Applied Energy 363 (2024) 123058

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Electricity market price forecasting using ELM and Bootstrap analysis: A case study of the German and Finnish Day-Ahead markets

Stylianos Loizidis [a,*], Andreas Kyprianou [b], George E. Georghiou [a]

[a] PHAETHON Centre of Excellence (CoE) for Intelligent, Efficient and Sustainable Energy Solutions, Department of Electrical and Computer Engineering, University of Cyprus, 1 Panepistimiou Avenue, 2109 Aglantzia, 20537, Nicosia, 1678, Cyprus

[b] PHAETHON Centre of Excellence (CoE) for Intelligent, Efficient and Sustainable Energy Solutions, Department of Mechanical and Manufacturing Engineering, University of Cyprus, 1 Panepistimiou Avenue, 2109 Aglantzia, 20537, Nicosia, 1678, Cyprus

A R T I C L E I N F O

Keywords:
Market price dynamics
Electricity market forecasting
Price classification
Extreme learning machine
Bootstrap analysis
GARCH
ARFIMA
Artificial neural network

A B S T R A C T

Electricity market liberalization and the absence of cost-efficient energy storage technologies have led to the transformation of state-owned electricity companies into complex electricity market entities, each having a different time horizon. Deregulation has intensified competition, giving rise to increased uncertainty caused by a multitude of interrelated exogenous factors, resulting in unexpected fluctuations in electricity prices. As a consequence, market participants encounter elevated risks and seek effective mitigation strategies. In this paper, the challenges described in the literature are addressed by studying price distribution histograms in the German and Finnish electricity markets. The objective is to identify normal price intervals that can serve as a foundation for an integrated Day-Ahead price forecasting methodology. A novel approach utilizing the Extreme Learning Machine in combination with Bootstrap intervals is proposed and applied to both markets. The findings demonstrate that Bootstrap intervals effectively capture normal prices, whereas extremely high prices typically align with the upper limits of Bootstrap intervals. Conversely, negative prices tend to fall outside the lower boundaries of the intervals. In order to assess the performance of the proposed methodology, a comparative analysis of its forecasting accuracy against the well-established Generalized AutoRegressive Conditional Heteroskedasticity and AutoRegressive Fractionally Integrated Moving Average models is conducted. In addition, both the computational efficiency and forecasting accuracy of the Extreme Learning Machine in comparison to the Artificial Neural Network are assessed. The results reveal the superior efficiency of the Extreme Learning Machine. The developed forecasting model could potentially assist market participants in making well-informed decisions and executing optimal bidding strategies in response to various scenarios before the Day-Ahead market closes. Notably, the proposed methodology transcends the limitations of fixed price thresholds and effectively addresses market nuances, including the occurrence of negative prices, thus offering a more comprehensive approach for electricity price forecasting.

1. Introduction

1.1. Motivation

In the era before the 1990s, electricity generation, transmission and distribution were managed by a single, usually state owned, company. Electricity market liberalization, in the late 1980s and beginning of the 1990s, in association with the lack of cost-efficient energy storage technology [1], had evolved the model of a single state-owned electricity company into a complicated nexus of electricity markets operating at different time horizons. These markets attracted new participants and companies specializing in different sectors of generation, transmission and distribution. The competition within these markets and the advent of renewable sources of energy facilitated the development of new notions and roles such as those of demand response and prosumers [2].

Market participants, amidst intensified competition due to deregulation, have to operate in an environment of increased uncertainty that is attributed to the fuel prices and demand, intermittent power production from the renewable energy sources and unexpected socio-political and public health events such as the war in Ukraine and Covid-19. This phenomenon subsequently induces heightened volatility, manifesting through abrupt and unforeseen fluctuations in electricity prices. Notably, these fluctuations can result in both exceptionally high and exceptionally low prices, as evidenced in 2021 due to heightened demand [3] and shifts in the price of Russian gas. The trend persisted

* Corresponding author.
E-mail address: loizidis.stylianos@ucy.ac.cy (S. Loizidis).

https://doi.org/10.1016/j.apenergy.2024.123058
Received 31 July 2023; Received in revised form 11 March 2024; Accepted 16 March 2024
Available online 21 March 2024

into 2022, exacerbated by the Ukrainian conflict, leading to further instances of extreme prices. The occurrence of negative prices also became more prevalent in 2020, attributed to the global pandemic [4,5]. It is important to underscore that the incorporation of negative prices is a deliberate strategy aimed at mitigating excess production, as outlined in [6]. This highlights the multifaceted nature of market dynamics and the measures undertaken to manage them. As a consequence, the exposure of market participants to risks increases and in order to avoid losses they seek ways to mitigate the effects from this exposure [7].

A tool that market participants use to face market volatility and minimize their risk for losses is market price forecasting. The information provided by forecasts may be exploited in devising better strategies that could lead to profit maximization. Moreover, the availability of dependable forecasts has the potential to incentivize the entry of new market participants, including aggregators and renewable energy providers. The rising prominence of renewables in the electricity production chain is particularly noteworthy. Many countries, in adherence to policy initiatives, give priority to renewables in meeting demand, a trend well-documented in literature [8]. This strategic emphasis on renewables not only aligns with environmental goals but also exerts a downward influence on prices, a phenomenon commonly known as the merit order effect (MOE) [6,9,10], thereby reshaping the supply curve towards more competitive pricing. An additional benefit at the market level is that it assists the market operator to better organize and operate the market [11].

This paper provides a comprehensive methodology for Day-Ahead forecasting of electricity prices. Initially, an in-depth analysis of the Day-Ahead markets in Germany and Finland was undertaken in order to identify the underlying features. The proposed methodology, which combines the ELM and the Bootstrap technique, is thoroughly evaluated across three price classes: normal, extremely high, and negative prices. The evaluation spans an entire year, facilitating a comprehensive assessment of the methodology's performance across different seasons. The choice of utilizing the ELM and Bootstrap techniques is motivated by the ELM's low runtime complexity. This characteristic enables the efficient generation of training replicates through the Bootstrap technique. This aspect is particularly beneficial for market participants as it allows them to test multiple scenarios prior to the closure of the Day-Ahead market, aiding in establishing of well-informed decisions.

Using the replicates generated by the bootstrap technique, forecasting ranges were created, referred to as bootstrap intervals. The analysis showed that the actual prices in the normal price and extremely high price classes consistently fell within these intervals. However, for the negative price class, there were instances of actual negative prices that fell outside the intervals. This suggests that the bootstrap intervals may not fully capture the extreme fluctuations in the negative price class.

In Section 2, we provide an overview of the methodology employed. In Section 3, we conduct a thorough histogram analysis of price distribution for both the German and Finnish Day-Ahead markets. Additionally, this section encompasses the determination of thresholds for normal price ranges. Section 4 is dedicated to presenting the forecasting results, which are systematically compared with the outcomes of GARCH and ARFIMA models. Furthermore, in Section 4, the computational time and performance of the ELM is analysed, drawing comparisons with those of the ANN. The paper concludes in Section 5, where the key findings are summarized.

1.2. Market forecasting literature survey

The literature survey on market price forecasting is divided into two parts for comprehensive analysis. Section 1.2.1 focuses on the literature related to forecasting normal market prices, while Section 1.2.2 discusses the literature that specifically addresses the forecasting of outliers, including extremely high or negative prices. Due to the extensive nature of the literature review, we opted to organize the models and methodologies into text and tables with six columns for normal price forecasting and seven columns for outliers forecasting. The first column specifies the model or methodology, the second column details the targeted market, the third column outlines the utilized data, the fourth column indicates the time span for both training and testing data, the fifth column reports the performance metrics and the sixth column provides the respective references (see Tables 1–3). For the case of outliers, the additional column concerns the separation of the outliers from the normal prices as seen in Table 4.

1.2.1. Normal market price forecasting models/methodologies

Volatility is a key defining trait of electricity market prices, significantly influencing the forecasting process. Over the years, a multitude of methodologies have emerged to predict these prices effectively. These approaches encompass statistical models, computational intelligence models, and hybrid models, with many demonstrating exceptional forecasting capabilities. In the category of statistical models, the most well-known models are Autoregressive (Integrated) Moving Average (AR(I)MA) models [58]. They are time series, statistical-based models that have found many applications in forecasting in general and have also been used in electricity forecasting over the years. In addition, many variants of these models were developed. Table 1 presents research studies employing statistical models like ARIMA, Autoregressive Exogenous (ARX), GARCH and combinations with techniques such as wavelet transform for electricity price forecasting.

Another significant category encompasses computational intelligence algorithms. These algorithms leverage advanced computational techniques to discern complex patterns and relationships within the data, contributing to their efficacy in providing accurate and timely predictions. Computational intelligence algorithms, including but not limited to artificial neural networks (ANN), genetic algorithms, recurrent networks (RN), i.e, neural networks that utilize a loop, deep neural networks (DNN), convolutional neural network (CNN), long-short term memoryless (LSTM), fuzzy logic, and evolutionary algorithms, stand out for their adaptability and capacity to model the non-linear market behaviours. Their application extends across various forecasting scenarios, making them valuable tools in electricity price forecasting.

A meticulous overview of computational intelligence algorithms that had been used in electricity price forecasting before 2014 is given in [59]. It is remarkable that in this review it is stated that the simultaneous forecasting of 24 or 48 values of electricity prices, corresponding to hourly or half-hourly prices for a given day, using ANNs was not widely employed. Nevertheless, in this paper, a survey of the literature on forecasting prices simultaneously for a given day is carried out and reported. In their study [60], the authors acknowledge the existence of numerous methodologies and models for Day-Ahead price forecasting, with no single method being definitively superior. In light of this, they conducted a comparative analysis of four forecasting models: SARIMA, LSTM, CNN-LSTM, and Vector Autoregressive. These models were evaluated using data from the German market during the period from October 2017 to October 2018, which encompassed Day-Ahead prices, load forecasts, emission prices, solar radiation, and wind speed data. Notably, the LSTM model outperformed the others. For the training, they used the data of 360 previous days of the day of interest. Their methodology yielded the lowest RMSE of 9.11 (€/MWh). Also, a comparison between various deep learning models that appear in the literature is given in [61]. More specifically a comparison between LSTM, CNN, Temporal CN (TCN), multilayer perception (MLP) and regression trees (such as random forest) models using Spanish market prices was carried out. The data spanned three periods: The Fraud period (Oct 2016–Jan 2017), the Normal Period (Sept–Dec 2019) and the Quarantine period (15 Mar–15 Jun 2020) It was concluded that in general, the LSTM, CNN and regression trees give better results than the other algorithms. Indicatively, the average MAE of the three models were 4.06 (€/MWh), 4.39 (€/MWh) and 4.08 (€/MWh) respectively.

In a study presented in [62], a combination of machine learning and agent modelling was proposed. The ANN served as the machine learning model, along with an outlier classifier, applied to the PowerACE

Table 1
Statistical models/methodologies for normal market price forecasting.

Model/Methodology	Market	Data	Train/Test periods	Performance	Ref.
ARIMA	Iberian	Prices 2008	Train: 26 Feb 2008–11 Mar 2008 Test: 12 Mar 2008	MAPE = 5.46% RMSE = 4.14 (€/MWh)	[12]
ARIMA	German	Prices 2000–2011	Train: First 2557 data points, Test: 1279 data points	MAPE = 3.55%, Max. Absolute Error = 33.10%	[13]
ARIMA	Danish	Prices, Temperature, Consumption, Production, Wind Power prognosis Hydro reserve, oil and natural gas prices, 2016–2017	Train: Data 2016, Test: Data 2017	MAE = 31.96 (DKK/MWh), RMSE = 38.88 (DKK/MWh)	[14]
ARIMA-Wavelet transform	Mainland Spain	Prices 2002	Train: 1 Jan-17 Feb, 2 Apr-19 May, 2 Jul-18 Aug, 1 Oct-17 Nov Test: 18–24 Feb, 20–26 May, 19–25 Aug, 18–24 Nov	Average Weekly Error = 8.11%	[15]
ARX	German	Day-Ahead and Intraday prices, temperature forecasts, 2016–2019	Train: 1 Oct 2015–30 Sept 2016, Val: 1 Oct 2016–30 Sept 2017, Test: 1 Oct 2017–30 Sept 2019	MAE = 5.92 (€/MWh) RMSE = 8.43 (€/MWh)	[16]
ARX-Augmentation methods ANN-Augmentation methods	Belgian (BE), Dutch (DU)	Prices, weather cond., production and load forecasts, 2016–2018	Train: Data 2016, Validation: Data 2017, Test: Data 2018	MAEs: ARX: BE = 10.02%, DU = 6.92%, ANN: BE = 9.30%, DU = 6.90%	[17]
GCARCH	Mainland Spanish, California	Prices 1999–2000	Test: Last week of each month	MSEs: Mainland Spanish: Dec = 5.49 : California: Jun = 57.42 ($/MWh), Jul = 47.72 ($/MWh), Dec = 43.98 ($/MWh)	[18]

market, an agent-based market simulation model. The methodology was implemented for ten interconnected European countries covering the time period 2020–2050. The data used were prices. The mean MAE (€/MWh) for the ten markets ranged from 0.17 to 16.20. In a study conducted by Tschora et al. [63], various machine learning models, including Support Vector Machines (SVM), Random Forest, Deep Neural Networks, and Convolutional Neural Networks, were employed. Three years of price data were used for training, while two years of price data were used for testing. The models were applied to the French, German, and Belgian markets, and the resulting MAE (€/MWh) ranged from 4.12–4.73, 3.22–4.11 and 6.55–7.28 respectively.

There are several studies suggesting the use of Extreme Learning Machines (ELMs) [64]. ELMs do not require intensive computations and this enables them to be easily combined with other methodologies, usually statistical-based or time series analysis-based, to increase forecasting accuracy. Table 2 presents a compilation of diverse computational intelligence algorithms employed in the domain of electricity price forecasting.

In addition, the errors reported in forecasting tasks indicate the extent of the deviation of the forecasts from the actual values. Usually, a prediction interval is estimated to indicate the range that the actual electricity prices would most likely lie in. This becomes very important when extremely high and/or negative prices occur since algorithms perform poorly when this happens. To this end in [65], a support vector machine (SVM) based algorithm is used to forecast the electricity market prices while, for the prediction interval a non-linear conditional heteroskedastic forecasting model was proposed. Historical electricity market price data from the Australian market in 2004 were used. The training period was the first ten days from May, August and December, and the testing period included the rest of the days. The result yielded an average MAPE of 6.32% for the months of May, August and December. A probabilistic interval forecasting methodology that is based on ELM and maximum likelihood estimation is proposed was [66]. This methodology was tested at four different trading intervals using historical electricity market price and demand data from the Australian market. The target market was NSW (Australian zone) and the training period was from January 2007 to December 2008 and the testing period was from January 2009 to December 2009.

