Day-Ahead market prices are divided into three classes, namely normal, extremely high and negative, and these algorithms are subsequently used to provide forecasts for the whole year 2020 of the German and Finnish Day-Ahead markets. The average yearly forecasting errors along with the computation time required by each methodology are reported. The findings indicate that the random forest algorithm performs best for the normal and extremely high price categories, while XGBoost demonstrates better results for the negative price category. The methodology based on Extreme Learning Machine requires the least computational time and achieves forecasting errors that are comparable to the best-performing methods.

Keywords: energy market; market conditions; production; demand; Day-Ahead forecasting; extreme learning machine; XGBoost; Random forest

Citation: Loizidis, S.; Konstantinidis, G.; Theocharides, S.; Kyprianou, A.; Georghiou, G.E. Electricity Day-Ahead Market Conditions and Their Effect on the Different Supervised Algorithms for Market Price Forecasting. *Energies* **2023**, *16*, 4617. https://doi.org/10.3390/en16124617

Academic Editor: Yuji Yamada

Received: 1 May 2023
Revised: 1 June 2023
Accepted: 7 June 2023
Published: 9 June 2023

1. Introduction

Towards the end of the last century, a transformation of the power sector business structure from a single, and often state owned company of generating, transmitting and distributing electricity to a market oriented structure has been initiated. This means that companies compete to sell electricity in response to demand. Moreover, with the advent of the technology of harvesting solar and wind energy from renewable sources and that of controlling demand facilitated the development of new market niches in which new business opportunities appeared. These market niches were evolved to a nexus of electricity markets operating in different time horizons. The overall system of markets in Europe consists of (a) future markets, with time-scale trading of weeks, months and years ahead, (b) the spot markets that include the day-ahead, with trading occurring for each hour of the next day, and intraday markets with continuous trading and (c) the balancing markets with trading occurring in real time. A recent review from the business research perspective can be found in [1].

The production and demand of electricity is greatly affected by weather conditions, implementation of tariff policies, policies to mitigate the effects of climate change and other unforeseeable events, such as pandemics and wars. These factors together with the lack of technology for efficient electricity storage render the electricity price in all the electricity markets volatile and thereby difficult to be forecast. A reliable market price forecasting tool is a valuable instrument that market participants may exploit to deal with market price volatility, as the availability of reliable forecasts enables better strategic planning. The unexpected existence of negative and extremely high prices renders any forecasting endeavour ever more challenging.

In this work, a comprehensive analysis is made of the conditions that affect market price forecasting and subsequently the prices are divided into three classes, namely, normal, extremely high and negative. Then, various machine learning-based prediction models are used in combination with the bootstrap method to give forecasts and are compared in terms of their performance. A novelty of this work is that these prediction models are used to forecast negative prices, which is not present in the existing literature.

Another novelty of this work is that forecasts are given for a whole year, 2020, for the German and Finnish Day-Ahead markets. This can be a benchmark for other studies because the behaviour of the price for each day of the year is captured.

In addition, in the literature the separation of prices is usually performed based on a fixed threshold. As explained in more detail in Section 3.3, this mode of separation is not very convenient. In this work, a different way of separating the prices is proposed. A final contribution of this work is the comparison of the various prediction models used in terms of computational time and performance.

Section 2 of this paper reviews the relevant to electricity market conditions literature whereas Section 3 describes the market conditions in Germany and Finland. In Section 4 a description of the prediction models used is given. The results of the comparison and the forecasting results with interpretation for each market price class can be found in Section 5. The paper concludes with Section 6.

2. Literature Review

A significant problem related to deregulated energy markets is the prediction of extremely high Market Clearing Prices. In [2], extreme price values are attributed to factors affecting the normal operation of the grid, such as device failures, to the bidding strategies of the market participants and to sudden increase in demand. Consumers can better manage risks associated with peak values if more choices for purchasing electricity are available, for example, from a centralized power pool or through bilateral contracts [3]. Extremely high prices can be stabilized by having elastic demand whereas a decrease in the expected price may be obtained, at the expense of increased volatility, by using renewable energy sources [4,5]. Demand elasticity is achieved by the demand-side management which at the same time has the effect of reducing the market price volatility.

The use of electricity demand management together with historical price data demonstrated that prediction methods must learn the actual relationships between prices and the factors affecting them [6]. To this end, neural networks and extreme learning machines were proposed in [7,8], respectively, to predict marginal prices, with forecasting Mean Absolute Error (MAE) ranging between 0 and 8 ($/MWh). These articles concluded that more detailed data concerning network structure and system operations is required for the development of simulation techniques and analytical approaches to yield lower forecasting errors. The use of predicting methodologies (ANN, ELM) is thus more suitable for cases where sufficient system operational data acquisition is not feasible [2].

Market power is defined as the ability of market actors to manipulate prices to their benefit for specific periods of time [9]. As enormous profits can be achieved by increased market prices [2] generators tend to exercise market power by changing their offer curves. Market power can be exercised through the so called economic withholding or/and physical withholding [10]. Economic withholding occurs when a producer submits an offer curve with relatively high prices compared to its marginal cost [11]. Physical withholding is when generators reduce the offers of their generation capacity in order to render a proportion of the capacity of their power plants unavailable [10]. The market risks can be effectively minimized via the design and utilization of monitoring and control mechanisms. These mechanisms require that the behaviour of market participants is constantly monitored to prevent power abuse. Specifically, the market prices of the supply curves offered by the various generators should always be representative of the reasonable expectation of their short-run marginal costs [12]. The supply market prices deemed unreasonable are replaced by the default market values. Another option for market monitoring and conditions identification is the examination of the generation and demand curves and how they both affect the clearing price. A less complicated and less time consuming method could be based on identifying signals in the demand and supply time series that indicate the possibility of occurrence of extremely high or negative prices.

Local market suppliers attempt to sell their excessive energy at the highest possible price [13]. On the other hand, buyers (consumers) in the same market are cost pruners who seek market price that is lower to the utility rate [14]. On some occasions market actors emphasize on gains obtained from prosumer models/profiles which enable users to maximize their utility via price signaling [15,16]. An ANN oriented approach to estimate the system marginal price (SMP) during weekends and public holidays was proposed in [17]. The conclusion made therein was that lower error values were obtained during Sundays due to the fact that SMP curve was less volatile as compared to that of Saturdays.

Over time, socio-economic factors, along with the global economy, have caused energy markets to undergo substantial transformations. A measure named predictive density which signals the likelihood of upward or downward trends of oil prices is given in [18]. It was also concluded that during periods of extreme volatile economic climates, such variables can be considered for MCP forecasting. In [19] it was demonstrated that forecasting methodologies that take into consideration such measures yield improved forecasts [19]. In order to further improve the forecasting results methodologies that classify the days to days with, normal, excessively high and negative prices have been proposed. These kinds of methodologies attempt to exploit the intrinsic characteristics of prices that appear in these categories. Many authors have used machine learning- and/or statistical-based methodologies to forecast normal and peak price values. To this end, the Extreme Machine Learning (ELM) was deployed to forecast normal and extremely high Day-Ahead MCP values [8,20].

Over the years, a variety of methodologies has been proposed for normal price forecasting. In [21] the asymmetric Takagi-Sugeno-Kang neuro-fuzzy model in combination with the fuzzy c-mean (FCM) data pre-processing method which classifies the patterns that may exist in the data was proposed. A two-stage methodology based on a cascaded neural network (CNN) that relies on a two-stage feature selection has been developed in [22]. In the first stage the modified relief algorithm is used to capture the relevant features, whereas in the second stage, the relevance values of the obtained features are further analysed to find the features to be used to train the network. Other types of neural networks that have been exploited are the recurrent neural networks (RNN) [23] and probabilistic neural networks (PNN) [24]. Deep learning-based algorithms such as the Lasso Estimated Auto-Regressive model and Deep Learning models were proposed in [25]. It was concluded that the Deep Learning model, in overall, could perform better than the LASSO model, but the LASSO model is suitable for short-term forecasts.

In [26] it was argued that using a model incorporating about 400 explanatory variables, a variance stabilizing transformation and a re-calibrated LASSO models gives better forecasts. An improvement of the previously mentioned method was proposed in [27], where it used the Seasonal Component Auto-Regressive (SCAR) model to decompose the electricity market price time series into trend-seasonal and a stochastic parts, and subsequently, model each one separately. It was observed that accuracy improved when the load forecasts were deseasonalized. The model was tested on Global Energy Forecasting Competition 2014 and Nord Pool data demonstrating lower weekly MAE than models proposed in other studies.

A hybrid model for accurate Day-Ahead forecasting was employed in [28]. In this paper the empirical mode decomposition filter and the maximum dependency and minimum redundancy criteria are together applied to construct features. This methodology gave lower average MAPE, as compared to other models, for both 1 h and 24 h ahead forecasting. However, the RMSE of the forecasts of the New South Wales (NSW) market prices was higher than the RMSE given by other methods (average RMSE ($/MWh): NSW market was 28.96 and for PJM market was 7.29). Another hybrid model, which consists of a multiple linear regression model, an ARIMA model and Holt-Winters model was proposed in [29].

There are cases in which extremely high or negative prices appear in the electricity markets. The need for accurate forecasts in these cases has lead many researchers to develop models or methodologies for spike forecasting, whereas, there are not reports of similar endeavours in the direction of negative price forecasting. In order to identify extremely high prices a fixed threshold ($Threshold = \mu \pm 2\sigma$, where μ and σ are the estimated mean and standard deviation of observed prices for a given period) is commonly used in the literature. That is if the price exceeds the threshold it is considered as a extremely high. The techniques that over the years have been proposed are based on: clustering analysis of the market clearing values [30], probabilistic neural networks (PNN) [31] and a combination of Bayesian experts and support vector machines (SVM) [32]. The techniques proposed vary in complexity, however the key point is that the majority of them identify price time series that contain extreme values and treat them in separate clusters from the other ones. Each cluster contains its unique features that are subsequently used to implement forecasting. A different approach was presented in [33] where a two-stage feature selection methodology, based on information theory, for forecasting occurrence and spike price value was proposed. The selected features were subsequently exploited by a methodology based on a combination of PNN with Hybrid Neuro Evolutionary System (HNES) to forecast the price values.

A combination of two ANNs was used in [34] to forecast normal market prices. The first network gives forecasts for the next day and the second gives forecasts for the next week. The forecasting of extremely high prices was performed using the Generalized Pareto Distribution (GPD). The reason they separated the days to days of normal and of exceedingly high market prices was that the networks could not capture the extremely high prices even though the training set that was used was containing data of 16 years. The market price data was for the period 7 December 1998–1 January 2014 and was related to the Australian market zones. An ELM-based market price classification methodology was proposed in [35]. More specifically, the training data was classified based on thresholds, while for testing three-dimensional vectors were used. The methodology was tested on the Ontario and PJM markets and it was found found are that the classification was more accurate for the Ontario market.

A support vector machine-based method of forecasting the occurrence of the extremely high prices was proposed in [36]. Therein the extremely high market prices were defined as those that exceeded the 95th percentile, which was estimated by fitting a Generalized Pareto distribution to the innovations an AR-EGARCH model. The data that was used were the log-transformed market prices, demand and wind production. The selection of the input features was conducted by finding the optimal number of lags of the log market prices. The proposed methodology was compared to NN and XGBoost-based methodologies which were unable to accurately classify extremely high or negative prices. A hybrid methodology for forecasting both the appearance and the actual value of extremely high prices was employed in [37]. The hybrid methodology was based on the wavelet transform and on certain time domain and calendar indicators. In addition, mutual information (MI) was used for the feature selection, whereas the forecasting of the appearance of the extremely high prices was carried out by a Probabilistic Neural Network (PNN). This methodology was on data obtained from the PJM and QLD (Australia) markets. Regarding the PJM market for threshold equal to 150, the extremely high price forecast accuracy was 97.3% with a forecast confidence interval of 87.7%, while for a threshold equal to 200 was 92% and confidence interval of 88.5%. The accuracy of the corresponding measures for the QLD market for the month of June 2004 was lower. Namely, the extremely high price forecast accuracy was 88.23% and the confidence interval was 83.33% and for January 2003 92.10% and 89.74%, respectively.

The interplay between grid engineering, congestion, and market prices, as supported by studies [38,39], emphasizes the importance of diligent market design [2,38,39]. Uncertainties, strategic behaviours, network failures, and renewable price volatility contribute to extreme price risks. Examining grid constraints, competition dynamics, and implementing effective measures [40–42] is crucial to prevent manipulation and ensure fairness, efficiency, and stability. Given the current market inefficiencies, machine learning-based forecasting tools attempt to learn historical data features preceding extreme prices to forecast and address them.

Most of the methodologies proposed in the literature are machine learning-based and the following conclusions are drawn from the literature survey presented above:

(a) The price behaviour in the markets on which forecasting implemented was not discussed.

(b) The training processing time of the proposed methodologies was not reported, making it difficult to assess their applicability in real-time scenarios.

(c) There is a lack of comprehensive study assessing the performance of the algorithms for each day of the year.

(d) In cases where data is separated into different price classes, the defined threshold may lead to misclassification of negative price days as normal days.

This article addresses these issues and presents a study that explores the behavior of the German and Finnish Day-Ahead electricity markets in Section 3. It proposes a new method of threshold definition. In the results Section 5, a detailed comparison of the methodologies proposed in Section 4 is provided, considering both performance and computational time. The comparison covers all days of the year for all three price classes of market price data separation.

3. Market Conditions

3.1. Brief Description of the Nature of Electricity Markets

The role of electricity markets is to ensure that at each instant the amount of electricity physically produced must equal the amount demanded to physically be supplied. As a consequence a cluster of different markets has over the years evolved to meet this physical condition. In such clusters normally one finds the so called Day-Ahead, Intraday and Balancing markets.

In the Day-Ahead market, electricity trading takes place one day before the actual transaction. The gate for buying and selling bids closes at 12:00 the day before the actual sale and purchase occurs. After the gate closes, buying bids and selling bids for each hour are collected. The buying bids are ranked in descending order, while selling bids in ascending order. The point of intersection of the curves determines the market price (or market equilibrium). It is based on auction and the pricing is uniform. Price forecasting in the Day-Ahead electricity market is time-sensitive, requiring predictions to be made before 12:00 of the next day. This time constraint necessitates the use of historical data and forecasted values of factors that influence prices, such as electricity demand and production from renewable energy sources. By incorporating this information, accurate forecasts can be generated to assist market participants in making informed decisions.

The Intraday market was created to facilitate the penetration of renewables into the grid. At a given time of day, the intermittent character of renewable sources causes deviation between the actual electricity amount available from the amount that was traded the day before. In the Intraday market, bids are adjusted based on the available updated forecasts. The Nord Pool offers the possibility of participating in the Intraday trading which starts at 14:00 and ends one hour before the physical delivery. The trading in this market is continuous and pricing is PaB (Pay-as-Bid).

Finally, the Balancing market corrects any deviations arising from either the Day-Ahead or the Intraday market. In this market the TSO (Transmission System Operator) buys Regulation power from Regulation market participants (or Balancing generators) to cover the deviations. Usually, the Balancing market gate closes 30 min to one hour before delivery. In addition, the Balancing market is usually based on auctions. The focus of the work described in this paper is on the Day-Ahead market.

As previously mentioned, the market price in Day-Ahead markets is established at the point where production and demand reach equilibrium. However, it is important to acknowledge that the bids to buy and sell are formulated based on load forecasts, which means that the market equilibrium point may deviate from the actual demand. From a micro-economic perspective, when production exceeds demand, the market price tends to decrease or even become negative. This scenario often occurs when there is a significantly higher contribution from renewable energy sources. Conversely, when demand surpasses production, market prices tend to rise.

3.2. Day-Ahead German and Finnish Markets: Exploratory Analysis

An exploratory data description and analysis of the German and Finnish Day-Ahead markets is presented in an attempt to decipher their nature. A notable observation is that Germany is a net exporter of electricity, whereas Finland imports electricity, specifically purchasing it from Germany. In order to facilitate the understanding of these markets Figures 1–8 plot the difference between production and consumption (i.e., Production minus Consumption) against the market price ranges for years (2019–2022).

In other words, the difference shows the energy required to balance production with consumption, which can come from the Intraday market if it is a renewable unit or from the Balancing market. The role of the Balancing market in recent years has become very important due to the increase in the penetration of renewables in the markets and in combination with the technical limitations of the electricity grid or congestion points that may arise balancing becomes even more difficult. The imbalance settlements are determined based on the overall imbalance of the grid and the imbalance caused by each market participant [43].

The wholesale Day-Ahead market prices, production and consumption data for the German market were retrieved from [44], while for the Finnish market the wholesale Day-Ahead market prices were taken from the Nord Pool [45] and the production and consumption data from the Finnish TSO [46].

Figures 1 and 2 show the difference between production and consumption for the year 2019 (before COVID-19) for the German and Finnish markets, respectively. In the German market, the equilibrium between production and consumption was within the market price range of 30–60 (€/MWh). That is, in this market price range the actual production equals the actual demand. On the contrary, in the Finnish market the equilibrium point is not clear but the market price was higher compared to that of the German market, which can be attributed to the fact that Finland buys electricity from Germany. An important feature that Figure 1 portrays for the German market is that when this difference is high the market price may be extremely negative. The excess production can be sold to other countries (e.g., Finland). In particular, for differences greater than 5000 MWh profit is obtained. Another important observation is that in the Finnish market (Figure 2) the distribution of the differences over the price range 30–180 (€/MWh) is uniform.

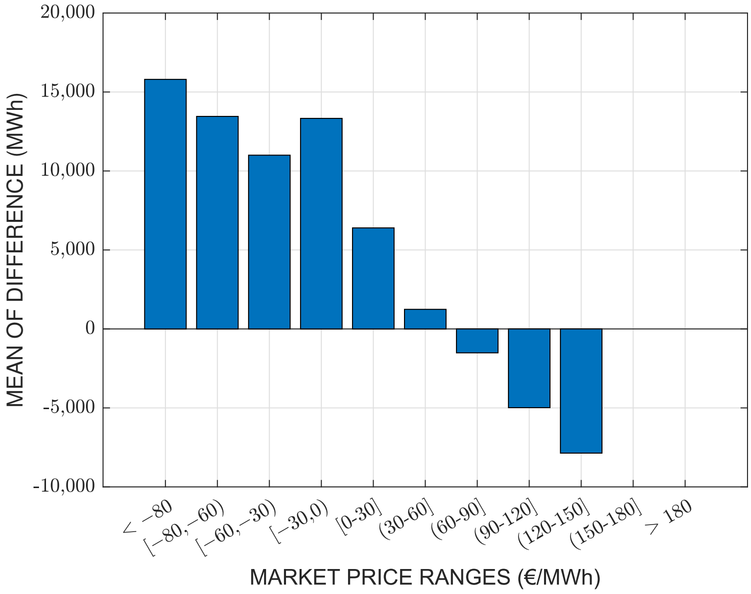

Figure 1. Mean of Difference for German Day-Ahead Market, 2019.

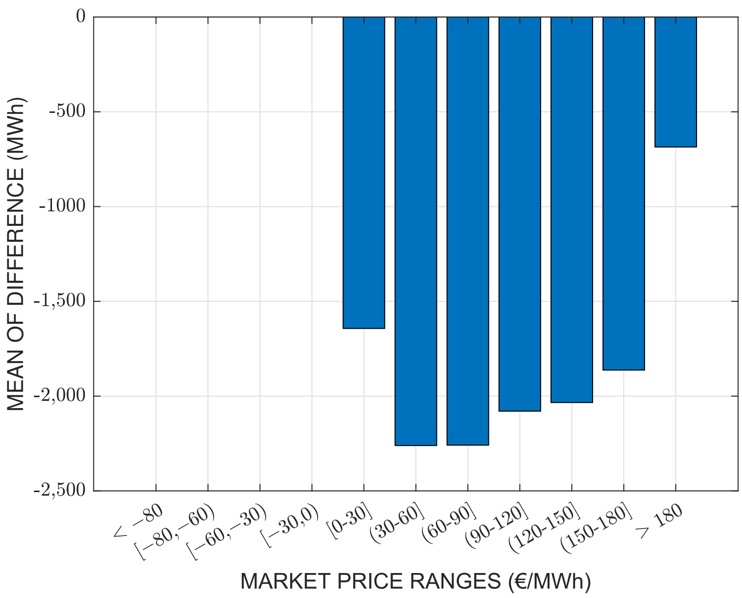

Figure 2. Mean of Difference for Finnish Day-Ahead Market, 2019.

The distributions over the market prices of the difference between production and consumption for the year 2020, during the COVID-19 period, for the German and Finish markets are shown in Figures 3 and 4, respectively. Strict lock-downs significantly decreased the demand and thereby, market prices in both countries were lower than those of 2019. The balance between production and consumption for the German market (Figure 3) was again in the market price range of 30–60 (€/MWh), while, the Finnish market (Figure 4) presented negative market prices for the first time, due to very low demand. In the Finnish market, negative prices never existed before 2020.

In addition, from Figures 3 and 4 it can be concluded that the negative market prices in 2019 appeared when the difference was approximately 14, 000 MWh, while in 2020 at the difference of approximately 9000 MWh. In addition, the excess production in the German market was 4000 MWh lower than that of 2019. Regarding the market prices between the two countries in 2020 Finland had slightly cheaper electricity than Germany.

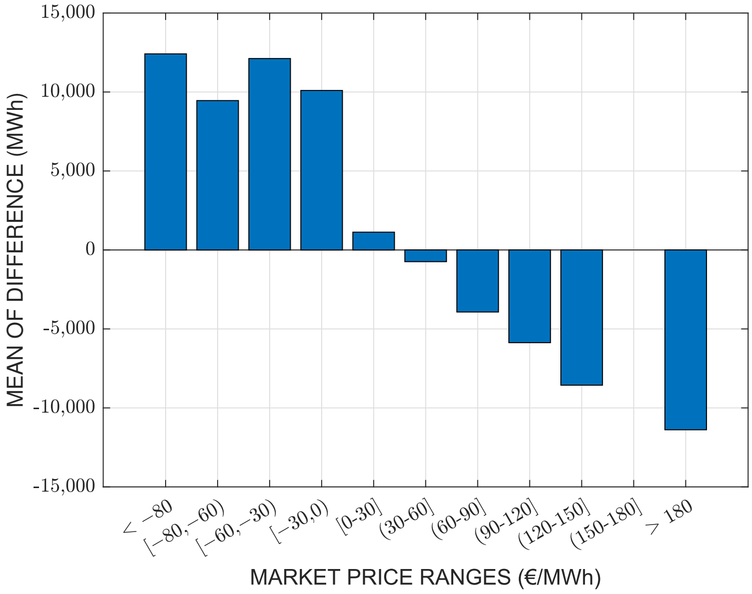

Figure 3. Mean of Difference for German Day-Ahead Market 2020.

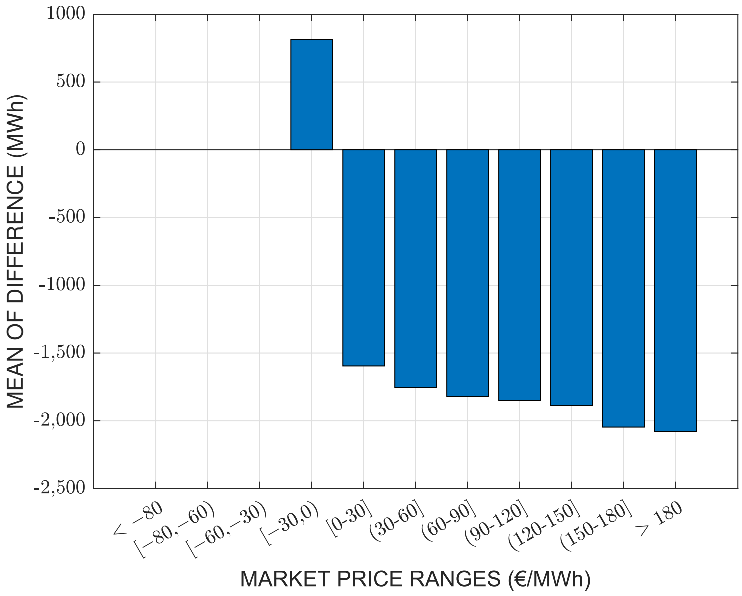

Figure 4. Mean of Difference for Finnish Day-Ahead Market, 2020.