Efforts to improve forecasting accuracy have facilitated the development of hybrid methodologies. These methodologies aim to improve the accuracy by combining two or more different techniques into a single hybrid approach. Such a case is the ELM in combination with the bootstrap method, which was also proposed for both electricity market forecasting and prediction interval determination [67]. The data used to train the ELM were historical electricity market prices from the QLD (Australian zone). Also, the testing period was seven consecutive days from the last month of each season. The average MAPE and RMSE for the four seasons were found to be 13.30% and 8.27%. A wavelet transform-based neural network in the above mentioned ELM-bootstrap based methodology was demonstrated in [68]. This model was tested using price data from the Ontario and the Australian electricity markets. The training period was two months before the day of interest and the testing period was the next 48 h (March, May, August, October and December) This methodology for the Ontario Market for the month May and year 2014 gave prediction interval coverage probability 89.88% and 95.83%, while the prediction interval nominal confidence was 90% and 95% respectively. The prediction interval coverage probability for the Australian market was 91.67% and 94.05%, while the prediction interval nominal confidence was 90% and 95%.

A similar methodology was proposed in [69] where the ELM is used in association with the quantile regression methodology to both forecast the market prices and calculate prediction intervals (PIs). The training data used were prices. The testing period was the twenty consecutive days of the last month of each season, and the training period was the rest of the price data of each season. A probabilistic forecasting methodology is described in [70] where a Bayesian deep learning model for market price forecasting and a heteroscedasticity model for data pre-processing are used. The deterministic parameters of this model were estimated from probability distribution functions in the weight space. The methodology was tested using price, load and production data from the Italian and Belgian electricity markets. The results demonstrated that this methodology had comparable performance to the deterministic neural network.

A probabilistic price forecasting methodology consisting of the variational mode decomposition (VMD), multi-objective sine cosine algorithm (IMOSCA) and regularized ELM was proposed in [71]. The VMD

Table 2
Computational intelligence algorithms for normal market price forecasting.

Model/Methodology	Market	Data	Train/Test periods	Performance	Ref.
ANN-Wavelet transform	Spanish, PJM	Prices, Demand, 2002	Train: 48 previous days	Average weekly MAPE: Spanish = 7.27% PJM = 5.74%	[19]
ANN-Wavelet transform- Bat alg-Scaled conjugate gradient alg.	Spanish	Prices 2002	Test: 18–24 Feb, 20–26 May, 19–25 Aug, 18–24 Nov	Weekly MAPEs: Winter = 0.57%, Spring = 1.11%, Summer = 1.08%, Fall = 1.46%	[20]
ANN-KNN	German	Prices, Fuel prices, Generation, Load, RES, Capacity, Calendar data, 2011–2013	Train: Jul 2011–Dec 2012, Test: Jan 2013–Sept 2013	Average RMSE = 6.58 (€/MWh)	[21]
Distributed DNN	German	Prices, Load and RES forecasts, EU emissions, Fuel prices, 2015–2020	Train+validation: Data 2015–2018, Test: 27 Dec 2018–31 Dec 2020	MAE = 3.54 (€/MWh), RMSE = 6.14 (€/MWh)	[22]
ANN-Transfer learning	Belgian, French, Germany	Prices, Encompassing load, production and RES forecasts, Weather data, 2013–2016	Train: Data 2013–2014, Validation: Data 2015, Test: Data 2016	MAEs: Belgium = 5.66 (€/MWh), French = 4.21 (€/MWh), German = 4.22 (€/MWh)	[23]
ELM-MRA	Ontario, PJM, Italian, New York	Prices 2004	Train: 49 previous days	Average Weekly MAPE: Ontario = 6.60%, PJM = 5.71% Average monthly mean daily error Italian = 8.57%, New York = 3.79%	[24]
Kernel ELM-Wavelet transform	Australian, PJM	Prices (Australian 2014, PJM 2004)	Train: Four previous weeks, Test: First week: Mar, Jun, Sept, Dec	Average MAPEs: Australian = 6.66%, PJM = 5.44%, ANN: BE = 9.30%, DU = 6.90%	[25]
Deep ELM-Sparrow search alg.	DK1 (Danish zone)	Load, Wind-to-Load, Pv-to-Load, 2020–2021	Train: 7 previous days before	MAE = 9.30 (€/MWh), RMSE = 11.90 (€/MWh)	[26]
Elma RNN	New York, Spanish	Prices (New York 2010, Spanish 2002)	Train: 42 previous days	Average MAPEs: New York = 3.82%, Spanish = 6.56%	[27]
NN-Fuzzy logic	Spanish	Prices 2002	Train: 48 previous days, Test: One week for each season	Average error: 7.80%	[28]
CNN-Mutual Information-Modified relief	PJM, Spanish	Prices (PJM 2006, Spanish 2002)	Train: 24 different CNNs using 50 samples	Average Weekly MAPE: PJM = 4.55%, Spanish = 5.88%	[29]
Random Forest-Pearson Corr.	Gansu (China zone)	Prices, forecast (Load, production, Day-Ahead tie line, RES) data, 2021	Train: First 20 days, Test: 21th day-D-1 day	MAE = 75.98 (Yen/MWh)	[30]
Two-stage supervised learning alg.	NSW	Prices, Wind production forecasts, 2021	Train: First 23 days of each month, Test: Next 7 days	MSE = 26.34 ($/MWh)	[31]
LSTM-Wavelet transform-Cicross optimization	DK1 (Danish zone)	Prices, Wind and Solar production, Forecasted Load, 2018–2019	Train: First 610 data points, Test: The rest data points	MAE = 1.31 (€/MWh)	[32]

method is used to decompose the training data, while the IMOSCA is used to optimize the regularized ELM parameters and compute the forecasting intervals. The data used were market prices relating to New South Wales and Queensland (Australian zones). The hybrid model was tested for one, two, four, and six forecasting steps. The training period was the first twenty-five days of June 2016, and the testing period was the last five days. The results obtained for New South Wales were MAE-$/MWh (RMSE-$/MWh) 5.05 (6.62), 5.04 (7.12), 10.98 (14.25) and 18.03 (22.52) respectively. For Queensland the results were MAE-$/MWh (RMSE-$/MWh) 4.91 (6.41), 4.92 (6.52), 11.57 (14.32) and 15.83 (20.00) respectively. A different approach to electricity price forecasting was employed in [72] where multivariate probabilistic forecasts derived from univariate forecasts. Their dataset consisted of historical Day-Ahead price data and load forecasting data. The training periods for each market were: Germany: 5 January 2015 to 3 January 2016, PJM: 29 December 2015 to 26 December 2016, Belgium and France: 5 January 2014 to 3 January 2015. The testing periods for each market were: Germany: 4 January 2016 to 31 December 2016, PJM: 27 December 216 to 24 December 2018, Belgium and France: 4 January 2015 to 31 December 2016. Notably, they achieved impressive results with continuous ranked probability scores of 3.14% (Germany), 2.39% (PJM), 5.50% (Belgium), and 3.31% (France). Given the extensive number of hybrid models developed, Table 3 presents additional hybrid methodologies in the context of electricity price forecasting.

1.2.2. Outliers forecasting: Models/methodologies

In the literature, there are works describing different endeavours to forecast very high market prices (spikes) and negative prices, with most papers presenting algorithms that attempt to forecast the spikes. To this end in [73] an ANN with wavelet to forecast normal electricity market prices was suggested, while for forecasting the range of spikes of electricity market prices and the level of confidence, a Bayesian method of classification was proposed. The data used were electricity market prices, demand, and production reserves of the Australian market. In order to detect high prices the threshold $\mu \pm 2\sigma$ was used and for this application was detected to the 75 ($/MWh). The training period was the data from January to June 2003, the testing period was from July to September. The results obtained demonstrated that in most cases, the forecasted market price was close to the actual electricity market price,

Table 3
Hybrid methodologies for normal market price forecasting.

Model/Methodology	Market	Data	Train/Test periods	Performance	Ref.
Kernel ELM-Wavelet transform-ARMA	PJM, Australian, Spanish	Prices 2013–2014	Train: PJM-Two weeks before, Australian-40 previous days, Spanish-Four previous weeks	MAPEs: PJM = 5.23%, Australian = 4.10%, Spanish = 4.35%	[33]
ELM-VMD-Ensemble empirical modal decomposition (EEMD)	NSW, Victoria, Spanish	Prices 2019–2020	NSW-VIC: Train: First 4911 data points, Test: Next 800 data points Spanish: Train: First 4984 data points, Test: Next 800 data points	MAEs (6-step ahead): NSW = 2.23 ($/MWh), VIC = 2.66 ($/MWh) and Spanish = 2.77 (€/MWh)	[34]
Kernel ELM-Variational Mode Decomposition-Chaotic sine-cosine alg.	New South Wales (NSW), Queensland (QLD), Singapore	Prices 2018	Train: 25 previous days	MAEs: NSW = 0.38 ($/MWh), QLD = 0.58 ($/MWh), Singapore = 0.76	[35]
BPNN-Fast (EEMD)-VMD	NSW	Prices 2016	Train: first 1200 data points, Test: Next 240 data points	MAEs ($/MWh): 1-ahead = 10.54, 2-ahead = 15.43, 4-ahead = 21.10	[36]
LSTM-Adaptive copula-based feature selection (ACBFS)-Decomposition	PJM	Prices 2015–2017	Train: First 1440 data points of each season, Test: Data 2017	RMSE = 2.19 ($/MWh)	[37]
CNN-LSTM	Iranian	Prices 2020–2021	Train: 80%, Test: 20%	Standardized RMSE = 0.09, Standardized MAE = 0.057	[38]
Bi-LSTM-Random Forest-Mahalanobis dist	DK1 (Danish zone)	Prices, Load, Consumption, RES actual and forecast data, 2020	Train: Randomly selected 50 days	RMSE = 1.48 (€/MWh), MAE = 1.77 (€/MWh)	[39]
LSTM-Random Forest	Greek (GR), Bulgarian (BL), German (DE), Austrian (AT)	(1) Forecasting data, Natural gas prices, Carbon emissions prices, (2) Data from (1) and neighbouring prices, 2021–2022	Not clarified	(1) MAEs(€/MWh): GR = 21.12, BL = 34.35, DE = 32.78, AT = 24.89, (2) GR = 18.11, DE = 30.56	[40]
LSTM-ANN	Iberian	Prices 2015–2019	Dat: Jan 2015–Aug 2018: Train: 77%, Validation 33%, Test: Data 2018–2019	RMSE = 5.03 (€/MWh)	[41]
DNN-LSTM-Wavelet transform	Spanish	Prices, Load 2002	Test: Last week of each season	Weekly MAPE = 2.20%	[42]
BiLSTM-EEMD	NSW	Prices, Load (Jul, Oct,Dec)2020, (Jan 2021)	Train: First 1190 data points, Test: Next 298 data points	RMSE($/MWh): Jul = 2.99, Oct = 2.43,Dec = 0.99, Jan = 1.34	[43]
EEMD-Max dependency and Min redundancy criterion	NSW, PJM	Prices (NSW:2018–2019, PJM:2017–2018)	Train: First 8 previous weeks, Test: First day of Mar, Jun, Sept, Dec	Average RMSE: NSW = 30.46 ($/MWh), PJM = 7.30 ($/MWh)	[44]
ANN-ANFIS-Multi-objective binary backtracking alg.	QLD	Prices 2018	Not clarified	Average MAPEs: Nov = 4.52%, Aug = 5.62%, May = 4.43%, Feb = 4.08%	[45]
ANFIS-Multi-objective binary backtracking alg.-Mutual Information	Ontario	Prices, Demand 2017	Train: Data of the first week, Test: Data of the last week	MAPE = 2.79% ($/MWh)	[46]
Extreme Gradient Boosting-Polynomial kernels-Net regression	ISO New England	Locational marginal prices 2012	Train: First 23 days of each month, Test: Next 7 days	Average MAE = 2.60 ($/MWh)	[47]
SVM-Cuckoo alg.-Singular Spectrum	NSW	Prices 2016	Train: First 20 days, Test: Next 10 days	MAPE = 0.83%	[48]

and therefore, the error was less than 30%, with one case depicting an error of 49.43%. A different approach to forecast electricity price spikes that utilizes the Autoregressive Conditional Hazard (ACH) was proposed in [74]. The proposed model used market prices, load and temperature recordings related to Australia. The price ranges that were considered as spikes were $100 \leq P_t \leq 300$ and $300 \leq P_t \leq 10,000$ ($/MWh). This study presented two cases that where dependent on actual and forecasted training load data respectively. The training period was from 1 March 2001 to 30 June 2007, while the testing period was from 1 July 2007 to 30 September 2007. In the first case it was found that using the ACH model the MAE and RMSE obtained were 0.078 ($/MWh) and 0.19 ($/MWh) respectively, while the benchmark model yielded an MAE and RMSE of 0.074 ($/MWh) and 0.22 ($/MWh) respectively. In the second case using the ACH model the obtained MAE and RMSE were 0.10 ($/MWh) and 0.20 ($/MWh) respectively, while those of the benchmark model were 0.078 ($/MWh) and 0.22 ($/MWh) respectively.

A methodology based on five different filters of detecting and replacing outliers was proposed in [75]. After the replacement, the time series were divided into deterministic and stochastic parts. In the deterministic part parametric and non-parametric models for forecasting are applied, while forecasting of the stochastic part was done using Auto-Regressive (AR), non-parametric Auto-Regressive (NPAR), Auto-Regressive moving average (ARMA) and vector Auto-Regressive

Table 4
Methodologies for outliers forecasting.