The difference distribution over the price range for post-COVID-19 year, 2020, for the two markets are shown in Figures 5 and 6. The market price was much higher compared to 2020, especially in Germany. The reason for this was that the demand increased due to the lifting of the strict lockdowns and consequently the difference between production and consumption is lower compared to that of 2020. However, the equilibrium point is again in the range of 30–60 (€/MWh). In addition, in Germany, the production from renewables was very low during the second half of the year and the price of natural gas increased. Additionally, since 1 January 2021, the European Union has increased the price for CO_2 certificates. Therefore, the combination of low production from renewables and high costs of coal or gas units had a direct impact on the market prices.

It is evident that for the year 2021 there were instances of excess production in Germany, Figure 5, which explains the occurrence of several negative market prices. On the other hand, Finland experienced an excess of demand (Figure 6), resulting in a different market dynamic. In terms of market prices, the electricity prices in Finland were generally lower than those in Germany, as indicated by the observed price difference. This discrepancy can be attributed to the difference in production and consumption levels between the two countries. The gap between production and consumption amounted to just over 2500 MWh, leading to the import of electricity from Germany at a lower cost. The difference distributions for market prices for the year 2022 are shown in Figures 7 and 8. It can be observed that the market prices remain significantly high and this can potentially be attributed to the Energy Crisis caused by the war in Ukraine. It is noteworthy that there were fewer instances of negative market prices in Germany compared to Finland, and no occurrences of negative prices were observed in Finland during the specified period.

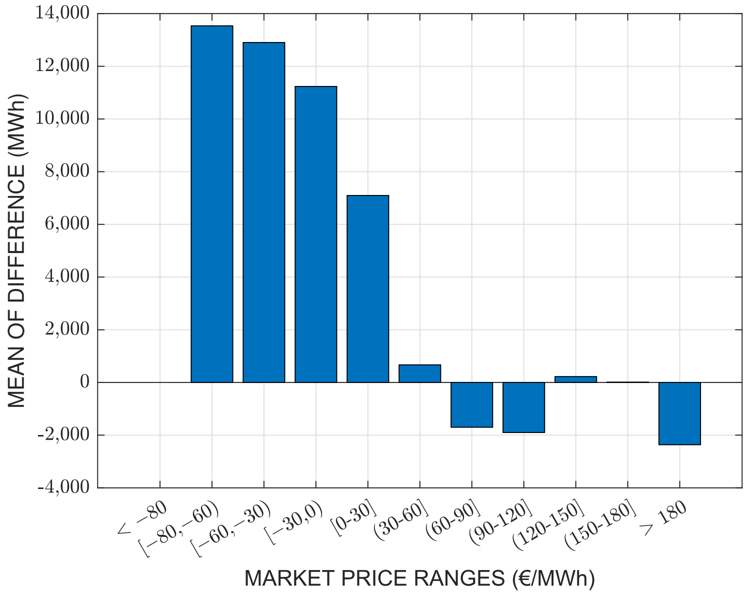

Figure 5. Mean of Difference for German Day-Ahead Market, 2021.

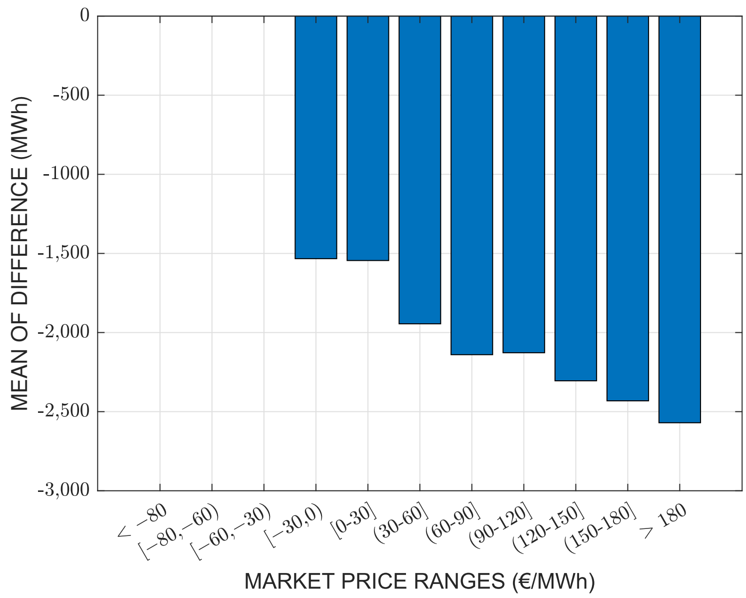

Figure 6. Mean of Difference for Finnish Day-Ahead Market, 2021.

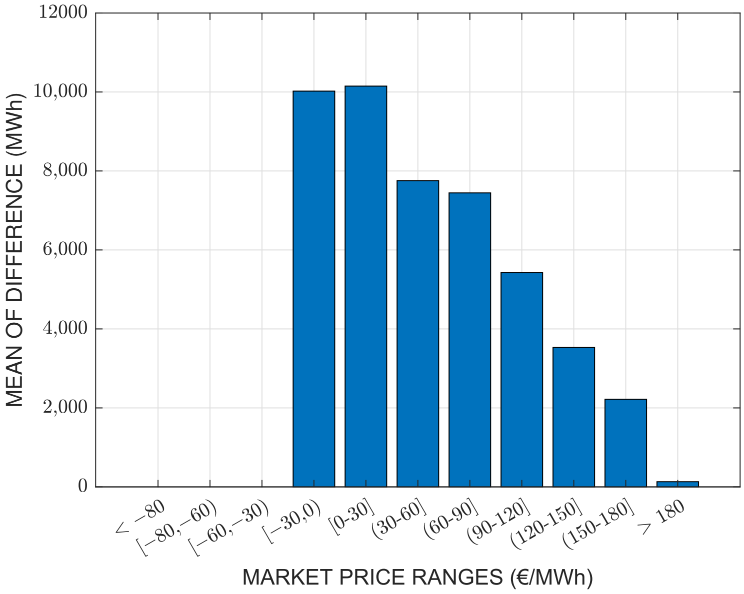

Figure 7. Mean of Difference between German Day-Ahead Market, 2022.

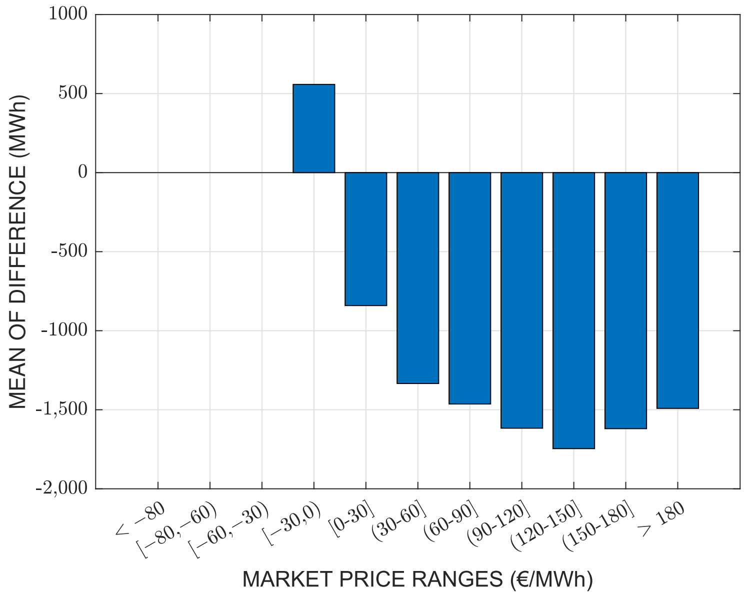

Figure 8. Mean of Difference between for Finnish Market, 2022.

3.3. Normal Market Price Range Determination

A threshold that is widely used to decide whether extremely high or low prices exist for a given period of time is given by $\mu \pm 2\sigma$ as discussed in Section 2. The prices that are outside the range $(\mu - 2\sigma, \mu + 2\sigma)$ are considered extremely low or high accordingly whereas those within this range as normal. However, there are cases that the lowest limit of this range is negative and thereby cannot be used, for negative prices are classified as normal. For example in the German market in 2022 the lower limit is -50.41 (€/MWh) and negative prices between this value and 0 (€/MWh) are considered normal.

In order to avoid cases where the range that is designated for normal prices includes negative prices the normal price range is determined by the number of price occurrences. To this end Table 1 lists the frequency of market price occurrences of some market price ranges for the German and Finnish wholesale Day-Ahead markets. The information in Table 1 is used as follows: given that the number of hours in one year is 8760, normal ranges are determined when the number of price occurrences within that range is above 1000. According to this definition the normal ranges per year listed in Table 2 were determined. Based on these ranges extremely high prices are decided to be those that exceed the upper limit of the normal price range and negative those that are lower than zero. Therefore, data is distributed into three classes. This classification facilitates the learning process and enables the methodology to generalize to new samples better.

Table 1. Frequency of wholesale Day-Ahead market price display per range.

Market Price Ranges (€/MWh)	2019	2020	2021	2022
German wholesale Day-Ahead Market				
<−80	4	3	-	-
[−80–−60)	11	18	4	-
[−60–−30)	30	28	21	-
[−30–0)	166	249	91	69
[0–30]	1700	3860	409	293
(30–60]	6501	4381	2449	222
(60–90]	326	207	2518	566
(90–120]	20	32	1254	710
(120–150]	2	4	488	694
(150–180]	-	-	415	816
>180	-	2	1088	5390
Finnish wholesale Day-Ahead Market				
<−80	-	-	-	-
[−80–−60)	-	-	-	-
[−60–−30)	-	-	-	-
[−30–0)	-	9	5	27
[0–30]	1010	5441	1748	1623
(30–60]	6570	2768	2880	844
(60–90]	1125	487	1983	834
(90–120]	36	38	1202	961
(120–150]	11	19	345	751
(150–180]	2	5	143	716
>180	5	16	454	3004

Table 2. Normal price ranges (€/MWh) determination.

2019	2020	2021	2022
German Day-Ahead Market			
$10 \leq Price \leq 75$	$7 \leq Price \leq 70$	$2 \leq Price \leq 65$	$60 \leq Price \leq 120$ or $Price > 160$
Finnish Day-Ahead Market			
$10 \leq Price \leq 70$	$1 \leq Price \leq 70$	$5 \leq Price \leq 120$	$10 \leq Price \leq 120$ or $Price > 160$

4. Supervised Algorithms for Market Price Forecasting

The methodologies proposed and compared in this work revolve around four fundamental machine learning algorithms: ELM, ANN, XGBoost, and RF. When these algorithms are individually applied to problems lacking clear cause-and-effect relationships, they typically exhibit lower performance. Given the absence of explicit cause-and-effect laws governing electricity market price behavior, the performance of these algorithms can be improved by incorporating more training samples. Bootstrapping, a statistically driven technique for generating additional samples, becomes valuable in this context. Since the number of samples can increase to the order of thousands, it is important to utilize computationally efficient machine learning algorithms. In this section, we describe the methodologies developed using the aforementioned algorithms and evaluate them in terms of both forecasting accuracy and computational performance.

The time complexity of the back propagation training algorithm is exponential meaning that the processing time increases exponentially with the increase in the number of ANN layers. In contrast, the ELM makes use of the Moore-Penrose pseudo-inverse to fit the randomly chosen weights of the hidden layer. This results to linear complexity increase. Each of these approaches, ELM or ANN has its advantages and disadvantages in terms of training time, accuracy, spatial and run-time complexity. As the ELM can be considered a variant of the ANN, a comparison can be performed between the two. The XGBoost and Random Forest machine learning-based models make use of bootstrapping, bagging and boosting and constitute a part of the methodology proposed herein. The rationale of this choice lies to the fact that their underlying mathematical formulation is similar. As a consequence their comparison provides insights on the behaviour of the various machine learning-based methods and how when used with fundamental tools such as bootstrapping and bagging the outcome is altered.

As it is mentioned above, the Bootstrap method is used in association with the machine learning-based methods. Details on Bootstrapping can be found in Section 4.5 and on training the forecasting models in Section 4.6.