Model/Methodology	Market	Data	Price separation	Train/Test periods	Performance	Ref.
NNs	Ontario	Spike prices, Demand, Temperature, Dew point temperature, Humidity 2012	Threshold $= \mu \pm 2\sigma$	Train: Two previous years, Test: 3 Jan, 3 Mar 2012	RMSEs: 3 Jan = 15.14, 3 Mar = 8.90	[49]
Decision trees	DK1 (Danish zone), SE1 (Sweden zone), Finnish	Prices, Load, production Allocated transmission capacity, Power flows, RES forecasts, 2012–2013	Threshold $= \mu \pm 2\sigma$	Train: One previous week	MAEs: DK1 = 3.58 (€/MWh), SE1 = 1.94 (€/MWh), Finnish = 3.57 (€/MWh)	[50]
ARIMA-NN-KNN	Finland	Prices, Demand, Supply, 2016–2017	Threshold $= \mu + 3\sigma$	Train: Data 2016, Test: Data 2017	Adapted mean average percentage error = 8.08%	[51]
ANN-Variance stabilization transformation	French	Prices, Demand, RES, Reserve capacity, Interconnector flow data 22 Sept 2016–21 Sept 2019	Positive spike: > 75.56 (€/MWh), Negative spike: <−3.90 (€/MWh)	Train: 626 spike data points, Test: 721 spike data points	MAE = 10.62 (€/MWh)	[52]
ELM	Ontario	Prices, Demand, 2010	$T_1 = -200$, $T_2 = 50$, $T_3 = 100$, $T_4 = 2000$ Class 1: if $T_1 < Price < T_2$ Class 2: if $T_2 < Price < T_3$ Class 3: if $T_3 < Price < T_4$	Test: Data 2010	Percentage Correct Classification: Jan = 91.80%, Mar = 99.73%, Jun = 91.53%, Sept = 93.19%	[53]
SVM	Australian	Prices, Demand, Supply, existence, Season, Time, net interchange, Dispatchable load, Sept 2003–June 2004	Threshold $= \mu \pm 2\sigma$ Spikes: Price > 75 ($/MWh)	Train: Data of Sept 2003–May 2004 Test: Data of June 2004	MAPE = 9.39%	[54]
ARX time Varying (Normal prices) SVM-Kernel regression (Positive spikes)	NSW	Prices, 2010–2015	Threshold $= \mu + 2\sigma$ Normal prices: ≤ 75 ($/MW) Positive spike: > 75 ($/MW)	Train: Data 2010–2014 Test: Data 2015	MAE = 3.75 ($/MWh), RMSE = 8.92 ($/MWh), MAPE = 6.47%	[55]
PNN-Wavelet transform-Mutual Information	PJM	Prices (Sept 2003–June 2004)	Threshold $= \mu \pm 2\sigma$ (1) Positive spike: > 150 ($/MWh) (2) Positive spike: > 200 ($/MWh)	Train: Data of Sept 2003–May 2004 Test: Data of June 2004	(1) Spike forecast accuracy = 97.3%, Confidence Interval = 87.70%, (2) Spike forecast accuracy = 92%, Confidence interval = 88.5%	[56]
CNN-Denoising autoencoders- EEMD	NSW	Prices, Consumption, Production, RES, Population economy, Weather data 25 Oct 2014–27 Jul 2020	Threshold $= \mu \pm 3\sigma$	Train: 25 Oct 2014–24 Oct 2019, 25 Jan 2015–24 Jan 2020, 26 Apr 2015–25 May 2020, 26 Jul 2015–25 Jul 2020, Test: 25–27 Oct 2019, 25–27 Jan 2020, 25–27 Apr 2020, 25–27 Jul 2020	MAPEs: Spring = 8.40%, Summer = 8.36%, Autumn = 9.50%, Winter = 9.70%	[57]

(VAR) time series models. It was concluded that using Recursive Filter and Moving window filter on prices in combination with threshold for peak replacement gave better results as compared to the rest of the proposed filters and peak replacement methods. The data they used were Italian market prices. The training period was between January 2012 to December 2016, while for testing it was the whole year of 2017. A Multivariate Logistic Regression (MLGR) based approach of forecasting the probability of occurrence of extremely high or low prices was proposed in [76]. The market price separation was done using thresholds. The threshold for identifying extreme low prices was determined at the 5th quantile, while for extreme high prices, it was set at the 95th quantile. Specifically, prices below 4.5 (AU$/MWh) were classified as extremely low values, and those exceeding 18.18 (AU$/MWh) were considered extremely high values. Also, the data used were historical market prices (one day before and one week before the forecasting day), reserve capacity, load, renewable production and inter-connector flow. The data was for the South Australia market, 2018–2019, and 75% was used for training and 25% for testing. This model gave better results compared to Multi-Layer Perceptron (MLP) and Radical Basis Function (RBF).

The authors in [77] studied the contributing factors to the appearance of extreme market prices in the German Day-Ahead market. They did this by first separating the spikes into positive and negative and referred to the reasons (or factors), such as supply, demand and the time, of each spike occurrence. In addition, they pointed out that renewable production affects the probability of spike occurrence. The forecast data analysed were related to the demand, supply, wind and pv and covered the period of 4 January 2010–31 May 2014. It was found that most of the positive spikes were due to very high demand, which was between 17:00–19:00. Negative market prices mostly occurred during the night when demand was very low and wind production was very high. Bello et al. [78] proposed a hybrid model for probabilistic low market price forecasting in the medium-term. The model incorporates logistic regression for rare events, decision trees, and multilayer perceptrons. Using the kernel method, they established a low price range of 10–15 (€/MWh). Their methodology was applied to the Spanish market. For training, they used the price data of the period January 2009–November 2011, while for testing December 2011–March 2012. The application of their methodology yielded the following number of occurrences of low prices: 4 in December, 7 in January, 11 in February and 15 in March. The actual number of low prices was correspondingly 6, 9, 11, and 11. Table 4 presents additional models/methodologies designed for outliers forecasting.

It is crucial to note that assessments of each highlighted methodology were conducted on specific markets, typically spanning less than a year, ranging from some days or weeks to a few months. The literature survey reveals a gap in comprehensive discussions regarding the performance of these models/methodologies across all days of an entire year. Additionally, in studies that classify the prices into different ranges, a fixed threshold is utilized to categorize them as normal or extremely high/low. In markets many negative prices occur, this method proves unsuitable, as days with negative prices are considered normal in such cases. Further discussion on this can be found in sub-Section 3.1.

The novelty of the work described in the paper is outlined below:

- In-depth Exploratory Analysis: The work presents an exploratory analysis of Day-Ahead electricity price histograms for both the German and Finnish markets. This analysis gives insights not previously reported in the literature.
- Novel Classification Framework: An innovative classification framework is presented that categorizes electricity prices into three distinct classes: normal, extremely high, and negative. This framework is based on precision theory and leverages price distribution histograms, offering a new perspective on price behaviour analysis.
- Advanced Forecasting Methodology: A forecasting methodology that combines ELM with Bootstrap intervals is introduced. This approach demonstrates superior predictive performance and robustness across various market conditions.
- Rigorous Comparative Analysis: A comprehensive comparative analysis that not only evaluates the proposed methodology's forecasting accuracy but also contrasts it with established models like GARCH and ARFIMA is provided. In addition, a detailed comparison of time execution and error performance against ANN architectures is concluded. This extensive analysis offers valuable insights into the efficacy and efficiency of the proposed approach when compared to existing models.

2. Overview of methodology

2.1. Extreme learning machine (ELM): Outline

The proposed forecasting algorithm is based on both the Extreme Learning Machine (ELM) [64] and bootstrapping [67]. In this section a brief outline of ELM is given. ELM is a single-layer Feedforward Neural Network (SLFN) that can be trained in a computationally efficient manner. This renders it amenable to be combined with other statistical techniques in order to either improve the forecast or unveil the statistical characteristics of the results. The motivation behind proposing the hybrid ELM-bootstrapping algorithm is the improvement of the forecast of the basic ELM algorithm and through statistical analysis to identify those market conditions that lead to the appearance of extremely high and negative prices.

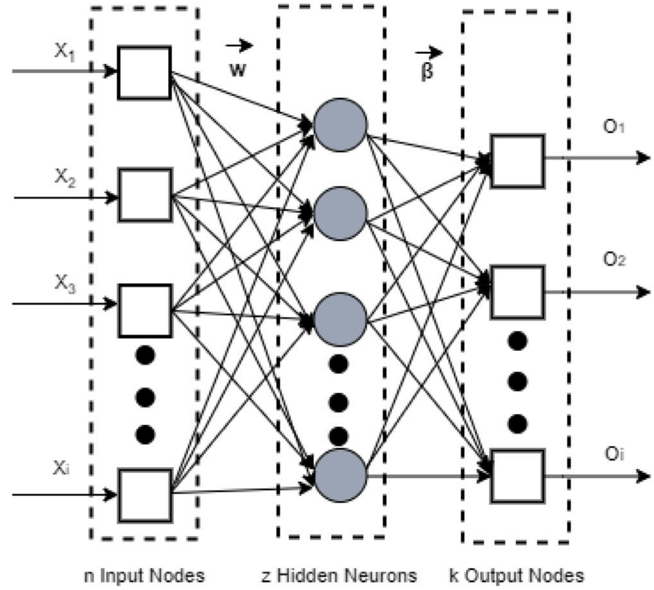


Fig. 1. A typical structure of ELM.

A brief description of the mathematical background of the ELM Single Hidden Layer Feed-Forward Neural Network is given for the sake of completeness. A neuron in an artificial neural network mimics the basic function of a physiological neuron, that of receiving a stimulus and responding with an action potential which is transmitted to other neurons or effector organs. Fig. 1 depicts an artificial neural network consisting of three layers of nodes, with the nodes of the middle layer designated to represent the artificial neurons. Each of the z neurons interconnect with the corresponding input node in the first layer and output node in the third. The first layer consists of n input nodes where n equals the dimension of the input vectors. These are passive nodes as they only indicate the entrance to the network of the corresponding component of the input. The nodes of the second layer correspond to the neurons and they are the active part of the neural network. The z neurons of the second layer are linked with all n input nodes and all k output nodes. Each neuron j is associated with two weights: the w_{ij} that multiplies the value of ith input component x_i and the β_{ij} of which the role is described later.

In order to facilitate the exposition of ELM parameters w_{ij} and β_{ij} for an ELM consisting of n input neurons, z hidden layers nodes and k output nodes can be cast in z n-dimensional vectors $\vec{w}_i$ and z k-dimensional vectors $\vec{\beta}_i$. The k-dimensional output of the ELM, $\vec{y}$, as shown in Fig. 1, to an n-dimensional input vector $\vec{x}$ is given by,

$$\vec{y} = \sum_{i=1}^{z} g\left(\vec{w}_i^T \vec{x} + b_i\right) \vec{\beta}_i \tag{1}$$

which is the superposition of the vectors $\vec{\beta}$ weighted by the values of the activation function $g(.)$.

In applications, the neuron weight vectors $\vec{\mathbf{w}}_\mathbf{i}$, neuron biases $\mathbf{b_i}$ and output weight vector $\vec{\beta}_\mathbf{i}$ must be estimated from the available observations in an optimal way. To this end for a given set of m observations consisting of n-dimensional vectors $\vec{x}_i$ and k-dimensional target values $\vec{O}_i$ the neuron weight vectors $\vec{w}_i$ and neuron biases are selected randomly. Subsequently, Eq. (1) is used for each of the m observations $\vec{x}_i$ and $\vec{O}_i$ and the known vectors $\vec{w}_i$, biases b_i and thereby creating the following system of m equations,

$$\sum_{i=1}^{z} g\left(\vec{w}_i^T \vec{x}_j + b_i\right) \vec{\beta}_i = \vec{O}_j, \ \ j = 1, 2, \dots, m \tag{2}$$

which in turn can be written in matrix form,

$$H\beta = O \tag{3}$$

where the matrix H is given by,

$$H = \begin{pmatrix} g\left(w_1, b_1, x_1\right) & \cdots & g\left(w_n, b_n, x_1\right) \\ \vdots & & \vdots \\ g\left(w_1, b_1, x_k\right) & & g\left(w_n, b_n, x_k\right) \end{pmatrix}_{k\times n} \tag{4}$$

and is known as the hidden layer output matrix, O is the $k \times m$ matrix of the observed target vectors, and β is the $z \times k$ matrix whose columns are the k dimensional vectors $\vec{\beta}$.

Training is achieved by computing the vectors $\vec{\beta}$ by solving Eq. (3) with respect to $\vec{\beta}$,

$$\vec{\beta} = H^{\dagger} O \tag{5}$$

where $H^{\dagger}$ is the Moore–Penrose generalized inverse (or pseudo-inverse) of the matrix H.

The accuracy of forecasting is assessed using the commonly used metrics of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) given by the Eqs. (6) and (7) respectively,

$$\text{MAE} = \frac{1}{\text{n}} \sum_{i=1}^{n} \mid y_{actual,i} - y_{forecasted,i} \mid \tag{6}$$

$$\text{RMSE} = \sqrt{\frac{1}{\text{n}} \sum_{i=1}^{n} (y_{actual,i} - y_{forecasted,i})^2} \tag{7}$$

The number of ELM input and output nodes is determined in Section 2.3, whereas the choice of the actual activation function and the number of hidden neurons and its rationale are given in Section 2.5.

2.2. Description of the bootstrapping method

The low run-time complexity implementation of ELM enables the simultaneous use of other statistical methods in view of enhancing the accuracy of the forecasts. A widely used method that is based on resampling and replacement is bootstrapping [67]. In the context of forecasting, the residuals, the difference between the forecasts given by an estimated model and the training data, are first computed as follows,

$$\epsilon_i = y_i - \hat{O}(x_i) \tag{8}$$

where $\hat{O}(x_i)$ are the forecasts obtained by the trained ELM. These residuals are subsequently recentered according to,

$$\hat{\epsilon}_i = \epsilon_i - \frac{1}{\text{n}} \sum_{j=1}^{k} \epsilon_j, \ \ i = 1, 2, \dots, k \tag{9}$$

and finally are resampled with replacement subsequently the newly generated residuals are added to the existing forecasts to create new training data as follows

$$y_i^{\star} = \hat{O}(x_i) + \hat{\epsilon}_i \tag{10}$$

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

2.3. Organizing the training data. Determination of the number of the ELM input and output nodes

The actual number of ELM input and output nodes is determined by the dimensions of the input vectors $\vec{x}_i$ and output vectors $\vec{O}_i$.

Table 5
Training data quantities.

Historical Day-Ahead market prices	Production forecast
Consumption Forecast	Forecasted Residual Load
Difference between Production Forecast and Consumption Forecast	Ratio: Production Forecast-to-Forecasted Residual Load
Ratio: Renewable Production Forecast-to-Consumption Forecast	

Table 6
German market, Average of ratio: $\frac{\text{RES Forecast}}{\text{Consumption Forecast}}$.

Time	2019	2020	2021	2022
00:00–06:00	0.31	0.33	0.28	0.30
07:00–13:00	0.39	0.45	0.41	0.46
14:00–23:00	0.32	0.33	0.29	0.32

This section describes the actual training data and the way they are organized to determine the input and output vectors.

A sample of training data consists of historical data of 24 consecutive days preceding the target day. The variables included in the training data for ELM are shown in Table 5. Another feature of the methodology is the division of the data to three different periods of the day, namely, 00:00–06:00, 07:00–13:00 and 14:00–23:00. This data separation follows the trend of the ratio renewable production forecast to consumption forecast for the German market as depicted in Table 6. During the 00:00–06:00 h, this ratio has lower values compared to the 07:00 to 13:00 period. Although the ratio in the 14:00–23:00 period exhibits similar values to the period 00:00–06:00 period, it has been decided to keep it as a separate class. This decision is based on the likelihood that the ratio in the 14:00–23:00 period is influenced by higher consumption forecast whereas in the 00:00–06:00 period it is affected by decreased consumption and renewable production.

A different ELM is trained for each training sample, in order to compute the vector $\vec{\beta}$, using randomly selected input weights, w, and biases, b, as described in Section 2.1. Different ELMs are trained for each period and the number of the input nodes depends on the number of hours that each period lasts. That is for the classes corresponding to the periods 00:00–06:00 and 07:00–13:00 the number of input nodes is 168 (number of hours (=7)×24 days), while for the 14:00–23:00 class this number is 240 (number of hours (=10)×24 days). As a consequence of this classification, the number of output nodes for the periods 00:00–06:00 and 07:00–1300 are 7 whereas for the period 14:00–2300 is 10.

The training data are scaled within the range $[0, 1]$ as follows,

$$X_{sc} = \frac{\text{x} - \text{x}_{\text{min}}}{\text{x}_{\text{max}} - \text{x}_{\text{min}}} \tag{11}$$

where x is the sample data, x_{min} is the minimum sample value and x_{max} is the maximum sample value.

2.4. Proposed forecasting methodology

The previous subsection, 2.3, describes how a sample is created. Briefly, a sample is composed of the input vector $\vec{x}_i$ of $n = 168$ observations for data corresponding to periods 00:00–0600 and 07:00–13:00 while for the period 14:00–23:00 of $n = 240$ observations and the target vectors $\vec{O}_i$ of $k = 7$ and $k = 10$ observations correspondingly. The dimensions of the input and output vectors, 168 and either 7 or 10 respectively, determine the number of ELM input and output nodes, namely 168 input nodes and either 7 or 10 output nodes. The next step of the methodology is to determine the number of hidden layer nodes (actual neurons) and their activation function (see 2.5) and finally to implement the actual forecasting. In the remaining of this section, the overall proposed forecasting methodology is described.

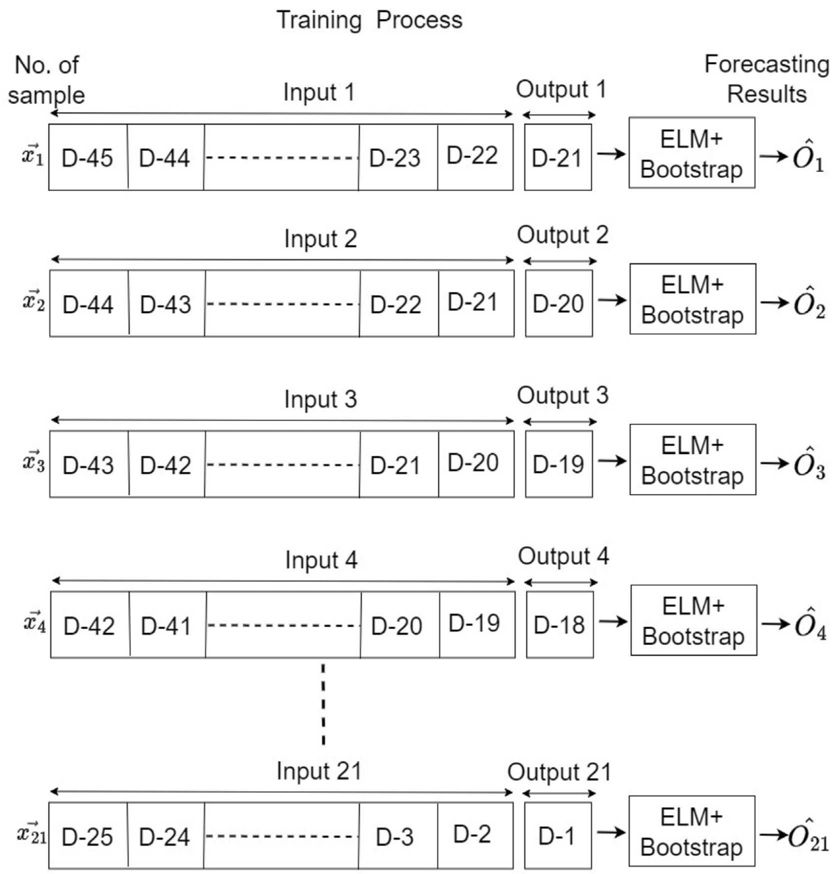


Fig. 2. Proposed methodology diagram.

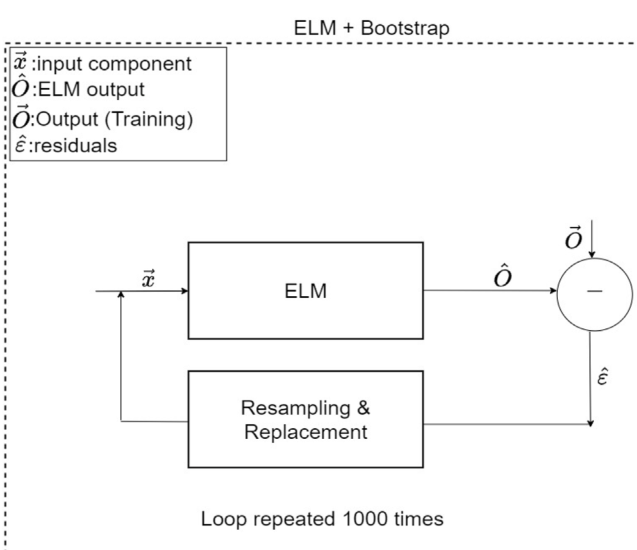


Fig. 3. ELM-Bootstrap process diagram.

The methodology is built around the ELM incorporating the bootstrapping technique described in Section 2.2. Next let us assume that a forecast for a target day, D is required. The prices and the other data for the target day D are not available and thereby to implement forecasting the available historical data must be exploited. To this end 21 training samples are constructed using the available data of the previous 45 days as follows. The nth day before the target day is denoted by $D-n$, for example, the 45th day before the target day is denoted by $D-45$. As it is explained in 2.3 a sample contains an input vector created by data of 24 days and an output vector which contains the actual prices of the 25th day. Starting from the day $D-45$ data is collected in the input vector $\vec{x}_1$, Fig. 2, for the next 23 days whereas the output vector $\vec{O}_1$, Fig. 2 consists of the actual prices of the $D-21$. The same is repeated for $D-44$ up to $D-24$ to create 21 samples containing the input data obtained from data of days $D-25$ to $D-2$ and output the prices of $D-1$. The 21 samples corresponding to the 21 input vectors $\vec{x}_i$ and output vectors $\vec{O}_i$ $i = 1 \ldots 21$ are depicted in Fig. 2. The number of 45 days was also chosen empirically by trial and error, as this number yielded the best forecasts.

Having created 21 samples instead of training a single ELM using all samples, an ELM for each sample is trained and thereby 21 ELMs are trained in total. The actual training of each ELM is finalized using the bootstrapping technique described in Section 2.2. Fig. 3 shows the ELM-Bootstrap process diagrammatically. More specifically, the ELM output ($\hat{O}$) is subtracted from the training data $\vec{O}$ to compute the residuals, in which resampling and replacement are implemented 1000 times for each one of the 21 ELMs, i.e. 1000 new training data samples are resulted according to Eq. (8). Once the ELMs are trained they are tested by using the sample corresponding to the last day D. That is, the sample of the day D consists of the input vector composed of data corresponding to day D-24 up to day $D-1$ and the output vector that contains the actual prices of the day D. The final result is the average vector of output vectors of the 21 trained ELMs, given by,

$$\hat{O} = \frac{1}{21}\sum_{i=1}^{21} \hat{O}_i \tag{12}$$

The proposed methodology is employed 365 times, corresponding to each day of the year in order to account for the unique characteristics of each day. This enables the methodology to adapt to the varying

Table 7
Defining the optimal value for α.

Neurons	1×10^{-3}	1.5×10^{-3}	2×10^{-3}	2.5×10^{-3}	3×10^{-3}	3.5×10^{-3}	4×10^{-3}	4.5×10^{-3}	5×10^{-3}
	$\times10^{-15}$	$\times10^{-15}$	$\times10^{-15}$	$\times10^{-15}$	$\times10^{-15}$	$\times10^{-15}$	$\times10^{-15}$	$\times10^{-15}$	$\times10^{-15}$
5	9,70	6,40	8,93	9,28	8,81	9,06	8,36	1,39	7,47
10	8,07	6,76	7,30	8,46	7,19	9,11	6,06	6,64	7,10
15	8,23	7,94	7,90	7,46	7,76	7,51	1,57	9,38	7,89
20	9,89	1,25	1,28	1,35	1,26	9,58	9,63	2,47	9,67
25	1,04	9,39	9,67	9,18	9,62	9,48	6,72	1,69	8,22
30	6,46	6,19	6,60	1,29	7,04	6,25	9,37	7,75	8,86
35	8,48	8,48	8,85	7,16	8,45	7,83	7,89	1,79	9,32
40	5,52	9,01	7,99	8,65	2,17	1,92	8,70	775	8,03
45	6,93	8,31	3,20	9,17	6,86	8,77	9,20	8,09	8,96
50	7,76	7,01	7,83	6,73	1,30	7,85	8,02	8,84	8,38
55	6,93	8,20	7,25	7,44	6,11	7,62	6,74	6,54	7,65
60	5,92	7,92	8,80	7,38	7,36	5,92	1,88	5,10	8,35
65	5,67	8,86	1,19	1,52	8,72	7,12	9,44	2,52	7,30
70	6,93	8,86	7,59	9,34	7,31	9,44	2,20	8,71	2,20
75	8,87	7,75	9,87	6,28	3,43	9,00	97,17	1,15	8,77
80	8,48	4,25	7,42	1,41	9,20	2,73	8,39	8,42	1,82
85	8,21	8,24	5,31	1,66	4,12	7,12	9,85	8,28	6,73
90	9,91	8,28	9,46	8,63	9,00	6,49	7,22	8,97	2,76
95	7,19	8,11	2,23	8,30	8,12	7,22	9,82	6,65	2,02
100	8,29	9,32	7,89	9,02	8,64	9,32	8,55	8,61	9,08

statistical behaviour of prices, especially during the years 2020 and 2022 when unexpected socio-political events such as the Covid-19 pandemic and the Ukrainian crisis led to substantial and structural changes in price dynamics. The day-by-day application of the proposed methodology aims to capture the nuances and regime-switching that occurred in the price levels and price volatility, thus enhancing the stability and validity of the forecasts.

2.5. Determining the number of hidden neurons and activation function

In order to determine the number of nodes in the hidden layer and the activation function, the methodology described in Section 2.4 is implemented by using the combinations of the number of hidden nodes and activation functions depicted in Table 8. The activation functions evaluated were the sigmoid function, arctan function, and the derivative Eliot function, given by Eqs. (13), (14), and (15) respectively. In these equations, α is a scaling parameter that needs to be determined based on the data.

$$g(x) = \frac{1}{1 + e^{-(\alpha x)}} \tag{13}$$

$$g(x) = \arctan^{-1}(ax) \tag{14}$$

$$g(x) = \frac{1}{(1 + |\alpha x|)^2} \tag{15}$$

To this end, the parameter α was varied from 1×10^{-3} to 5×10^{-3} in steps of 0.5×10^{-3} and tested, using the proposed methodology, on ELMs whose number of internal neurons was varied from 5 to 100 in steps of 5. The errors of these combinations are tabulated in Table 7. The smallest MAE error is obtained when the number of hidden neurons is 25 and the value of the parameter α is 1×10^{-3}. The actual error is 1.04×10^{-15} (€/MWh). In the subsequent comparison of the sigmoid function with the other activation functions the parameter α is set to this value.

The errors upon completion of the proposed methodology on the training data themselves are also listed in Table 8. The smallest error is obtained when the number of hidden nodes is 20 and the activation function is the sigmoid. The error in this case is 1.43×10^{-15} (€/MWh). The errors listed in first column of Table 7, corresponding to the sigmoid function with $\alpha = 1 \times 10^{-3}$, are different to those listed also in the first column of Table 8. Even though one expects these errors to be the same, they are not as the weights that are used by the ELMs are randomly selected. Since using 25 hidden neurons in both cases gives similar errors close to the smallest this number is selected. As a result

Table 8
Defining the number of hidden nodes and activation function.

Number of hidden nodes	Sigmoid MAE (€/MWh) $\times 10^{-15}$	Arctan MAE (€/MWh) $\times 10^{-15}$	Deriv. Eliot Sig. MAE (€/MWh) $\times 10^{-15}$
5	6,87	4,46	5,87
10	1,54	5,15	5,56
15	6,70	5,89	6,42
20	1,43	7,32	6,22
25	1,56	5,01	5,87
30	6,96	5,90	5,41
35	5,01	6,69	5,65
40	1,96	4,55	5,34
45	2,13	4,93	5,63
50	2,11	5,29	5,90
55	3,30	5,03	5,22
60	5,63	5,63	6,37
65	5,84	5,31	5,50
70	3,08	5,31	5,52
75	6,12	4,99	5,74
80	4,07	5,72	5,73
85	5,13	5,51	5,74
90	4,20	5,75	6,36
95	3,81	5,49	5,50
100	6,88	5,03	5,60

of this, the ELMs used for the subsequent study are composed of 25 hidden neurons having sigmoid as their activation function with the parameter α set to the value 0.001.

2.6. Methodology overview

The flowchart presented in Fig. 4 outlines the sequential steps of the proposed methodology. Initially, data, including historical Day-Ahead prices and forecast data, is acquired. Subsequently, data corresponding to each 24-h period is subdivided into three distinct time intervals: 00:00–06:00, 07:00–13:00, and 14:00–23:00. The training data is then subjected to scaling, and subsequently, the optimized parameters of the ELM are obtained, namely, the coefficient α, the activation function, and the number of internal neurons.

3. Data selection and organization

3.1. Data sources

The data used in this study relate to the German [79] and Finnish Day Ahead markets. The data concerning the German market are the

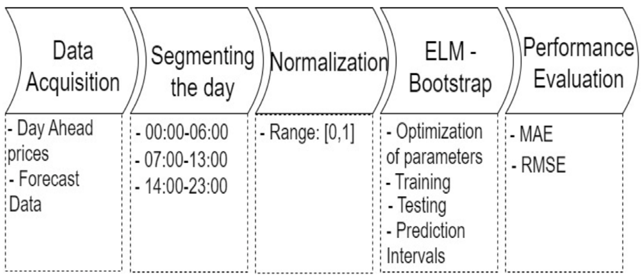


Fig. 4. Flow chart of the proposed methodology.

historical actual electricity market prices and the historical forecasts of the production, consumption, renewable production and residual load (or net load). The training data was obtained by computing: the difference between the production and consumption forecast, the ratio of production forecast to forecasted residual load and the ratio between the renewable production forecast and the consumption forecast. For the Finnish Day-Ahead market, the electricity prices were obtained from the Estonian TSO (Elering) [80], which in turn draws the data from the Nord Pool, while the rest of the data from the Finnish TSO (FinGrid) [81]. The data used is openly accessible.

3.2. Price histograms

The market conditions that affect the electricity price forecasting were analysed in [2] and it was found that the histograms of the prices provide information about their distribution that is useful in designing the overall forecasting methodology. In this section, the histograms of electricity prices of German and Finnish day ahead markets for the period 2019 – 2022 are plotted in Figs. 5(a)–8(b). In what follows the features appearing in these histograms are discussed and interpreted in the context of forecasting.

The 2019 histograms which correspond to a normal year, Figs. 5(a) and 5(b), indicate that the prices for both markets occur mostly within the range of 30 to 60 (€/MWh). During the following year, 2020, see Figs. 6(a) and 6(b), the prices are significantly redistributed towards lower values, 0–30 and 30–60 (€/MWh). This is even more pronounced for the Finnish market as can be seen in Fig. 6(b). The redistribution is attributed to Covid-19 and the strict lock-downs which resulted in very low demand. What is also worth noting is the appearance of negative prices as can be seen in Figs. 6(a) and 6(b). The number of negative market price occurrences increased in 2020 in Germany from 211 in 2019 to 298 in 2020 (lowest market price in 2020 was −83.94 (€/MWh) on 21/4/2020 at 13:00), while in Finland negative market prices appeared for the first time (lowest market price was −1.73 (€/MWh) on Monday 2/11 at 03:00) and continued to appear in 2021 and 2022. Also, another important factor that played a role in increasing the negative prices in the German market in 2020 was renewable production. Specifically, pv production reached record levels.

The prices of 2021, Figs. 7(a) and 7(b), are distributed to higher values than those corresponding to years 2019 and 2020, Figs. 5(a) to 6(b). It is noteworthy that the highest price that occurred in the German market in 2021 was 620.00 (€/MWh), (Tuesday 21/12/2021 at 17:00), and in the Finnish market 100.07 (€/MWh), (Tuesday 07/12/2021 at 07:00). These extremely high prices are attributed to excessively high demand, lower production from renewables and higher gas prices. In addition, from January 1, 2021, the European Union increased the cost of certificates for emissions and thereby it is assumed that this factor also played a role in the price rising in 2021 [2].