4.1. Extreme Learning Machine (ELM) Model

An ELM [20] is a Single Layer Neural Network with a single hidden layer consisting of N neurons, Figure 9. Given that its activation function is $c(x)$ it learns to model Z arbitrary data samples (k_i,t_i), $k_i \in \Re^n$ using the following equation:

$$\sum_{j=1}^{h} c_j(k_j)\beta_j = \sum_{j=1}^{h} c\left(\vec{w}_j^T \cdot \vec{k} + b_j\right)\beta_j = \vec{t} \tag{1}$$

where w_j is the weight vector connecting the jth hidden neuron to the z input neurons, β_j is the weight vector connecting the jth hidden neuron with the output neuron. The bias of each hidden neuron is denoted by b_j.

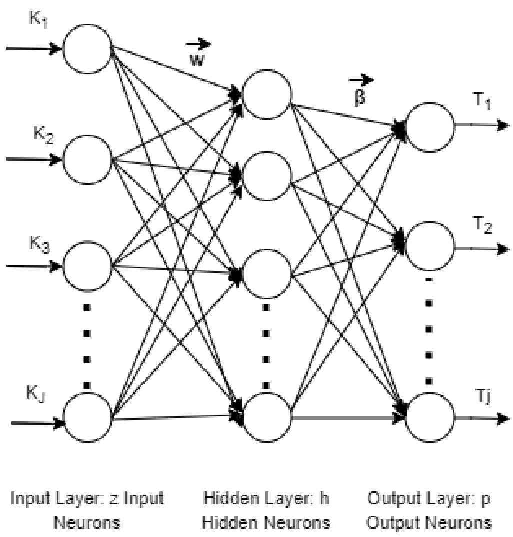

Figure 9. Extreme Learning Machine (ELM) diagram.

Equation (1) can be cast in matrix form as:

$$H\vec{\beta} = T \tag{2}$$

where H is the hidden layer matrix T is the matrix containing the target vectors $\vec{t}_i$.

The output weight vector β is estimated by solving the following equation,

$$\vec{\beta} = H^{\dagger}T \tag{3}$$

where $H^{\dagger}$ is the pseudo-inverse.

4.2. Artificial Neural Network (ANN) Model

ANNs are universal fitting machines that can be utilized as non-linear approximators for applications with available input-output data. Their structure mimics the architecture of the human brain with some simplifications. Their basic structure includes weights, neurons and activation functions. Each of the former components described above represents the actual brain's synapses and neurons. The mathematical representation of an ANN is similar to that of an ELM with a significant difference on the number of hidden layers and the training/fitting algorithm. The most commonly used algorithm for setting the weights of a Neural Network is the back-propagation method. More details about its implementation can be found in [47]. Figure 10 shows the typical structure of the ANN.

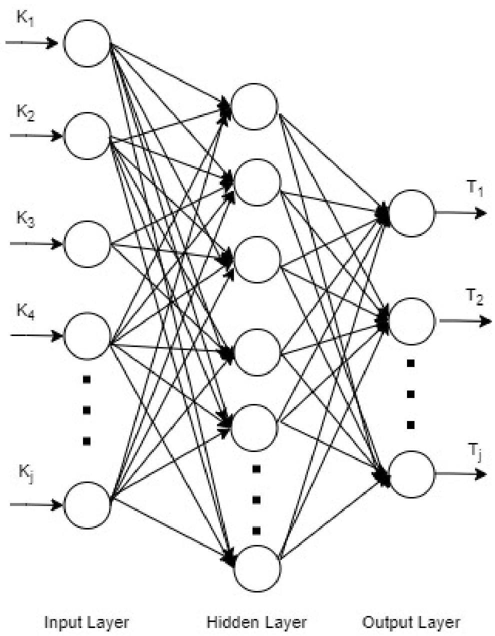

Figure 10. Artificial Neural Network (ANN) diagram.

4.3. Extreme Gradient Boosting (XGBoost) Model

The XGBoost model (see Figure 11 the typical structure) belongs to the category of boosting algorithms. Its main advantages are the very low runtime complexity and its accuracy. The XGBoost model is based on the gradient lifting decision trees of Classification and Regression type, abbreviated as (CART) in the relevant literature [48]. The model is described by:

$$\hat{y} = \sum_{i=1}^{T} f_t(x_t), f_t \in F \tag{4}$$

where T is the number of trees, f_t is specific CART tree and F all possible CART trees [48].

4.4. Random Forest (RF) Model

Random Forest (RF), Figure 12 is a supervised learning algorithm consisting of multiple decision trees trained via the bagging method. It can be utilized for classification, regression and forecasting problems. A RF has nearly the same hyper-parameters as a decision tree or a bagging classifier. A RF structure adds additional randomness to a model as the number of trees grows. It searches for the best feature among a random subset of features [49].

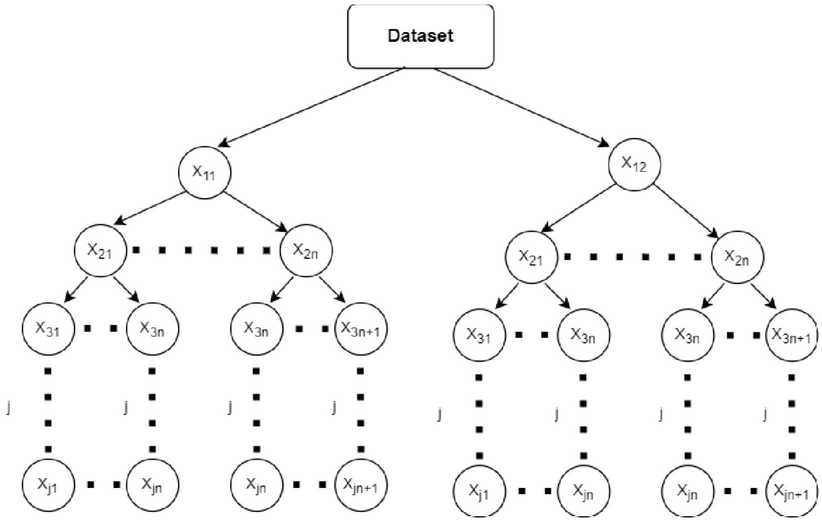

Figure 11. XGBoost diagram.

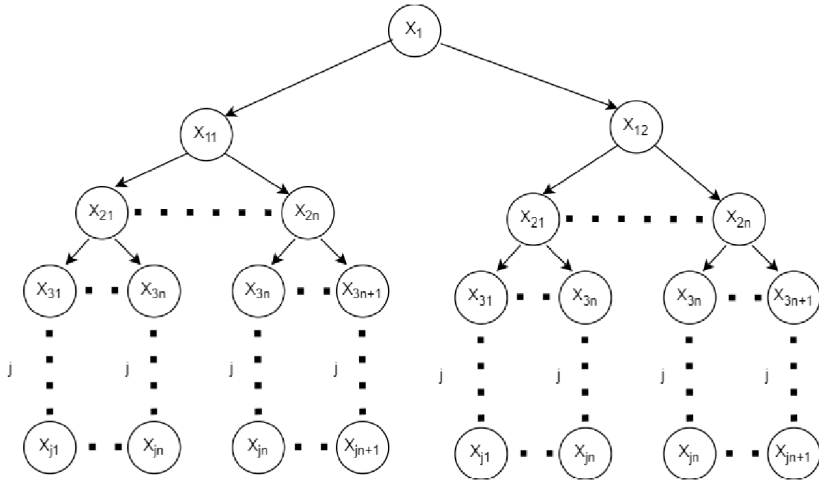

Figure 12. Random Forest diagram.

4.5. Bootstrap Method

The advantage of short learning time of the algorithms described above enables the use of statistical methods that can improve the forecasts. One such widely used method, that is based on resampling and replacement, is bootstrapping [8]. In the context of forecasting, the residuals, which are the difference between the forecasts given by an estimated model and the training data are computed as follows,

$$\epsilon_j = t_j - \hat{T}(x_j) \tag{5}$$

where $\hat{T}(k_i)$ are the forecasts obtained by the the trained model. Then, the residuals are subsequently re-centered according to,

$$\hat{\epsilon_j} = \epsilon_j - \frac{1}{\mathrm{n}} \sum_{i=1}^{p} \epsilon_i,\ j = 1,2,3,\ldots,p \tag{6}$$

and finally are re-sampled with replacement. The newly generated residuals are added to the existing forecasts to create new training data as follows

$$t_j^{\star} = \hat{T}(k_j) + \hat{\epsilon}_j \tag{7}$$

where $t_j^{\star}$ is the newly generated bootstrap data.

4.6. Description of the Training Process

The forecasting models employed in this study were used for Day-Ahead market price forecasting. The stochastic and highly erratic nature of electricity markets, renders any forecasting attempt challenging. In addition, the forecasting methodologies should be of low runtime complexity in order to facilitate the testing of multiple scenarios.

The models were trained using normalized to (0-1) data including historical Day-Ahead prices, production forecast, consumption forecast, renewable production forecast, and forecasted residual load data. The training data set was consisted of data of approximately 45 days leading up to the day of interest. This amounted to 21 training samples. The input vector contained data of 24 days while the output included the actual Day-Ahead prices of the 25th day.

The choice of activation functions, the number of hidden neurons, the number of training samples, the number of estimators, and the number of input neurons are determined by trial and error.

Regarding the training process, the first input vector, $\vec{k}_1$, included data of day I-45 up to its subsequent 23 days. The corresponding output vector, $\vec{T}_1$ included the actual Day-Ahead prices for the day I-21, where I, is the day of interest. Subsequently, the next training sample contained the data from day I-44 to day I-21, while the output vector, $\vec{T}_2$ the actual Day-Ahead prices of day I-20. This procedure was repeated, giving 21 training samples with their corresponding 21 outputs. Each training sample trained one prediction model, resulting to 21 trained models. For each of the 21 models, the bootstrap method was applied by calculating the residuals between actual and predicted prices, followed by 1000 iterations of sampling with replacement for each of the 21 models. The final forecasts were obtained by calculating the mean of the output vectors, $\vec{T}$ of the 21 models according to the equation below,

$$\hat{T} = \frac{1}{21}\sum_{j=1}^{21}\hat{T}_j \tag{8}$$

5. Results

This section presents the forecasting results. Section 5.1 presents and discusses the results of the comparison in terms of computing time and performance of the methodologies proposed in this study. Section 5.2 presents the actual results obtained for typical days from each price class, namely, normal price days, days with extremely high prices, and days with negative prices.

5.1. Comparison of the Proposed Methodologies

The prediction models were compared in terms of computational time and performance. The data used were those of German market for the year 2020. In the case of methodologies involving the ELM and ANN, the number of internal neurons was changed and tested, while for those involving the XGBoost and the RF, the number of estimators was changed. Table 3 lists the results of ELM and ANN whereas Table 4 those corresponding to the XGBoost and RF methodologies. The time reported in these tables corresponds to the duration required to implement forecasting for the entire year.

As it can be seen from the results, the ELM-based methodology required the least computational time. It took 83.40 min to implement forecasting for the entire year and this time was about 3 and 2.60 times less than the time taken by the ANN-XGBoost and RF, respectively. ELM (Extreme Learning Machine) required a total of 9.40 min to implement the forecasting for the extremely high price class whereas for the other methodologies took approximately 7 to 8 times longer to complete the same forecasting task. Finally, for the negative price class, the ELM runtime was 19.17 min which was 8, 4.70 and 4.30 times less than the implementation runtime of ANN, XGBoost and RF, respectively.

Table 3. Average yearly RMSE (€/MWh) for ELM and ANN-based methodologies.