The increasing trend of market prices in both the markets continues throughout 2022 as it is evident by the skewed to the right histograms in Figs. 8(a) and 8(b), with the German market exhibiting prices over 180 (€/MWh) for more than 5000 h and the Finnish market for 3000. The more frequent appearance of these higher prices can be attributed to the energy crisis.

Table 9
Price ranges determining normal price behaviour (€/MWh).

2019	2020	2021	2022
German market			
$10 \le Price \le 75$	$7 \le Price \le 70$	$2 \le Price \le 65$	$60 \le Price \le 250$
Finnish market			
$10 \le Price \le 70$	$1 \le Price \le 70$	$5 \le Price \le 120$	$10 \le Price \le 180$

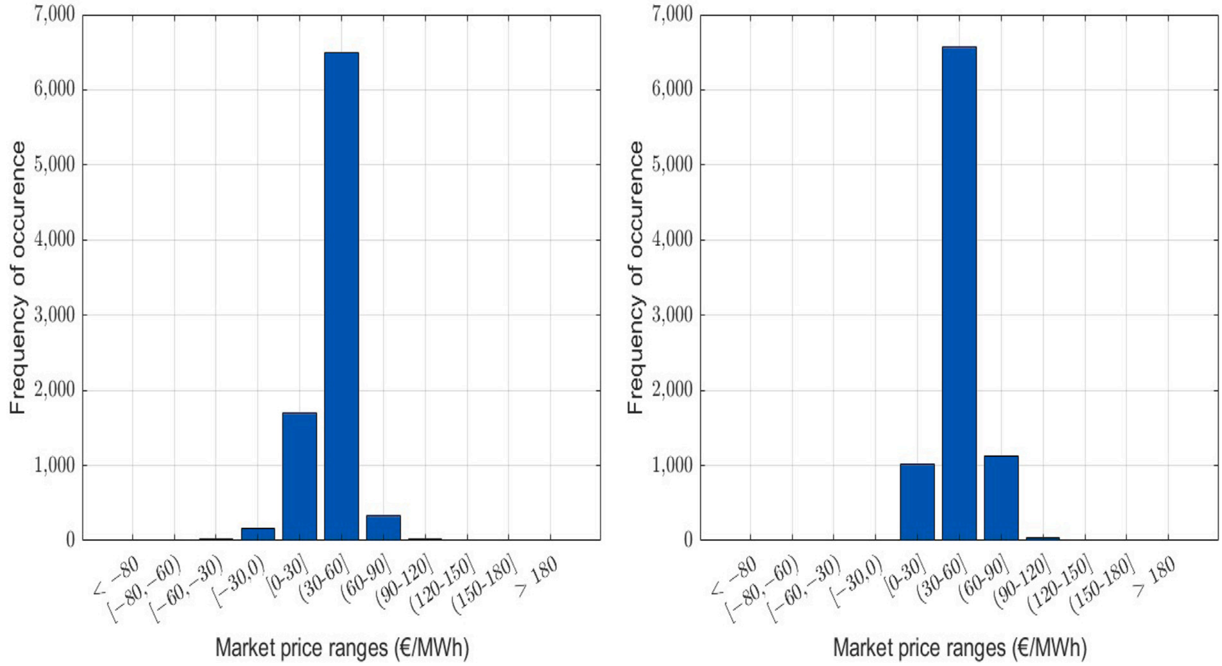
3.3. Threshold definition

The market prices were classified into three different categories, namely, normal, extremely high positive and negative market prices. In the literature a threshold is determined based on the expression, $Threshold = \mu \pm 2\sigma$ [73], where μ is the sample mean and σ the sample standard deviation. However, its use for the electricity markets under study allows negative price values to be considered as normal. For example, this expression gives, for the German market in 2021, the price −4.53 (€/MWh) as normal. In order to avoid situations where the range that is designated for normal prices includes negative prices the normal price range is determined by the number of price occurrences that appears in the histograms in Figs. 5(a)–8(b). It was decided that 10% of the prices of the whole year were normal prices. This is motivated by the precision theory in prevalence studies as described in [82]. According to this theory 10% sampling from population of 5000 and above gives an acceptable error margin of 3%. That is, a price range that appears in the histogram is considered normal if the number of price occurrences within that range exceeds 1000. Given that a year consists of 8760 h the value of 1000 indicates that prices lie in the range, to be considered as normal. In addition, in this way more samples for extremely high and negative prices are created which can better train the model.

The normal market price ranges for the German market are as follows: In 2019, the normal price range obtained is 10 to 75 (€/MWh). The histogram in Fig. 5(a) indicates that most of the prices are in the range of 30–60 (€/MWh) but the number of occurrences in the ranges 10–30 (€/MWh) and 60–75 (€/MWh) is such that these ranges are also considered normal. Using the histogram in Fig. 6(a) and by making a similar observation as in the case of 2019 the normal price range for the year 2020 is determined to be 7 to 70 (€/MWh) whereas for the year 2021 the normal price range obtained is 2–65 (€/MWh). The price histogram for the year 2022, as in Fig. 8(a), indicates that the majority of price occurrences are above 180 (€/MWh), with a significant number falling in the range 60–250 (€/MWh). Consequently, for the year 2022, the normal range for the German market is 60–250 (€/MWh) and 10–180 (€/MWh) for the Finnish market. The normal price ranges for the Finnish market were also chosen using similar reasoning upon information extracted from the corresponding histograms of Figs. 5(b), 6(b), 7(b) and 8(b). The normal price ranges obtained for each year are shown in Table 9.

In addition, negative prices in electricity markets can also be considered extreme. Supply flexibility depends on market conditions, forecasts, and previous hours' supply [83] and must be flexible for both, upwards and downwards. Non-flexible units and the intermittent nature of renewables contribute to negative prices since excess production from renewables cannot be stored [6]. This phenomenon can arise when renewables production exceeds demand [84–86].

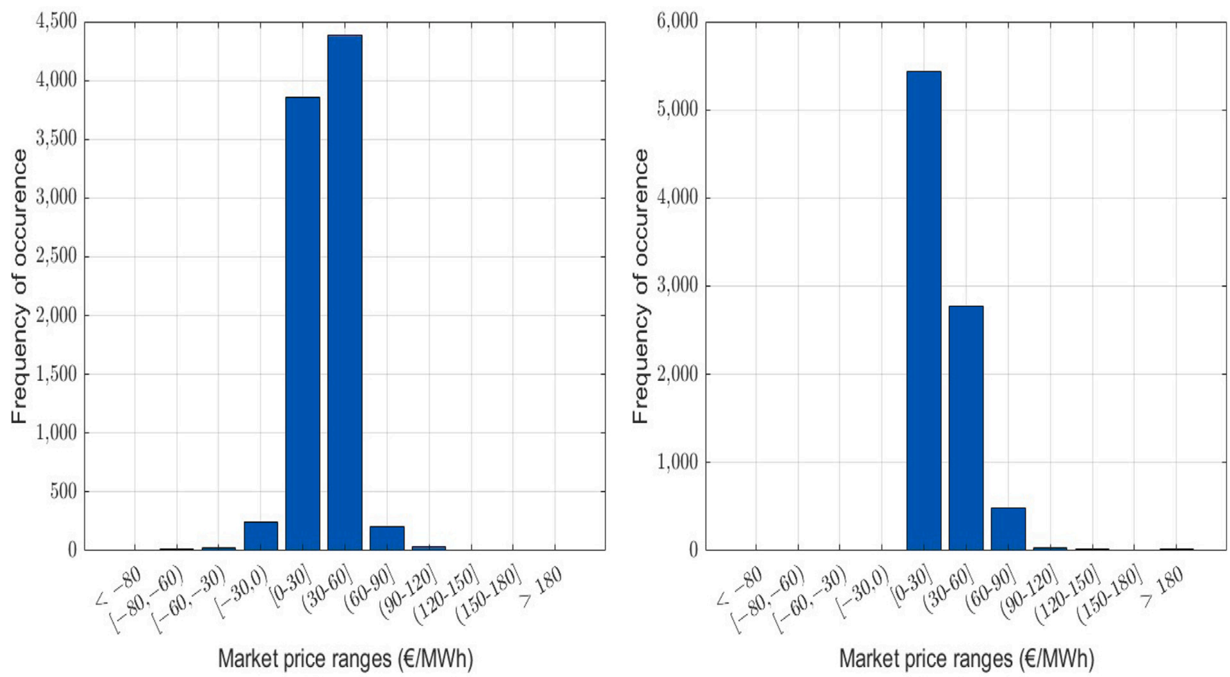
Producers provide incentives for consumers to consume the surplus through negative prices [83], without posing a threat to businesses, as their decision to maintain production signifies its economic viability [87,88]. Negative prices can also incentivize producers to invest in more flexible units [87].



(a) German Day-Ahead Market

(b) Finnish Day-Ahead Market

Fig. 5. Market price distribution for 2019 (before Covid-19).



(a) German Day-Ahead Market

(b) Finnish Day-Ahead Market

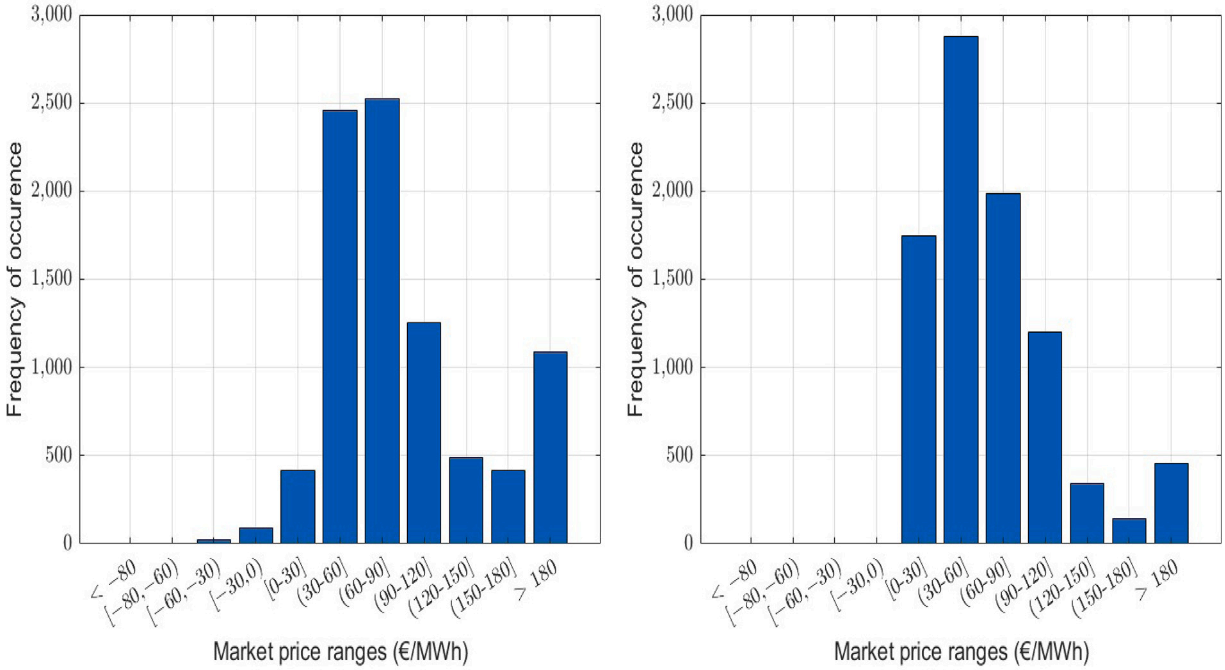
Fig. 6. Market price distribution for 2020 (during Covid-19).

4. Day-Ahead market forecasting results

In this section the methodology proposed in Section 2 is applied to forecast the next day prices in the German and Finnish Day-Ahead markets. The data, as described in Section 3.3, is separated into three classes labelled as normal, high and negative price and the methodology is tested for these three different classes.

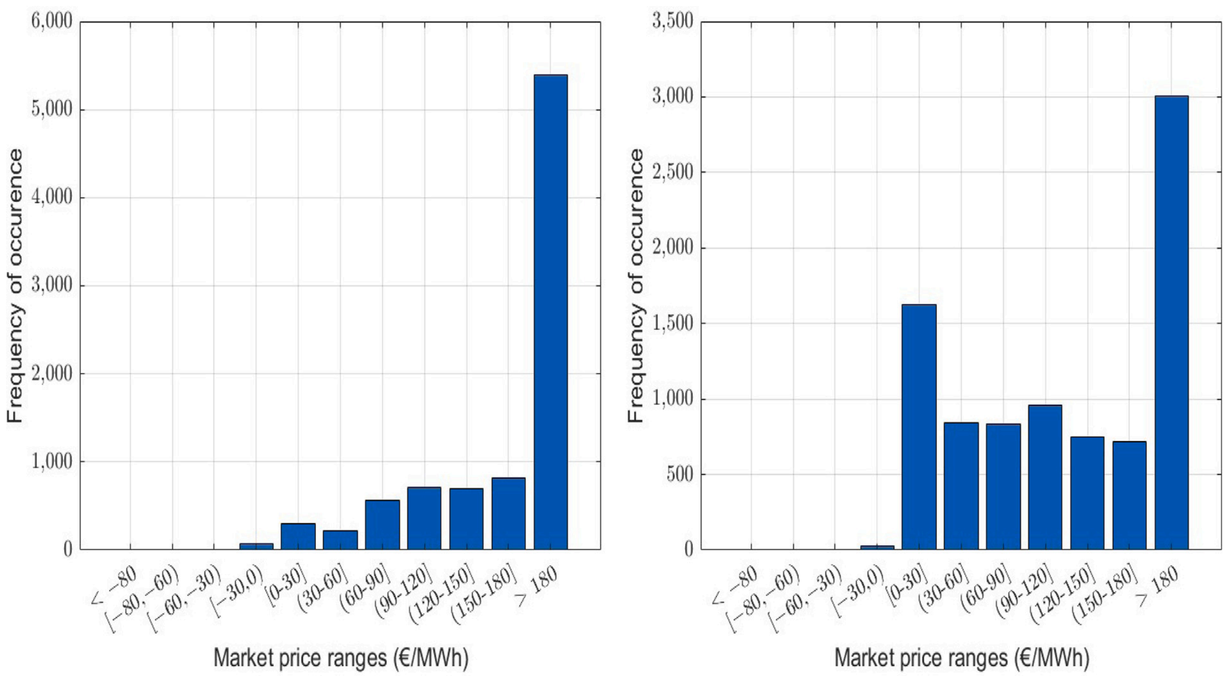
In addition, the proposed methodology was systematically compared with the GARCH and ARFIMA models that are well-established time series analysis models, which have been used in the field of financial economics for years. These models are commonly employed for forecasting the actual prices or volatility in Day-Ahead electricity markets [18,89–94], stock markets [95], and fuel markets [96]. They are often used individually or in combination with other forecasting approaches.

The ARFIMA model is an extension of the ARIMA model that incorporates fractional differencing. This feature is particularly valuable as it allows the modelling of time series having long memory or persistent



(a) German Day-Ahead Market (b) Finnish Day-Ahead Market

Fig. 7. Market price distribution for 2021 (post Covid-19).