	Normal Prices		Extremely High Prices		Negative Prices	
Internal Neurons	**Time (mm.ss)**	**Average RMSE (€/MWh)**	**Time (mm.ss)**	**Average RMSE (€/MWh)**	**Time (mm.ss)**	**Average RMSE (€/MWh)**
			ELM			
10	87.00	7.29	9.46	13.67	19.32	17.88
20	85.80	7.26	9.45	13.44	19.17	17.40
30	83.40	7.39	9.40	14.02	19.44	17.51
40	89.40	7.42	9.43	13.52	20.12	17.92
50	93.00	7.41	9.50	13.48	19.40	17.98
60	91.20	7.33	9.48	13.63	20.22	17.72
			ANN			
10	252.00	12.13	69.00	16.04	153.60	20.07
20	252.60	10.97	73.20	15.11	181.80	17.55
30	256.80	10.82	67.20	15.82	184.80	18.23
40	255.60	11.74	75.60	16.41	155.40	18.07
50	259.20	11.05	79.80	16.00	187.20	18.50
60	255.00	11.07	77.40	16.13	189.60	18.62

Table 4. Average yearly RMSE (€/MWh) for XGBoost and RF-based methodologies.

	Normal Prices		Extremely High Prices		Negative Prices	
Number of Estimators	**Time (mm.ss)**	**Average RMSE (€/MWh)**	**Time (mm.ss)**	**Average RMSE (€/MWh)**	**Time (mm.ss)**	**Average RMSE (€/MWh)**
			XGBoost			
50	266.40	15.50	70.80	22.19	90.00	21.90
100	311.40	13.01	84.60	20.57	138.60	20.08
150	331.20	12.39	95.40	17.41	154.20	19.60
200	379.80	9.10	129.60	14.77	192.20	18.77
250	423.00	7.55	148.80	13.90	214.80	17.95
300	454.20	7.12	187.20	13.32	252.00	17.05
			RF			
50	215.40	8.02	75.00	14.07	82.00	18.23
100	253.20	7.26	88.20	13.50	126.60	17.82
150	268.20	7.00	121.80	13.17	139.20	17.27
200	307.80	7.14	137.40	13.72	150.00	17.36
250	330.00	7.62	144.00	14.60	184.80	17.96
300	369.60	7.85	155.40	15.03	195.60	18.06

As it has been already mentioned, the proposed methodologies were tested for the entire year 2020 for the German Day-Ahead market. As far as the normal price class concerned, the RF gave the best result, average RMSE = 7 (€/MWh) using 150 estimators. The XGBoost and ELM-based methodologies gave close results to that of the RF. The advantage of the ELM-based methodology is that the time it requires to give these results is three and four times less that the times required by the RF and XGBoost-based methodologies. The ANN-based methodology gave the highest errors. For the extremely high price class, the RF gave a slightly better result, average RMSE = 13.17 (€/MWh) than those obtained by implementing the XGBoost and ELM-based methodologies whereas for the class of negative prices, the XGBoost-based methodology gave the best result, average RMSE = 17.05 (€/MWh). Overall the best results obtained from the XGBoost and RF-based methodologies are comparable with those obtained by the ELM-based methodology. The ELM-based methodology offers several advantages, including significantly lower runtimes and consistent errors. The ELM-based methodology offers several key advantages, including significantly lower runtimes and consistent errors, meaning that the average root mean square error (RMSE) remains relatively stable with the change of the number of internal neurons. The significantly lower runtimes of the ELM-based methodology indicate that it is more efficient in terms of computational resources and time required for implementation, an advantage that can be beneficial when dealing with large datasets and/or time-sensitive forecasting tasks.

5.2. Forecasting Results for Each Class

As it is mentioned in Section 4.6, 1000 forecasts were obtained through the bootstrap methodology, and the final actual forecasts are taken to be their mean. Note that the number of hidden of neurons used for the ELM and ANN models was 20, while the number estimated for the XGBoost and RF was 300 and 100, respectively.

5.2.1. Normal Market Price Class

Forecasting results for 6 August 2019 for the German Day-Ahead market for prediction models are shown in Figures 13–16 and the corresponding ones for 11 April 2020 for the Finnish D.A. market are shown in Figures 17–20. In all cases, the forecasted prices are close to the actual ones. The best results for both markets are given by the ELM, MAE = 1.37 (€/MWh), RMSE = 1.60 (€/MWh) for the German market, and MAE = 1.84 (€/MWh), RMSE = 2.39 (€/MWh) for the Finnish market. The ELM can capture accurately the behavior (peaks: 05:00–07:00 and 16:00–20:00) of the two markets.

The results for normal prices obtained by the ELM-based methodology, when compared to those given in the literature, exhibit lower error and short computational runtimes. Specifically, in [8] an ELM-based methodology was used with price data with load data (QLD-Australia 2002) for training demonstrated an average RMSE (for four seasons) 8.27 ($/WMh). In addition in [23] a RNN was used and tested using data obtained from different American markets gave average RMSE (four seasons) 2.13 ($/MWh). In [24] the PNN was tested to give forecasts for some days of 2002 with reported RMSEs ($/MWh) 2.80, 2.69 and 2.54, while RMSEs in [28] were 7.30 ($/MWh) for PJM market, 2018. Comparing results from different studies in the literature can indeed be challenging due to variations in the datasets used and the methodologies employed. According to [25], the lack of clear information about how the data is separated into training and testing sets in some studies can lead to misrepresentations of the actual price behavior over the course of the year.

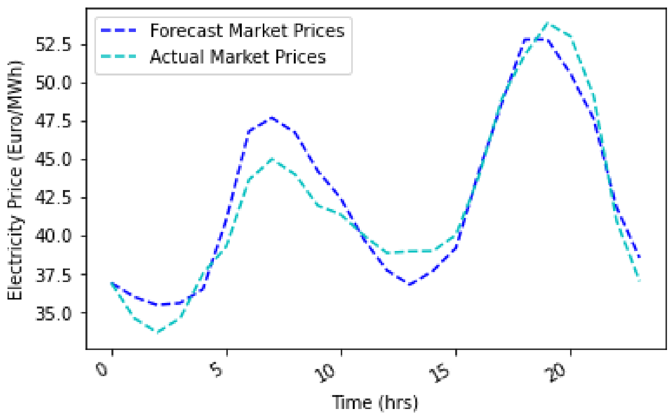

Figure 13. 6 August 2019 (normal prices), German D.A. market, ELM model: MAE = 1.37 (€/MWh), RMSE = 1.60 (€/MWh).

5.2.2. Extremely High Market Price Class

Forecasting results for 19 November 2019 for the German D.A. market are shown in Figures 21–24 and the corresponding ones for 16 February 2021 for the Finnish D.A. market are shown in Figures 25–28. In general, the results given by the prediction models are higher compared to those of the normal market prices. The better result for both markets is given by the RF, MAE = 7.83 (€/MWh), RMSE = 8.19 (€/MWh) for the German market and MAE = 7.36 (€/MWh), RMSE = 8.09 (€/MWh) for the Finnish market. In [50] where the ACH model is used on different data set corresponding to the Australian market zones.

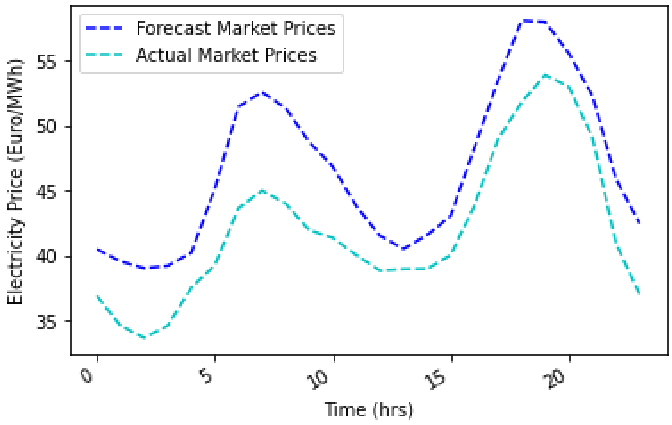

Figure 14. 6 August 2019 (normal prices), German D.A. market, ANN model: MAE = 4.63 (€/MWh), RMSE = 4.93 (€/MWh).

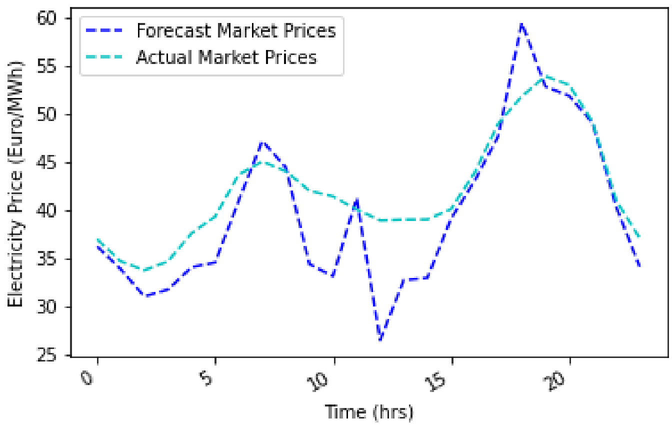

Figure 15. 6 August 2019 (normal prices), German D.A. market, XGBoost model: MAE = 3.31 (€/MWh), RMSE = 4.56 (€/MWh).

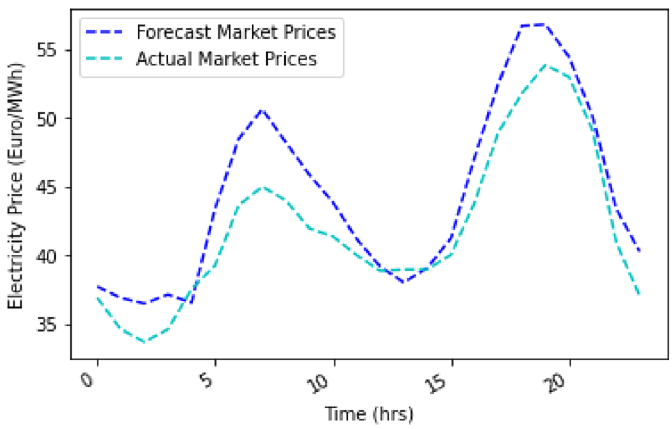

Figure 16. 6 August 2019 (normal prices), German D.A. market, RF model: MAE = 1.93 (€/MWh), RMSE = 2.39 (€/MWh).

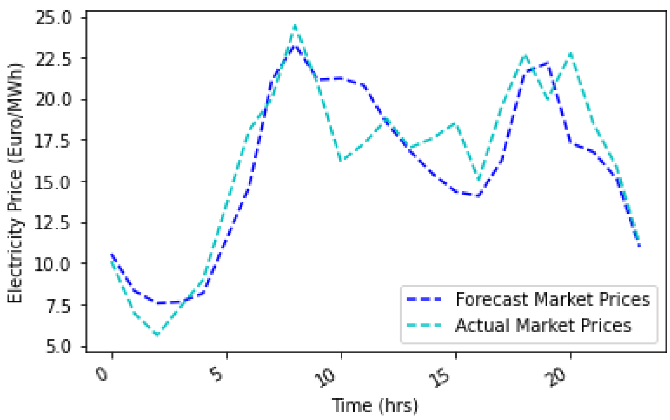

Figure 17. 11 April 2020 (normal prices), Finnish D.A. market, ELM model: MAE = 1.84 (€/MWh), RMSE = 2.39 (€/MWh).

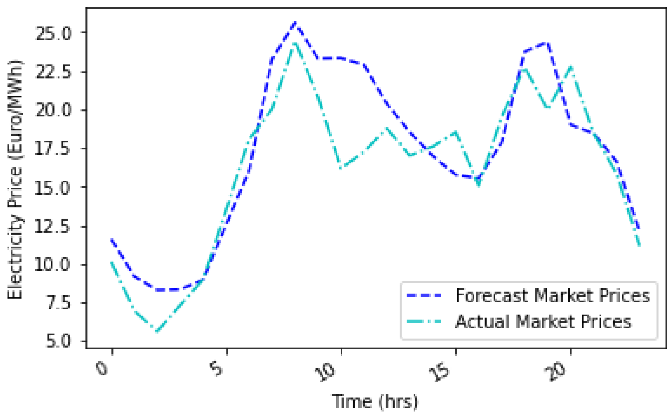

Figure 18. 11 April 2020 (normal prices), Finnish D.A. market, ANN model: MAE = 2.06 (€/MWh), RMSE = 2.70 (€/MWh).

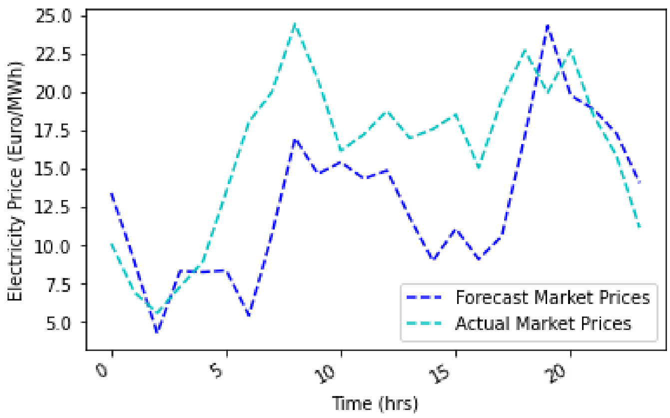

Figure 19. 11 April 2020 (normal prices), Finnish D.A. market, XGBoost model: MAE = 4.60 (€/MWh), RMSE = 5.33 (€/MWh).

In addition, none of the methodologies could capture the exceedingly high price that appears at 19:00 in the German market, however, the rest of the forecast prices are close to the actual ones. With regards to the Finnish market, the RF captured the peak price that occurs between 07:00 and 10:00 in contrast to the other methodologies that give higher forecasted prices.

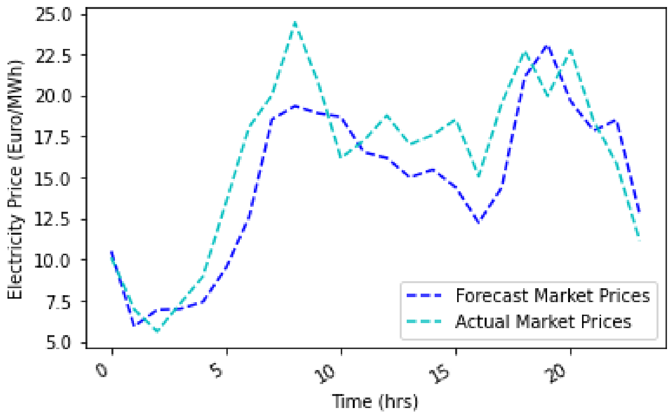

Figure 20. 11 April 2020 (normal prices), Finnish D.A. market, ELM model: MAE = 2.39 (€/MWh), RMSE = 3.04 (€/MWh).

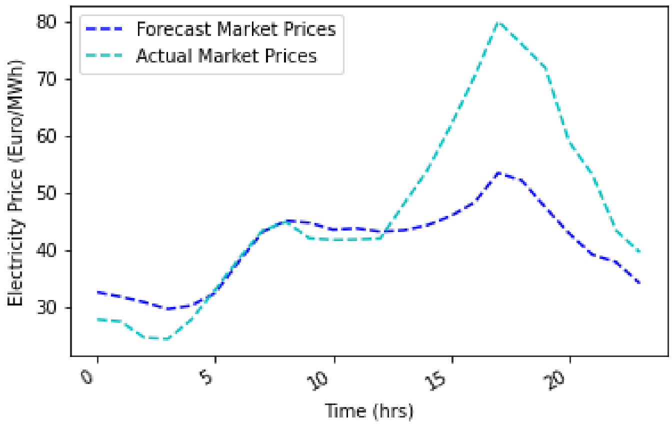

Figure 21. 19 November 2019 (extremely high prices), German D.A. market, ELM model: MAE = 8.36 (€/MWh), RMSE = 9.52 (€/MWh).

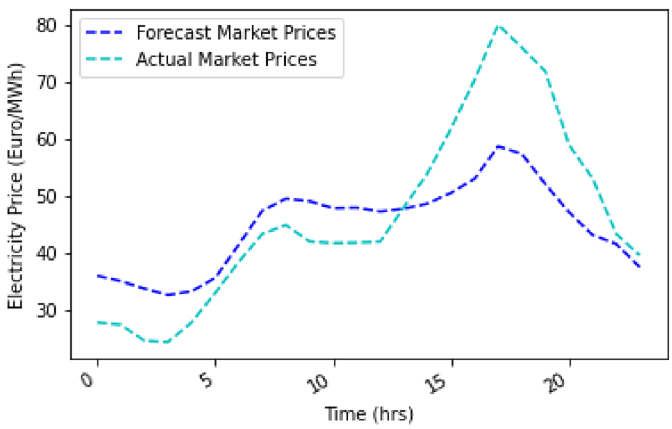

Figure 22. 19 November 2019 (extremely high prices), German D.A. market, ANN model: MAE = 8.23 (€/MWh), RMSE = 9.17 (€/MWh).

5.2.3. Negative Market Price Class

Forecasting results for 24 December 2019 for the German D.A. market are shown in Figures 29–32 and the corresponding ones for 6 July 2020 for the Finnish D.A. market are shown in Figures 33–36. Based on the results, the forecasting models do not accurately predict the occurrence of negative market prices. These prices are relatively rare and unpredictable, which renders the training of forecasting models insufficient. As a consequence, the errors tend to be higher when compared to the forecasting of other price classes.

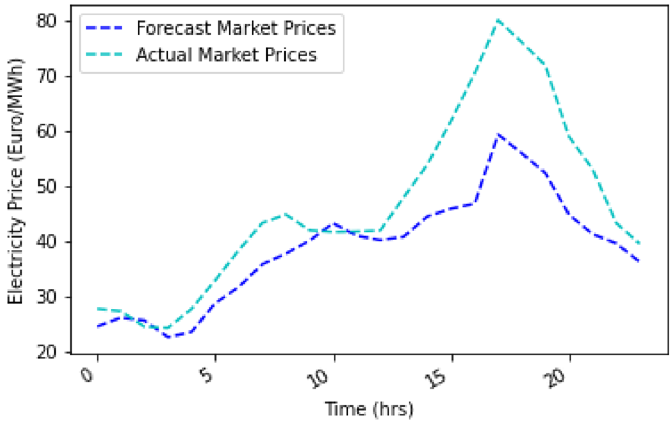

Figure 23. 19 November 2019 (extremely high prices), German D.A. market, XGBoost model: MAE = 8.00 (€/MWh), RMSE = 8.71 (€/MWh).

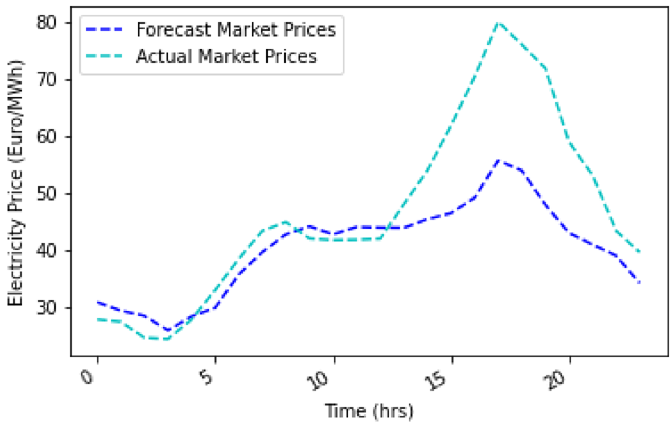

Figure 24. 19 November 2019 (extremely high prices), German D.A. market, RF model: MAE = 7.83 (€/MWh), RMSE = 8.19 (€/MWh).

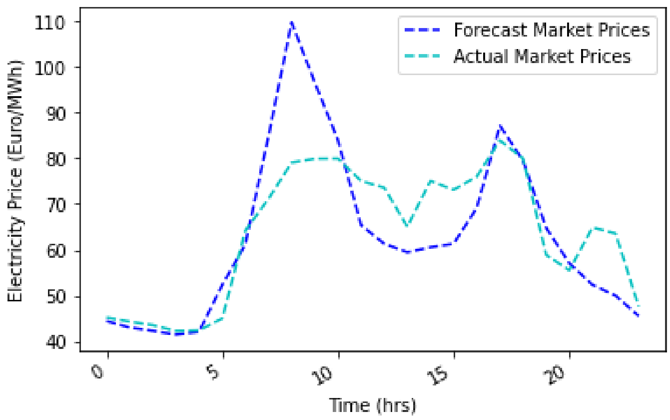

Figure 25. 16 February 2021 (extremely high prices), Finnish D.A. market, ELM model: MAE = 7.49 (€/MWh), RMSE = 8.23 (€/MWh).

For the German market, the RF gave the best results, MAE = 7.88 (€/MWh) and RMSE = 8.23 (€/MWh) while for the Finnish market, the ELM gave the best results, MAE = 17.88 (€/MWh), RMSE = 20.03 (€/MWh). An exception is the XGBoost (Figure 31) which gave negative prices in the time period 03:00–06:00.

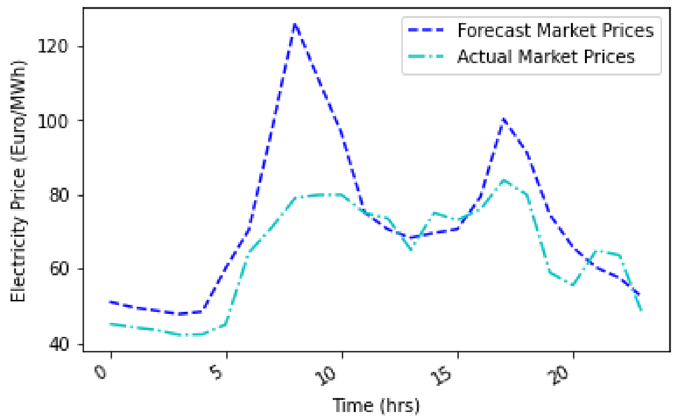

Figure 26. 16 February 2021 (extremely high prices), Finnish D.A. market, ANN model: MAE = 10.71 (€/MWh), RMSE = 11.87 (€/MWh).

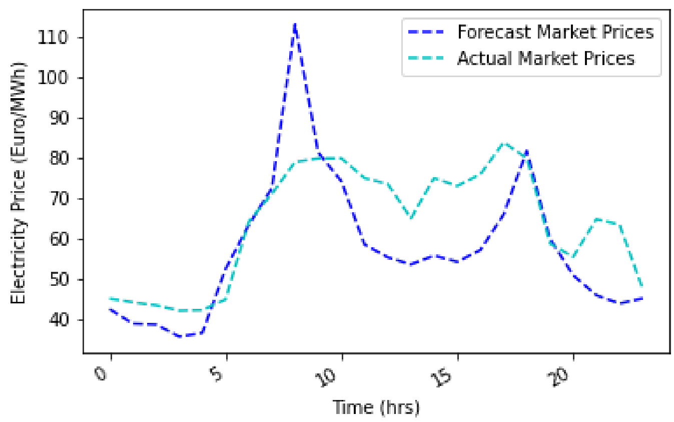

Figure 27. 16 February 2021 (extremely high prices), Finnish D.A. market, XGBoost model: MAE = 10.20 (€/MWh), RMSE = 11.55 (€/MWh).

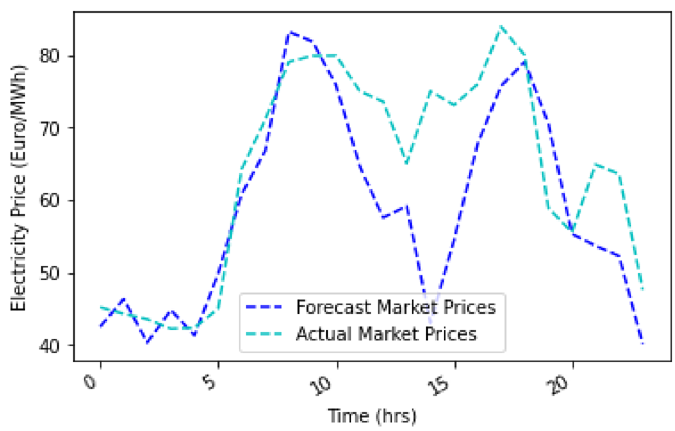

Figure 28. 16 February 2021 (extremely high prices), Finnish D.A. market, RF model: MAE = 7.36 (€/MWh), RMSE = 8.09 (€/MWh).

The appearance of negative prices is limited and thereby their forecasting using historical data becomes challenging due to the lack of satisfactory amount of training data. The ELM-based methodology has the potential to be generalized to forecasting negative values, given that sufficient training information is provided. The number of training samples should at least be equal to the number of hidden neurons.