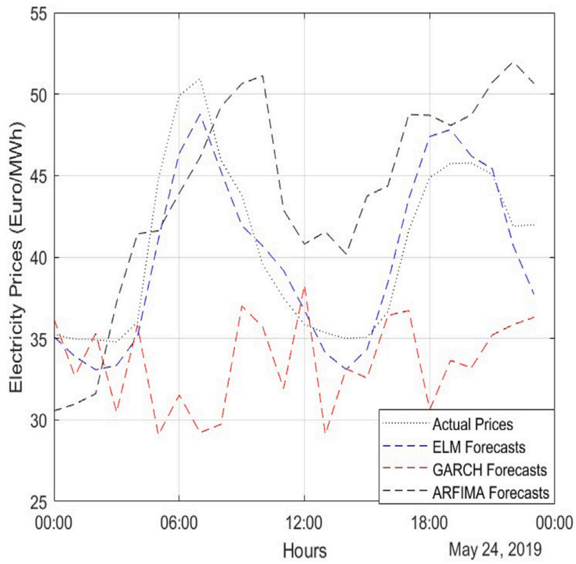
(a) German Day-Ahead Market (b) Finnish Day-Ahead Market

Fig. 8. Market price distribution for 2022 (Energy Crisis).

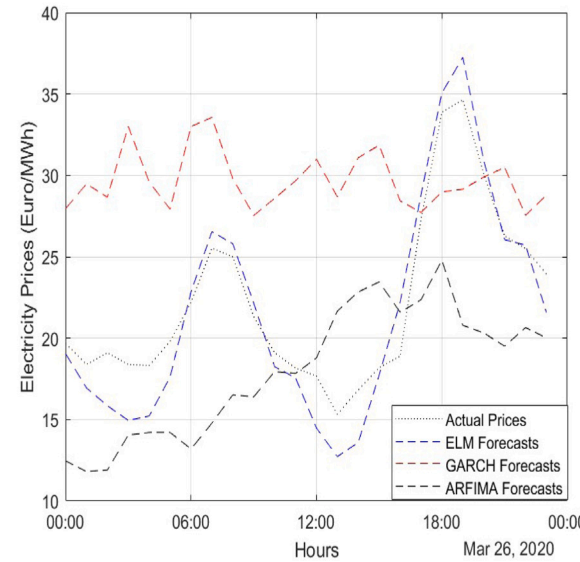
dependence. On the other hand, GARCH models extend ARCH models by introducing a moving average component for conditional variance. This enhancement enables GARCH models to capture the time-varying volatility that might be present in the time series under study.

The methodology presented in [18], which utilizes historical price data served as the basis for training the time series models. However, the logarithmic transformation employed in [18] was omitted to avoid complications with the negative prices. The orders p and q of the auto-regressive and moving average part, respectively, of the underlying ARMA model were determined by analysing of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. The value for both p and q orders obtained from this analysis was 6. The training process for the GARCH model utilized price data from the preceding 180 days of the target day.

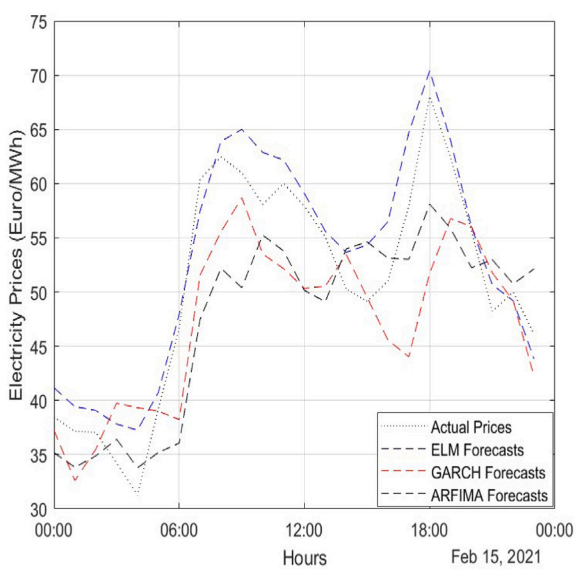
The fractional differencing exponent (d) of the ARFIMA model was estimated to be approximately 0.5. Additionally, using the AIC and BIC criteria [94], the values of the orders p and q were both found to be 2.



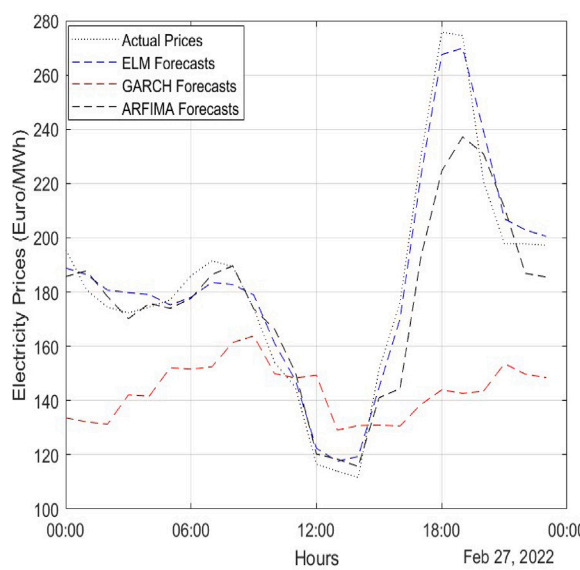
(a) ELM: MAE=1.65 (€/MWh), RMSE=1.94 (€/MWh), GARCH: MAE=7.12 (€/MWh), RMSE=8.90 (€/MWh), ARFIMA: MAE=5.60, RMSE=6.08 (€/MWh).



(b) ELM: MAE=1.96 (€/MWh), RMSE=2.03 (€/MWh), GARCH: MAE=8.21 (€/MWh), RMSE=8.96 (€/MWh), ARFIMA: MAE=5.01 (€/MWh), RMSE=6.77 (€/MWh).



(c) ELM: MAE=2.80 (€/MWh), RMSE=3.29 (€/MWh), GARCH: MAE=5.13 (€/MWh), RMSE=6.38 (€/MWh), ARFIMA: MAE=5.51 (€/MWh), RMSE=6.39 (€/MWh).



(d) ELM: MAE=6.52 (€/MWh), RMSE=7.11 (€/MWh), GARCH: MAE=12.69 (€/MWh), RMSE=14.24 (€/MWh), ARFIMA: MAE=11.87 (€/MWh), RMSE=12.72 (€/MWh).

Fig. 9. Days with normal prices, German market, 2019–2022.

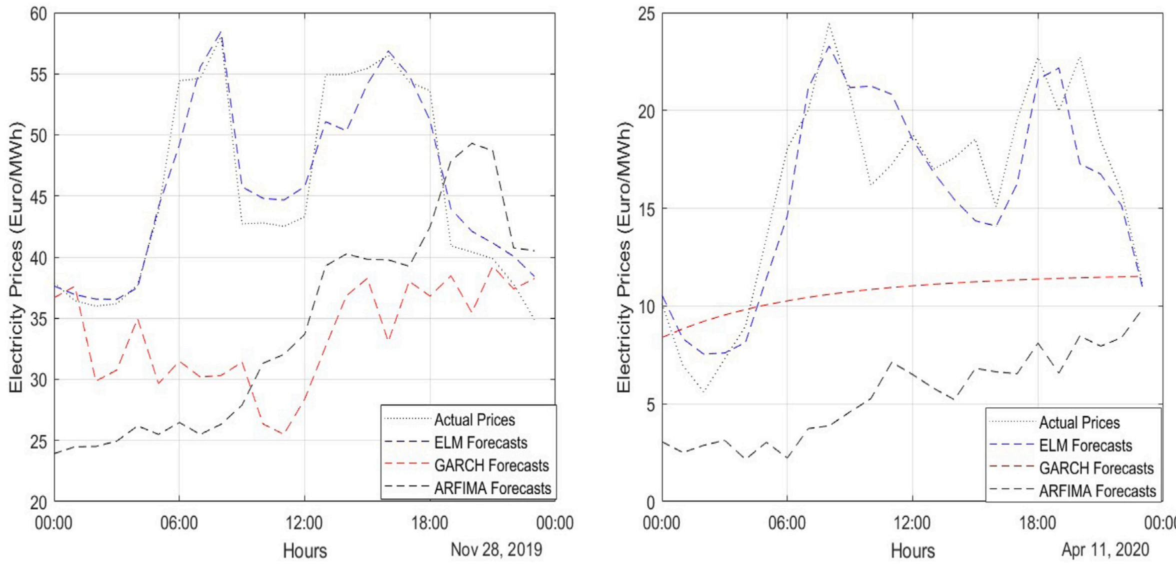
Furthermore, during the training of the ARFIMA model, data from the preceding 60 days leading up to the target day was utilized.

4.1. Normal market price results

The forecasting results for typical normal price days for the German market 2019–2022, are shown in Figs. 9(a)–9(d) and the corresponding ones for the Finnish market in Figs. 10(a)–10(d). The forecasted price values obtained using the proposed ELM based methodology for the cases of 2019–2021 are very close to the actual ones. The MAE and RMSE is below 5 (€/MWh), with the lowest MAE (RMSE) for the German-2019 and Finnish-2019 market being 1.65 (1.94) (€/MWh) and 1.81 (2.32) (€/MWh) respectively. For the case of 2022 the errors are higher than 5, German market: 5.02 (6.11) (€/MWh), Finnish market: 5.55 (6.30) (€/MWh).

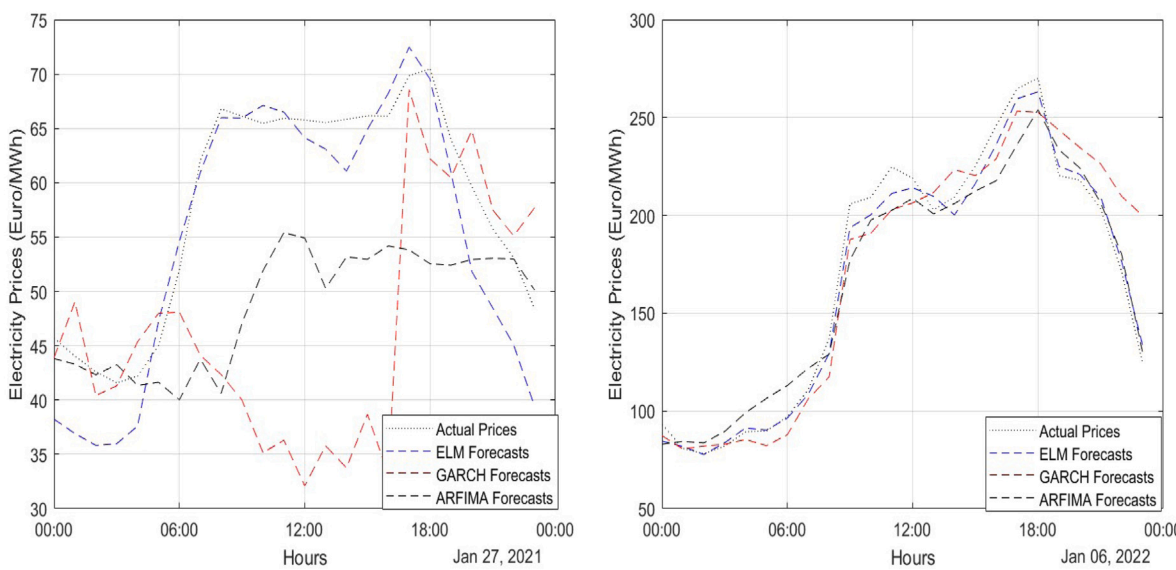
In contrast, the statistical models exhibited higher forecasting errors for both markets. Specifically, the GARCH model yielded MAE and RMSE values ranging between 5.13–12.69 (€/MWh) and 6.38–14.24 (€/MWh), respectively, for the German market. Similarly, for the Finnish market, the GARCH model yielded errors in the range of 7.12–14.97 (MAE) (€/MWh) and 8.82–16.31 (RMSE) (€/MWh). The forecasting errors obtained using the ARFIMA model were 5.60–11.87 (MAE) (€/MWh) and 6.08–12.72 (RMSE) (€/MWh) for the German market, and 9.55–13.99 (MAE) (€/MWh) and 11.03–15.02 (RMSE) (€/MWh) for the Finnish market.

The twenty-four-hour price behaviour, 2019–2021 in the Finnish market during a day of normal prices is characterized by three distinct



(a) ELM: MAE=1.81 (€/MWh), RMSE=2.32 (€/MWh), GARCH: MAE=10.76 (€/MWh), RMSE=12.08 (€/MWh), ARFIMA: MAE=13.99 (€/MWh), RMSE=15.02 (€/MWh).

(b) ELM: MAE=1.84 (€/MWh), RMSE=2.39 (€/MWh), GARCH: MAE=7.12 (€/MWh), RMSE=8.82 (€/MWh), ARFIMA: MAE=10.67 (€/MWh), RMSE=11.62 (€/MWh).



(c) ELM: MAE=3.82 (€/MWh), RMSE=4.72 (€/MWh), GARCH: MAE=12.66 (€/MWh), RMSE= 14.29 (€/MWh), ARFIMA: MAE=9.55 (€/MWh), RMSE=11.03 (€/MWh).

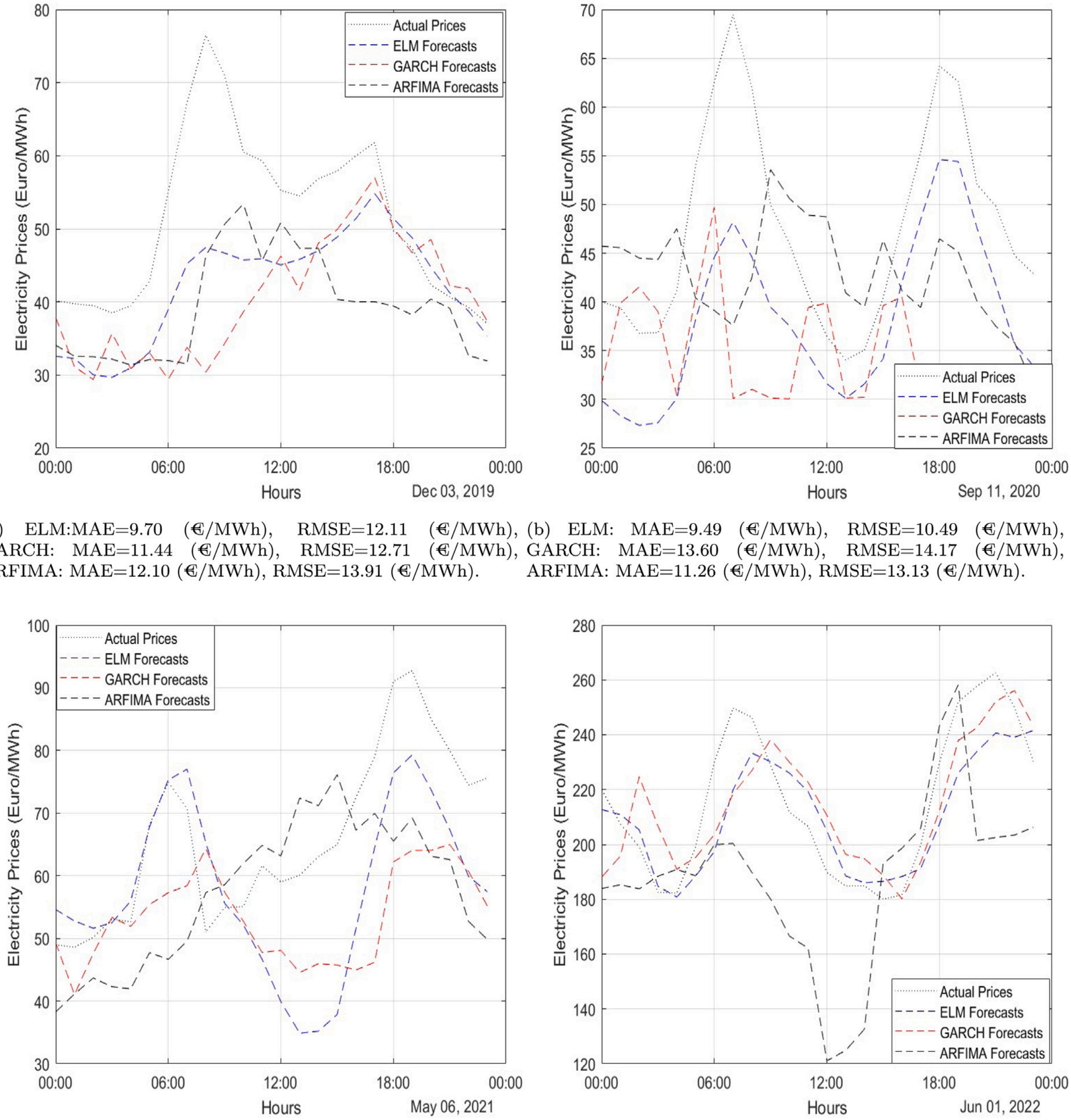
(d) ELM: MAE=5.82 (€/MWh), RMSE=6.61 (€/MWh), GARCH: MAE=14.97 (€/MWh), RMSE=16.31 (€/MWh), ARFIMA: MAE=12.04 (€/MWh), RMSE=14.37 (€/MWh).