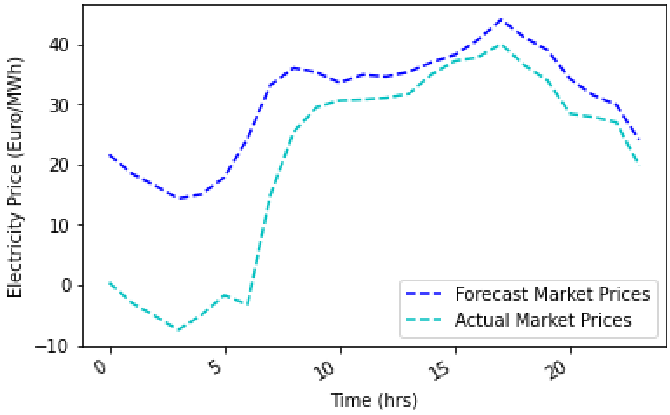

Figure 29. 24 December 2019 (negative prices), German D.A. market, ELM model: MAE = 9.91 (€/MWh), RMSE = 10.44 (€/MWh).

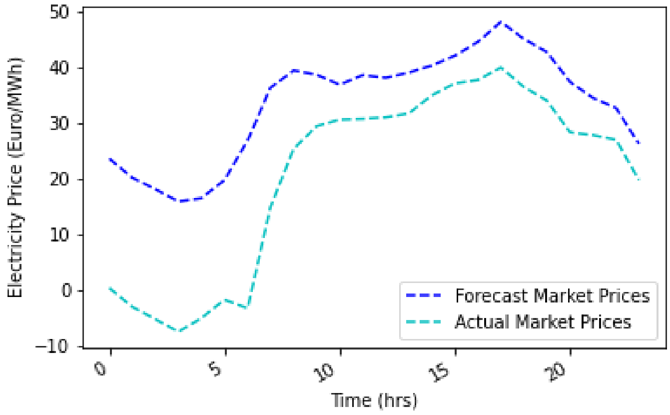

Figure 30. 24 December 2019 (negative prices), German D.A. market, ANN model: MAE = 12.93 (€/MWh), RMSE = 13.60 (€/MWh).

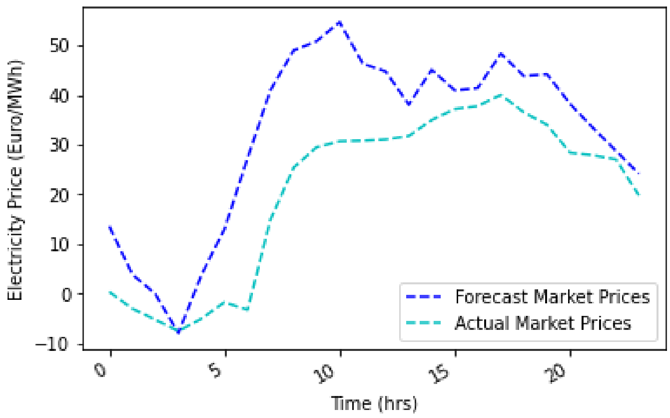

Figure 31. 24 December 2019 (negative prices), German D.A. market, XGBoost model: MAE = 11.45 (€/MWh), RMSE = 13.02 (€/MWh).

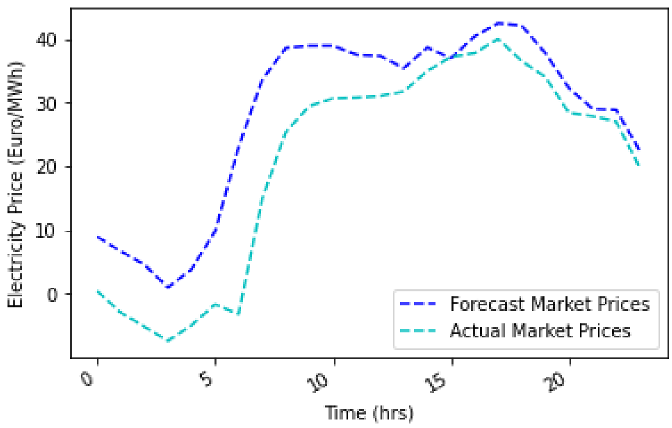

Figure 32. 24 December 2019 (negative prices), German D.A. market, ELM model: MAE = 7.88 (€/MWh), RMSE = 8.23 (€/MWh).

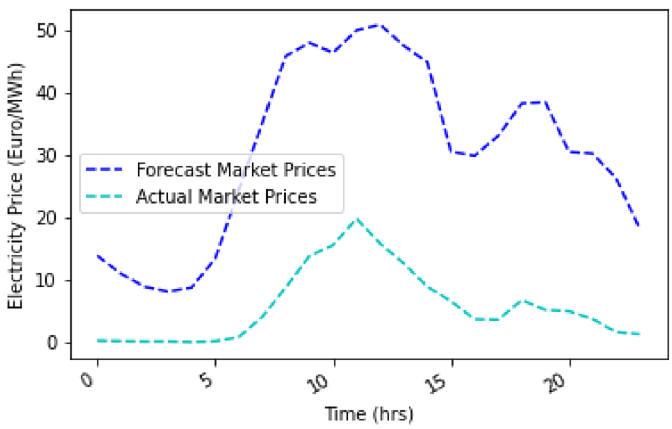

Figure 33. 6 July 2020 (negative prices), Finnish D.A. market, ELM model: MAE = 17.88 (€/MWh), RMSE = 20.03 (€/MWh).

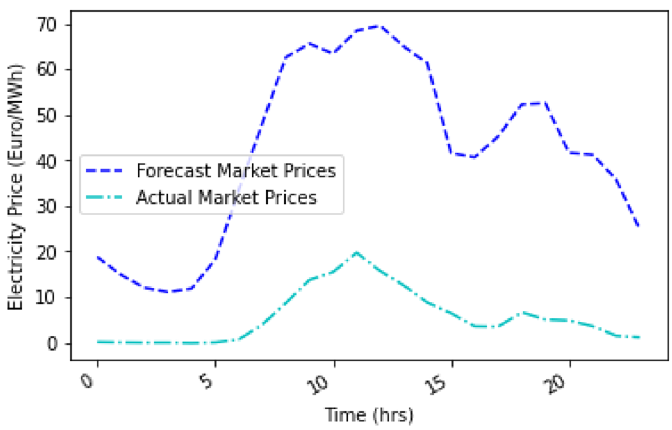

Figure 34. 6 July 2020 (negative prices), Finnish D.A. market, ANN model: MAE = 20.42 (€/MWh), RMSE = 21.02 (€/MWh).

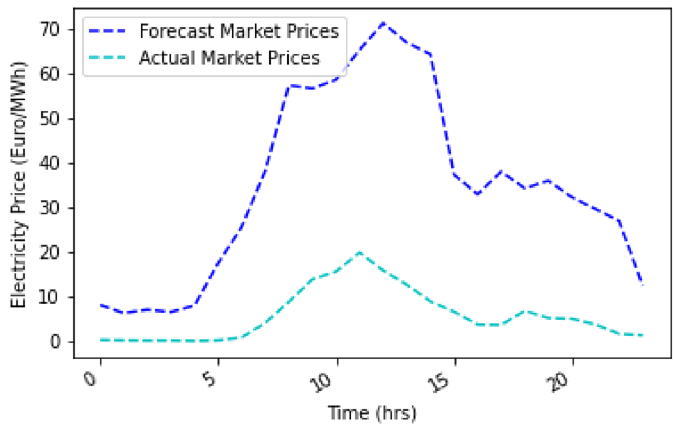

Figure 35. 6 July 2020 (negative prices), Finnish D.A. market, XGBoost model: MAE = 20.90 (€/MWh), RMSE = 21.93 (€/MWh).

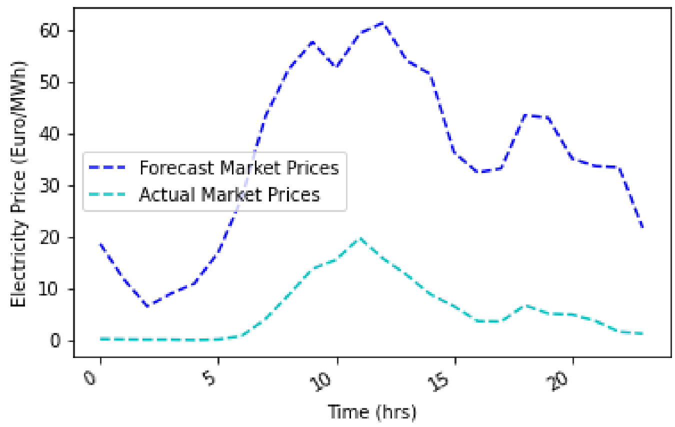

Figure 36. 6 July 2020 (negative prices), Finnish D.A. market, RF model: MAE = 21.01 (€/MWh), RMSE = 22.15 (€/MWh).

6. Conclusions

The Day-Ahead market prices are influenced by many factors and events and as a consequence highly volatile and unpredictable behavior is observed. As the exploratory study depicted in Section 3, the COVID-19 and the energy crisis have played and still play an important role in determining the market price.

As a result of the above, and amongst other factors, electricity markets and the global economy have changed their state (fuel prices, production, demand, etc.) and this has an impact on the forecasting performance of the various models. In this paper, the market prices were divided into three classes: normal, extremely high and negative market prices. The separation was based on the frequency of the appearance of the prices. That is, the range of prices that appears most often is determined to be the normal price range, the prices that were higher than the upper limit of the normal price range were deemed to be extremely high prices, while the prices that were smaller than zero constituted the class of negative prices.

The forecasting methodologies used in this work were, ELM-, ANN-, XGBoost- and RF-based. The best results were obtained for the normal price class. The ELM-based methodology had the lowest computational time, namely, for a period of an entire year it took about 90 min, to give the average RMSE. This time was about three times shorter than the time taken by the ANN and four times shorter than the time taken by the XGBoost and RF. The RF gave the lowest average RMSE, 7.00 (€/MWh) but the XGBoost and ELM gave 7.12 and 7.26 (€/MWh), respectively. All the methodologies gave MAE (RMSE) of less than 5 (5.40) (€/MWh) for both markets for a typical normal price day.

The runtime required by the ELM-based methodology to forecast prices in the extremely high price class was about 9.45 min. This runtime was about 7–7.5 times shorter than the runtimes required by the other methods. The RF gave the lowest mean RMSE, 13.17 (€/MWh) which was 1.90 times higher than the average RMSE of normal prices. In the German Day-Ahead market an extremely high price appears around 19:00 and none of the methods could accurately capture it. The errors obtained were over 8 (€/MWh). Similar observations were made for the Finnish market.

The runtime of the ELM-based methodology for negative prices class was approximately 20 min. This time was 7.5 and 4.5 times lower than the time taken by the ANN and the XGBoost-RF, respectively. The XGBoost gave the lowest average RMSE, but in general it is higher than the errors obtained for the other two classes, 2.43 and 1.29 times higher than the lowest average RMSE for the class of normal and extremely high prices, respectively.

It was demonstrated that machine learning-based methodologies give accurate forecasts on normal price days whereas for days with extremely high or/and negative prices the forecasting error increases. Based on the results presented in this study, future work on improving the performance on days with high and negative prices could follow two directions. One direction is to develop a two stage methodology of which the first stage would perform next day class classification. Then, the machine learning-based algorithms discussed earlier could be used to generate price intervals. Additionally, according to the next day class forecast, the actual price would be determined as follows: by the upper prices of the interval if the next day is classified as day with high prices, lower prices of the interval for days classified as days with negative prices. Finally, additional study must be carried out to determine the actual relationship between the current market state with the next day forecasting. For example, high positive difference between production and consumption signals lower or negative prices. These signalling instances could be rigorously defined, using notion from information theory and signalling, and exploited in forecasting both the actual time of occurrence and actual value of extremely low or negative price.

Author Contributions: Writing—original draft, S.L., G.K., S.T., A.K. and G.E.G.; Writing—review & editing, S.L., G.K., S.T., A.K. and G.E.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 863876, project title: "A novel smart grid architecture that facilitates high RES penetration through innovative markets towards efficient interaction between advanced electricity grid management and intelligent stakeholders" (project acronym: FLEXGRID).