Fig. 10. Days with normal prices, Finnish market, 2019–2022.

time periods. For the first seven hours of the day the price increases steadily followed by a second twelve hour period of a price plateau. Then for the remaining hours of the day the prices drop steadily. This is in contrast to the behaviour observed in the German market where within a day two peak prices are observed, one between 06:00 and 12:00 and the other one between 17:00 and 19:00. The proposed methodology in both markets not only captures the overall behaviour but also forecasts values very close to the actual ones. This means that under normal price conditions the algorithm can be generalized well across different markets. An exception is the 2022 price behaviour in the Finnish market. This is illustrated in Fig. 10(d), where the three distinct time periods do not exist. Additionally, around 16:00 a spike appeared.

4.2. Extremely high market price results

All the days of 2019–2022 that exhibited extremely high prices in both the German and Finnish markets were studied. Indicative forecasting results concerning such days are shown in Figs. 11(a)–11(d) for the German market and in Figs. 12(a)–12(d) for the Finnish market. The MAE (RMSE) errors, corresponding to the proposed ELM based methodology, for the German market are 9.49 (10.49) (€/MWh) whereas for the Finnish market are 7.82 (9.42) (€/MWh). In the German market, 2019–2022 the proposed methodology predicts correctly the time of occurrence of the peak values which are usually isolated prices and one appears in the morning between 06:00–09:00 and the other in the afternoon between 17:00–19:00. In addition, there are a few cases where the high prices were prolonged for some time period. Overall,



(a) ELM:MAE=9.70 (€/MWh), RMSE=12.11 (€/MWh), GARCH: MAE=11.44 (€/MWh), RMSE=12.71 (€/MWh), ARFIMA: MAE=12.10 (€/MWh), RMSE=13.91 (€/MWh).

(b) ELM: MAE=9.49 (€/MWh), RMSE=10.49 (€/MWh), GARCH: MAE=13.60 (€/MWh), RMSE=14.17 (€/MWh), ARFIMA: MAE=11.26 (€/MWh), RMSE=13.13 (€/MWh).

(c) ELM:MAE=11.42 (€/MWh), RMSE=13.31 (€/MWh), GARCH: MAE=13.81 (€/MWh), RMSE=14.97 (€/MWh), ARFIMA: MAE=13.40 (€/MWh), RMSE=15.44 (€/MWh).

(d) ELM: MAE=12.20 (€/MWh), RMSE=14.26 (€/MWh), GARCH: MAE=14.90 (€/MWh) RMSE=16.53 (€/MWh), ARFIMA: MAE=14.19 (€/MWh), RMSE=15.69 (€/MWh).

Fig. 11. Days with extremely high prices, German market, 2019–2022.

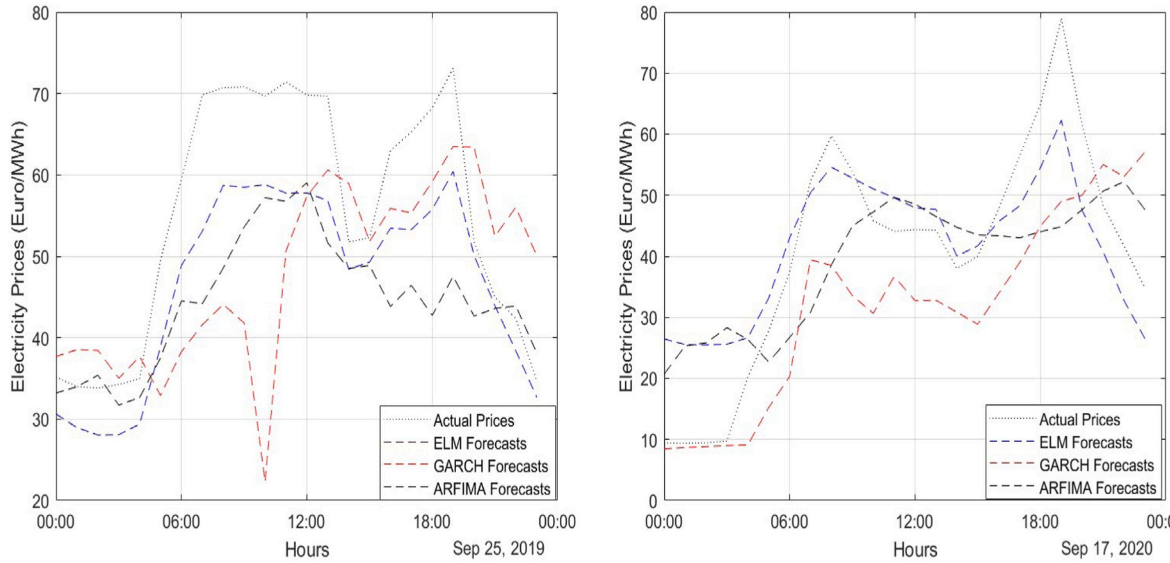
for days that extremely high prices appear in the German market, the methodology forecasts lower prices.

In the Finnish market, 2019–2022 extremely high prices were observed to occur either for a prolonged period starting at around 07:00 and lasting up until around 13:00, or isolated at either around 06:00 or occurring within a day both for a prolonged period and isolated. The proposed methodology, even though in some cases predicts higher price values, Fig. 12(c), in most of the cases predicts lower values.

As for the case of the normal price class, the time series models generated forecasts with higher errors for the class of extremely high prices. Specifically, the GARCH model yielded MAE and RMSE values ranging between 11.44–14.90 (€/MWh) and 12.71–16.53 (€/MWh), respectively, for the German market. In the case of the Finnish market, the GARCH model produced errors within the range of 11.17–15.06 (MAE) (€/MWh) and 12.59–16.61 (RMSE) (€/MWh). Meanwhile, the ARFIMA model exhibited errors between 11.26–14.19 (MAE) (€/MWh) and 13.13–15.69 (RMSE) (€/MWh) for the German market, and 11.14–15.48 (MAE) (€/MWh) and 14.13–17.09 (RMSE) (€/MWh) for the Finnish market.

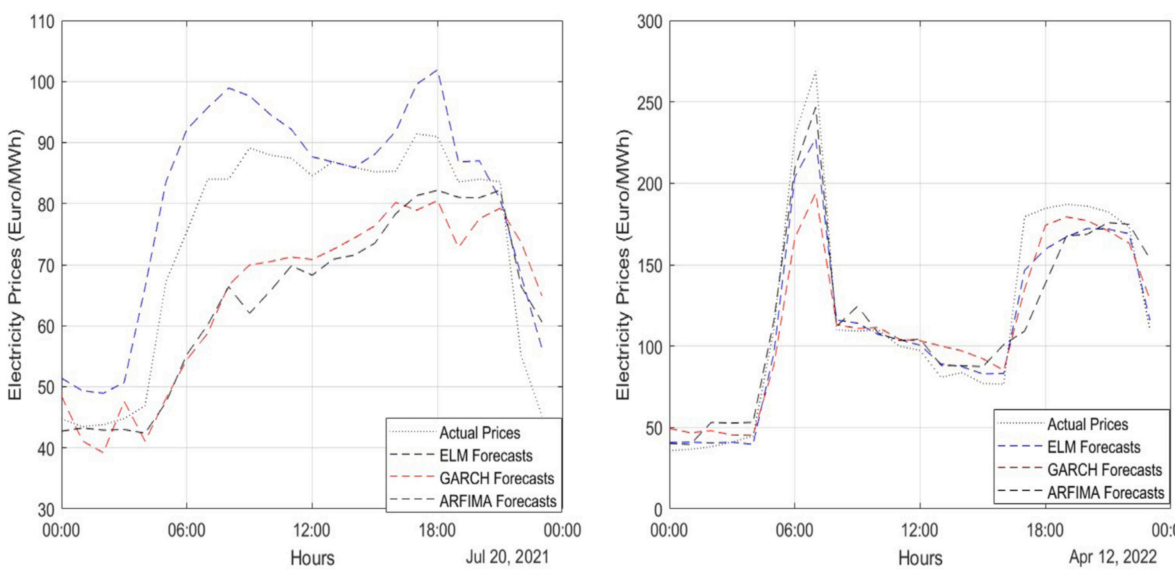
4.3. Negative market price results

Similarly to the case of extremely high prices, the study of negative prices focuses to the days that negative prices appeared for the German and Finnish markets from 2019 to 2022. Figs. 13(a)–13(d) depict the results of the days with negative prices in Germany. It is worth noting that there were no instances of negative prices in the Finnish market before 2020, so only indicative results for the years 2020–2022 are shown in Figs. 14(a)–14(c) respectively.



(a) ELM: MAE=8.35 (€/MWh), RMSE=9.48 (€/MWh), GARCH: MAE=12.97 (€/MWh), RMSE=13.48 (€/MWh), ARFIMA: MAE=11.14 (€/MWh), RMSE=14.13 (€/MWh).

(b) ELM: MAE=7.84 (€/MWh), RMSE=9.45 (€/MWh), GARCH: MAE=11.92 (€/MWh), RME=13.89 (€/MWh), ARFIMA: MAE=11.31 (€/MWh), RMSE=13.76 (€/MWh).



(c) ELM: MAE=7.82 (€/MWh), RMSE=9.42 (€/MWh), GARCH: MAE=11.17 (€/MWh), RMSE=12.59 (€/MWh), ARFIMA: MAE=11.48 (€/MWh), RMSE=14.01 (€/MWh).

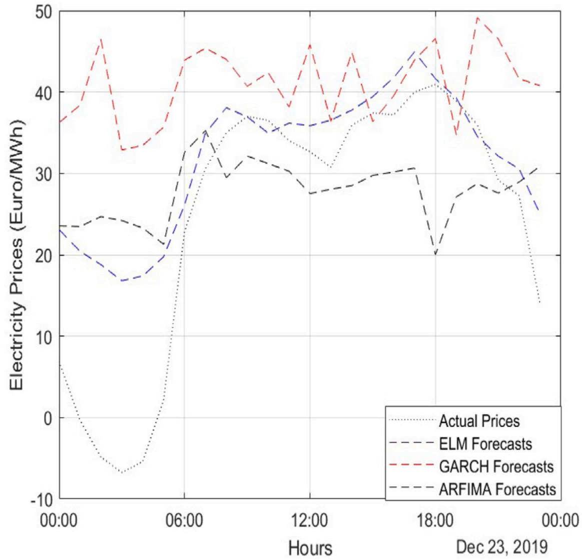
(d) ELM: MAE=10.65 (€/MWh), RMSE=12.27 (€/MWh), GARCH: MAE=15.06 (€/MWh), RMSE=16.61 (€/MWh), ARFIMA: MAE=15.48 (€/MWh), RMSE=17.09 (€/MWh).

Fig. 12. Days with extremely high prices, Finnish market, 2019–2022.

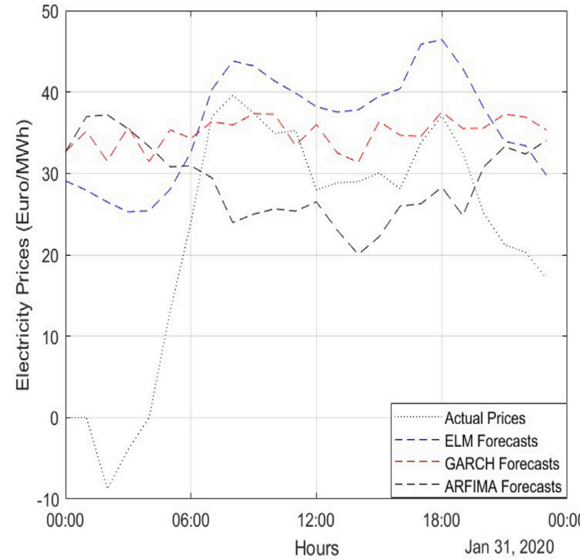
Also, the prediction model inherently accommodates negative data without necessitating specific modifications. It is designed to adeptly process and generate forecasts for a diverse range of price fluctuations, including negative values, without the need for tailored adjustments. This intentional design choice aims to showcase the dynamic adaptability of our methodology, particularly in handling the nuances associated with this specific price class.

The negative prices observed in the German market span a prolonged period from 2019 to 2022 and are not isolated occurrences or spikes. They are most prominent in the early morning hours and on weekends, but negative prices can also occur on weekdays at different times throughout the day. Another notable characteristic of the German market is the significant price dip observed between 11:00 and 16:00, as depicted in Fig. 13(c) or Fig. 13(d). Furthermore, in the Finnish market, the negative prices observed in 2020 occurred exclusively on Mondays between 00:00 and 04:00. Similarly, in 2021, there was only one day (Monday) with negative prices during the same time frame. In 2022 there were six days with negative prices, 6 to 8 of October, 11 and 12 of November and 31 of December. The 6 of October was Thursday, while the other days in which negative prices appeared were either Fridays or Saturdays. Additionally, as depicted in Figs. 14(a)–14(c), the negative prices in the Finnish market were close to zero.