Data Availability Statement: The production and demand data used to create the histograms of Section 3 come from the SMARD (German online resources data) and the Fingrid (Finnish TSO, Online resources data). Below in the bibliography are the corresponding urls. To create Table 1, the data from the SMARD website (German Day-Ahead prices) and the Nord Pool (Finnish Day-Ahead prices) were used. The Nord Pool's url is in the bibliography.

Conflicts of Interest: The authors declare no conflict of interest. Also, The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

DA	Day-Ahead
MCP	Market Clearing Price
SMP	System Marginal Price
PaB	Pay-as-Bid
TSO	Transmission System Operator

COVID	Coronavirus Disease
ANN	Artificial neural Network
ELM	Extreme Learning Machine
XGBoost	Extreme Gradient Boosting
RF	Random Forest
FCM	Fuzzy C-Mean
RNN	Recurrent Neural Network
SVM	Support vector Machine
PNN	Probabilistic Neural Network
HNES	Hybrid Neuro Evolutionary System
CART	Classification and Regression Type
MAE	Mean Absolute Error
RMSE	Root Mean Square Error

References

1. Bichler, M.; Buhl, H.U.; Knörr, J.; Maldonado, F.; Schott, P.; Waldherr, S.; Weibelzahl, M. Electricity Markets in a Time of Change: A Call to Arms for Business Research. *Schmalenbach J. Bus. Res.* **2022**, *74*, 77–102. [CrossRef]
2. Bao, M.; Ding, Y.; Zhou, X.; Guo, C.; Shao, C. Risk assessment and management of electricity markets: A review with suggestions. *CSEE J. Power Energy Syst.* **2021**, *7*, 1322–1333. [CrossRef]
3. Wang, P.; Billinton, R. Reliability assessment of a restructured power system using reliability network equivalent techniques. *IET* **2003**, *150*, 555–560. [CrossRef]
4. Zhao, Q.; Wang, P.; Goel, L.; Ding, Y. Impacts of renewable energy penetration on nodal price and nodal reliability in deregulated power system. In Proceedings of the 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 24–28 July 2011; pp. 1–6. [CrossRef]
5. Wang, Y.; Ding, Y. Nodal price uncertainty analysis considering random failures and elastic demand. In Proceedings of the IEEE PES Power Systems Conference and Exposition, New York, NY, USA, 10–13 October 2004; Volume 1, pp. 174–178. [CrossRef]
6. Feuerriegel, S.; Strüker, J.; Neumann, D. Reducing price uncertainty through demand side management. In Proceedings of the Thirty Third International Conference on Information Systems, Orlando, FL, USA, 16–19 December 2012; pp. 1–20.
7. Hong, Y.Y.; Hslao, C.Y. Locational marginal price forecasting in deregulated electricity markets using artificial intelligence. *IEEE Trans. Power Syst.* **2002**, *149*, 621–626. [CrossRef]
8. Chen, X.; Dong, Z.Y.; Meng, K.; Xu, Y.; Wong, K.P.; Ngan, H.W. Electricity Price Forecasting With Extreme Learning Machine and Bootstrapping. *IEEE Trans. Power Syst.* **2012**, *27*, 2055–2062. [CrossRef]
9. Wang, P.; Xiao, Y.; Ding, Y. Nodal market power assessment in electricity markets. *IEEE Trans. Power Syst.* **2004**, *19*, 1373–1379. [CrossRef]
10. Lakić, E.; Medved, T.; Zupančič, J.; Gubina, A.F. The review of market power detection tools in organised electricity markets. In Proceedings of the 2017 14th International Conference on the European Energy Market (EEM), Dresden, Germany, 6–9 June 2017; pp. 1–6. [CrossRef]
11. qiang Zhang, F.; Zhou, H. Research on Economic Withholding in Wholesale Markets Based on Incremental Heat Rate. In Proceedings of the 2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific, Dalian, China, 18 August 2005; pp. 1–7. [CrossRef]
12. Wholesale. Wholesale Electricity Market Rules. 2020. Available online: https://www.erawa.com.au/rule-change-panel/wholesaleelectricity-market-rules (accessed on 30 March 2020).
13. Yu, Z.; Razzaq, A.; Rehman, A.; Shah, A.; Jameel, K.; Mor, R.S. Disruption in global supply chain and socio-economic shocks: A lesson from COVID-19 for sustainable production and consumption. *Oper. Manag. Res.* **2022**, *15*, 233–248. [CrossRef]
14. Cali, U.; Çakir, O. Energy Policy Instruments for Distributed Ledger Technology Empowered Peer-to-Peer Local Energy Markets. *IEEE Access* **2019**, *7*, 82888–82900. [CrossRef]
15. Bampoulas, A.; Saffari, M.; Pallonetto, F.; Mangina, E.; Finn, D.P. A fundamental unified framework to quantify and characterise energy flexibility of residential buildings with multiple electrical and thermal energy systems. *Appl. Energy* **2021**, *282*, 116096. [CrossRef]
16. Tahersima, F.; Stoustrup, J.; Meybodi, S.A.; Rasmussen, H. Contribution of domestic heating systems to smart grid control. In Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA, 12–15 December 2011; pp. 3677–3681. [CrossRef]
17. Kalfa, V.R.; Arslan, B.; Ertuğrul, İ. Determining the Factors Affecting the Market Clearing Price by Using Multiple Linear Regression Method. *Alphanumeric* **2021**, *9*, 35–48. [CrossRef]
18. Baumeister, C., Korobilis, D.; Lee, T.K. Energy Markets and Global Economic Conditions. *Rev. Econ. Stat.* **2022**, *104*, 828–844. [CrossRef]
19. Halkos, G.E.; Tsirivis, A.S. Energy Commodities: A Review of Optimal Hedging Strategies. *Energies* **2019**, *12*, 3979. [CrossRef]
20. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: Theory and applications. *Neurocomputing* **2006**, *70*, 489–501. [CrossRef]

21. Alshejari, A.; Kodogiannis, V.S. Electricity price forecasting using asymmetric fuzzy neural network systems. In Proceedings of the 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Naples, Italy, 9–12 July 2017; pp. 1–6. [CrossRef]
22. Amjady, N.; Daraeepour, A. Design of input vector for day-ahead price forecasting of electricity markets. *Expert Syst. Appl.* **2009**, *36*, 12281–12294. [CrossRef]
23. Ghayekhloo, M.; Azimi, R.; Ghofrani, M.; Menhaj, M.; Shekari, E. A combination approach based on a novel data clustering method and Bayesian recurrent neural network for day-ahead price forecasting of electricity markets. *Electr. Power Syst. Res.* **2019**, *168*, 184–199. [CrossRef]
24. Lin, W.M.; Gow, H.J.; Tsai, M.T. Electricity price forecasting using Enhanced Probability Neural Network. *Energy Convers. Manag.* **2010**, *51*, 2707–2714. [CrossRef]
25. Lago, J.; Marcjasz, G.; De Schutter, B.; Weron, R. Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark. *Appl. Energy* **2021**, *293*, 116983. [CrossRef]
26. Uniejewski, B.; Weron, R. Efficient Forecasting of Electricity Spot Prices with Expert and LASSO Models. *Energies* **2018**, *11*, 2039. [CrossRef]
27. Uniejewski, B.; Marcjasz, G.; Weron, R. On the importance of the long-term seasonal component in day-ahead electricity price forecasting: Part II — Probabilistic forecasting. *Energy Econ.* **2019**, *79*, 171–182. [CrossRef]
28. Shao, Z.; Zheng, Q.; Liu, C.; Gao, S.; Wang, G.; Chu, Y. A feature extraction- and ranking-based framework for electricity spot price forecasting using a hybrid deep neural network. *Electr. Power Syst. Res.* **2021**, *200*, 107453. . [CrossRef]
29. Bissing, D.; Klein, M.T.; Chinnathambi, R.A.; Selvaraj, D.F.; Ranganathan, P. A Hybrid Regression Model for Day-Ahead Energy Price Forecasting. *IEEE Access* **2019**, *7*, 36833–36842. [CrossRef]
30. He, D.; Chen, W.P. A real-time electricity price forecasting based on the spike clustering analysis. In Proceedings of the 2016 IEEE/PES Transmission and Distribution Conference and Exposition (T&D), Dallas, TX, USA, 3–5 May 2016; pp. 1–5.
31. Wang, Y.; Li, L.; Ni, J.; Huang, S. Feature selection using tabu search with long-term memories and probabilistic neural networks. *Pattern Recognit. Lett.* **2009**, *30*, 661–670. [CrossRef]
32. Wu, W.; Zhou, J.; Mo, L.; Zhu, C. Forecasting electricity market price spikes based on bayesian expert with support vector machines. In *Advanced Data Mining and Applications, Proceedings of the International Conference on Advanced Data Mining and Applications, Xi'an, China, 14–16 August 2006*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 205–212.
33. Amjady, N.; Keynia, F. A new prediction strategy for price spike forecasting of day-ahead electricity markets. *Appl. Soft Comput.* **2011**, *11*, 4246–4256. [CrossRef]
34. Dev, P.; Martin, M.A. Using neural networks and extreme value distributions to model electricity pool prices: Evidence from the Australian National Electricity Market 1998–2013. *Energy Convers. Manag.* **2014**, *84*, 122–132. [CrossRef]
35. Shrivastava, N.A., Panigraphi, B.K.; Lim, M.H. Electricity price classification using extreme learning machines. *Neural Comput. Appl.* **2016**, *27*, 9–18. [CrossRef]
36. Stathakis, E.; Papadimitriou, T.; Gogas, P. Forecasting Price Spikes in Electricity Prices. *Rev. Econ. Anal.* **2021**, *13*, 65–87. [CrossRef]
37. 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. [CrossRef]
38. Wang, P.; Goel, L.; Ding, Y. The impact of random failures on nodal price and nodal reliability in restructured power systems. *Electr. Power Syst. Res.* **2004**, *71*, 129–134. [CrossRef]
39. Strategic bidding and rebidding in electricity markets. *Energy Econ.* **2016**, *59*, 24–36. .: 10.1016/j.eneco.2016.07.011. [CrossRef]
40. Liu, Y.; Wu, F.F. Impacts of Network Constraints on Electricity Market Equilibrium. *IEEE Trans. Power Syst.* **2007**, *22*, 126–135. [CrossRef]
41. Chattopadhyay, D. Multicommodity spatial Cournot model for generator bidding analysis. *IEEE Trans. Power Syst.* **2004**, *19*, 267–275. [CrossRef]
42. Peng, T.; Tomsovic, K. Congestion influence on bidding strategies in an electricity market. *IEEE Trans. Power Syst.* **2003**, *18*, 1054–1061. [CrossRef]
43. Morales, J.M.; Conejo, A.J.; Madsen, H.; Pinson, P.; Zugno, M. *Integrating Renewables in Electricity Markets*; Springer: New York, NY, USA, 2014; Volume 205.
44. SMARD. German Market Data. Available online: https://www.smard.de/en/downloadcenter/download-market-data (accessed on 30 March 2020).
45. Nord Pool. Finnish Day-Ahead Market Prices. Available online: https://www.nordpoolgroup.com/en/Market-data1/Dayahead/Area-Prices/ALL1/Hourly/?view=table (accessed on 30 March 2020).
46. Fingrid. Finnish Market Data. Available online: https://data.fingrid.fi/open-data-forms/search/en/ (accessed on 30 March 2020).
47. Kubat, M. Neural networks: A comprehensive foundation by Simon Haykin, Macmillan, 1994, ISBN 0-02-352781-7. *Knowl. Eng. Rev.* **1999**, *13*, 409–412. [CrossRef]
48. Wang, Y.; Guo, Y. Forecasting method of stock market volatility in time series data based on mixed model of ARIMA and XGBoost. *China Commun.* **2020**, *17*, 205–221. [CrossRef]

49. Pavlov, Y.L. Random forests. In *Random Forests*; De Gruyter: Berlin, Germany, 2019.
50. Christensen, T.M.; Hurn, A.S.; Lindsay, K.A. Forecasting spikes in electricity prices. *Int. J. Forecast.* **2012**, *28*, 400–411. [CrossRef]