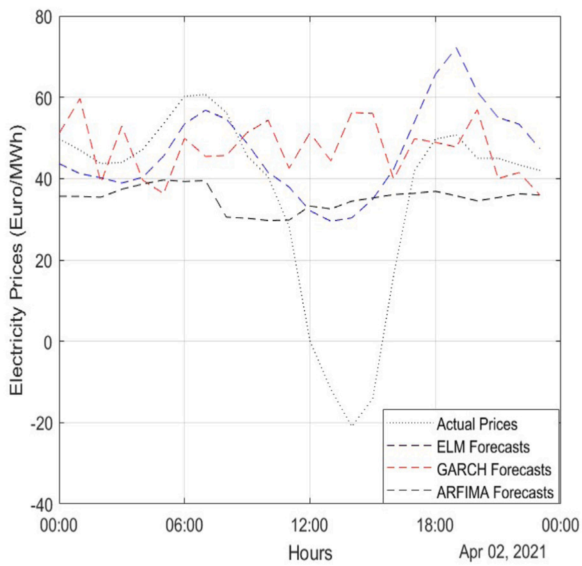
The MAE and RMSE errors corresponding to the ELM based proposed methodology are also higher than those occurring for days of normal and extremely high prices. Namely, MAE (RMSE) 7.54–15.41 (11.03–16.68) (€/MWh) for the German market and 15.89–18.19 (16.64–19.24) (€/MWh) for the Finnish market. The GARCH model exhibited MAE and RMSE values ranging between 14.89–17.93 (€/MWh) and 15.78–20.34 (€/MWh) for the German market, while for the Finnish market, the range was 15.52–22.63 (MAE) (€/MWh) and



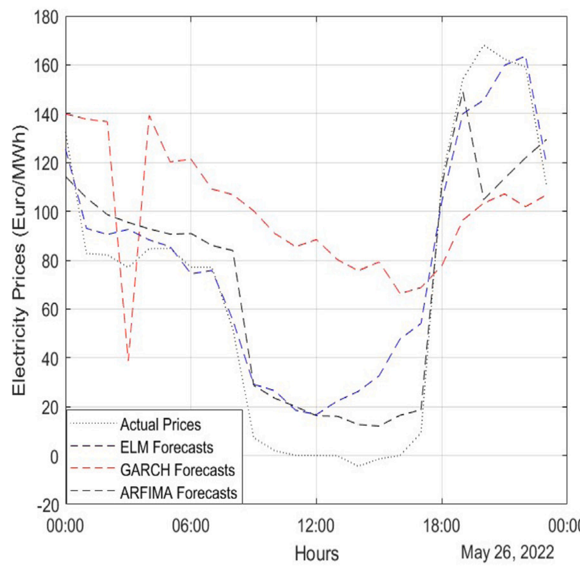
(a) ELM: MAE=7.54 (€/MWh), RMSE=11.03 (€/MWh), GARCH: MAE=15.88 (€/MWh), RMSE=17.93 (€/MWh), ARFIMA: 11.74 (€/MWh), RMSE=14.83 (€/MWh).



(b) ELM: MAE=13.72 (€/MWh), RMSE=16.24 (€/MWh), GARCH: MAE=14.89 (€/MWh), RMSE=15.78 (€/MWh), ARFIMA: MAE=15.19 (€/MWh), RMSE=17.57 (€/MWh).



(c) ELM: MAE=14.08 (€/MWh), RMSE=15.23 (€/MWh), GARCH: MAE=17.93 (€/MWh), RMSE=20.34 (€/MWh), ARFIMA: MAE=17.40 (€/MWh), RMSE=19.12 (€/MWh).



(d) ELM: MAE=15.41 (€/MWh), RMSE=16.68 (€/MWh), CARCH: MAE=22.52 (€/MWh), RMSE=24.85 (€/MWh), ARFIMA: MAE=19.66 (€/MWh), RMSE=21.97 (€/MWh).

Fig. 13. Days with negative prices, German market, 2019–2022.

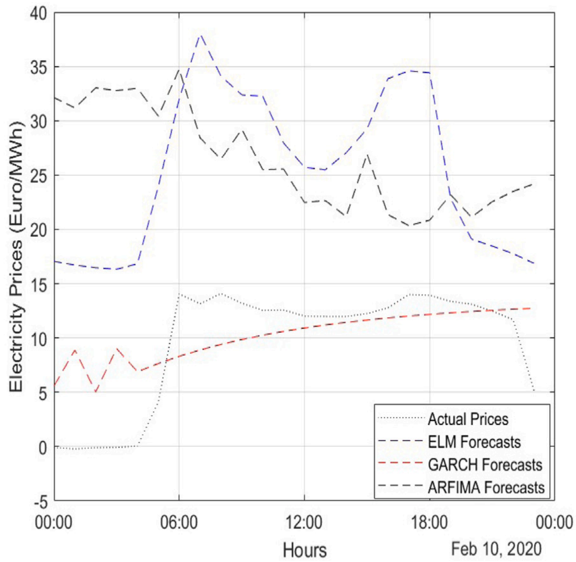
16.97–23.73 (RMSE) (€/MWh). On the other hand, the ARFIMA model demonstrated errors of 11.74–15.48 (MAE) (€/MWh) and 14.83–21.97 (RMSE) (€/MWh) for the German market, and 16.86–20.16 (MAE) (€/MWh) and 19.26–22.81 (RMSE) (€/MWh) for the Finnish market. Once again, the proposed methodology exhibited superior performance compared to both the GARCH and ARFIMA models.

The methodology does not predict the occurrence of negative prices but for the case of the German market, the prediction of the non-negative price values is close to the actual ones. The negative price values in the Finnish market are not isolated but they appear for a period of time, usually at the hours 00:00 to 04:00. This feature does not normally happen and it is difficult to be forecasted.

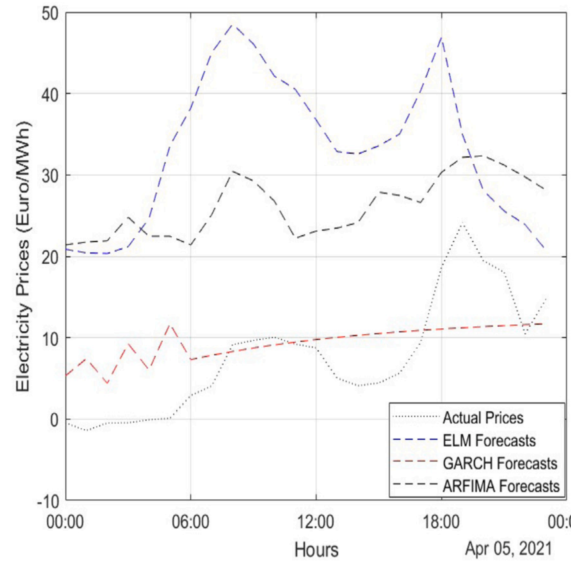
Negative prices indicate excess production usually due to increased production from renewable sources [84]. Additionally, a different interpretation of negative prices is provided in the Ref. [6]. According to the authors, negative prices can occur in two scenarios: when wind generation is moderate and the load is low, or when wind generation is high and the load is low. This interpretation suggests a lack of flexibility in the electricity system. In order to improve the price forecasts for extremely high prices and negative prices, the bootstrap intervals that are described in the next section can be used.

4.4. Bootstrap intervals

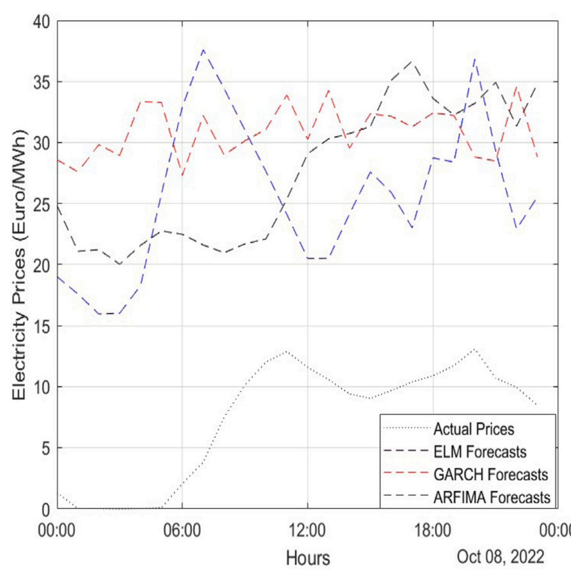
The bootstrap methodology, Section 2.2, generates 1000 forecasts for each hour and the actual forecast is given by their average value. These forecasts are considered to lie in an interval determined by the lowest and highest generated values which for the purposes of the work herein is referred to as bootstrapping interval. Examples of these intervals for days of normal prices in the German and Finish markets



(a) ELM: MAE=15.89 (€/MWh), RMSE=16.64 (€/MWh), GARCH: MAE=15.52 (€/MWh), RMSE=16.97 (€/MWh), ARFIMA: MAE=16.86 (€/MWh), RMSE=19.28 (€/MWh).



(b) ELM: MAE=17.88 (€/MWh), RMSE=20.03 (€/MWh), GARCH: MAE=18.40 (€/MWh), RSME=19.44 (€/MWh), ARFIMA: MAE=18.38 (€/MWh), RMSE=20.23 (€/MWh).



(c) ELM: MAE=18.19 (€/MWh), RMSE=19.24 (€/MWh), GARCH: MAE=22.63 (€/MWh), RMSE=23.73 (€/MWh), ARFIMA: MAE=20.16 (€/MWh), RMSE=22.81 (€/MWh).

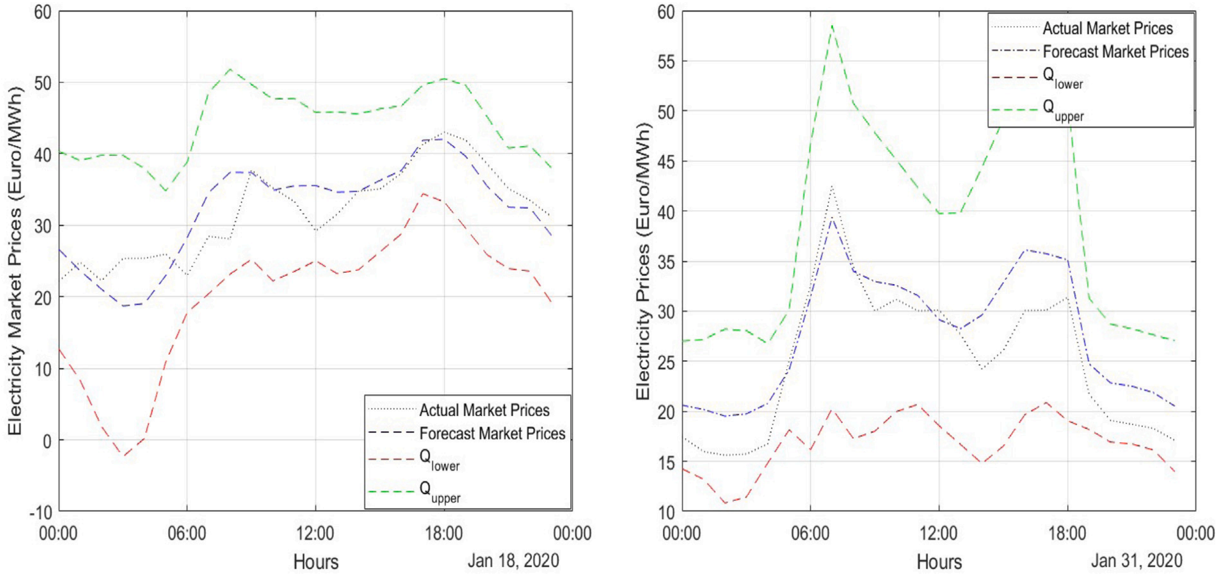
Fig. 14. Days with negative prices, Finnish market, 2019–2022.

are shown in Figs. 15(a) and 15(b) respectively, whereas, for days with extremely high prices are shown in Figs. 16(a) and 16(b) and for days with negative prices in Figs. 17(a) and 17(b).

As anticipated, the actual price values of days with normal price values are entirely within the bootstrapping intervals, as depicted in Figs. 15(a) and 15(b). The same holds true for the German market, where days with extremely high prices are represented in Fig. 16(a). However, in the case of the Finnish market shown in Fig. 16(b), the actual values during the first hours of the day fall outside the bootstrapping interval. Specifically, these values are lower than the lower bound of the interval. It is worth noting that the prices that determine the upper range of the bootstrapping interval for the extremely high prices in the German market, as illustrated in Fig. 16(a), are in closer proximity to the actual price spikes.

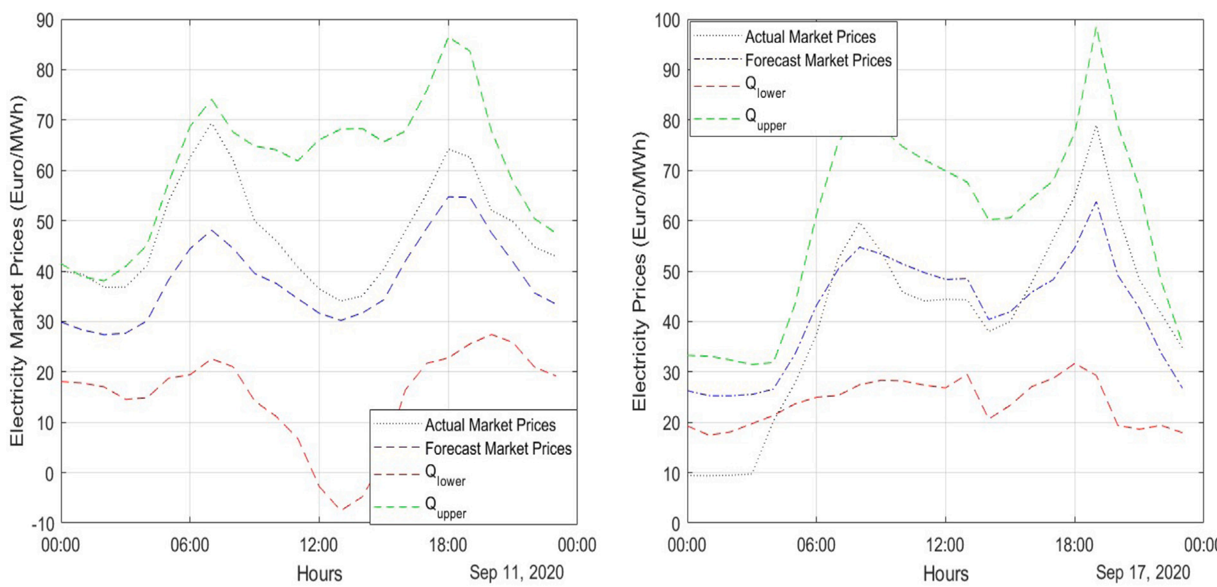
The negative prices do not lie in the bootstrap intervals. In the case of the German market, Fig. 17(a), even though the lower part of the interval includes negative prices the actual value is much more negative than the lower limit of the interval. In the Finnish market, Fig. 17(b), the actual prices are closer to the lower limit of the bootstrap interval.

In conclusion, it can be inferred that the bootstrap intervals provide a reliable expected range within which the actual prices are observed. It is noteworthy that the prices do not surpass the upper limit of the interval, even for days with extreme price values. Conversely, for days with negative prices, the actual prices tend to be closer to the lower part of the bootstrap interval. So, the effectiveness of the bootstrapping is that it can fully capture the normal prices, while for the extremely high or negative prices the upper limit and the lower limit respectively are closest to them. The bootstrap intervals offer a potential application for forecasting the values of negative and extremely positive prices.



(a) German market.
(b) Finnish market.

Fig. 15. Days with normal prices.



(a) German market.
(b) Finnish market.

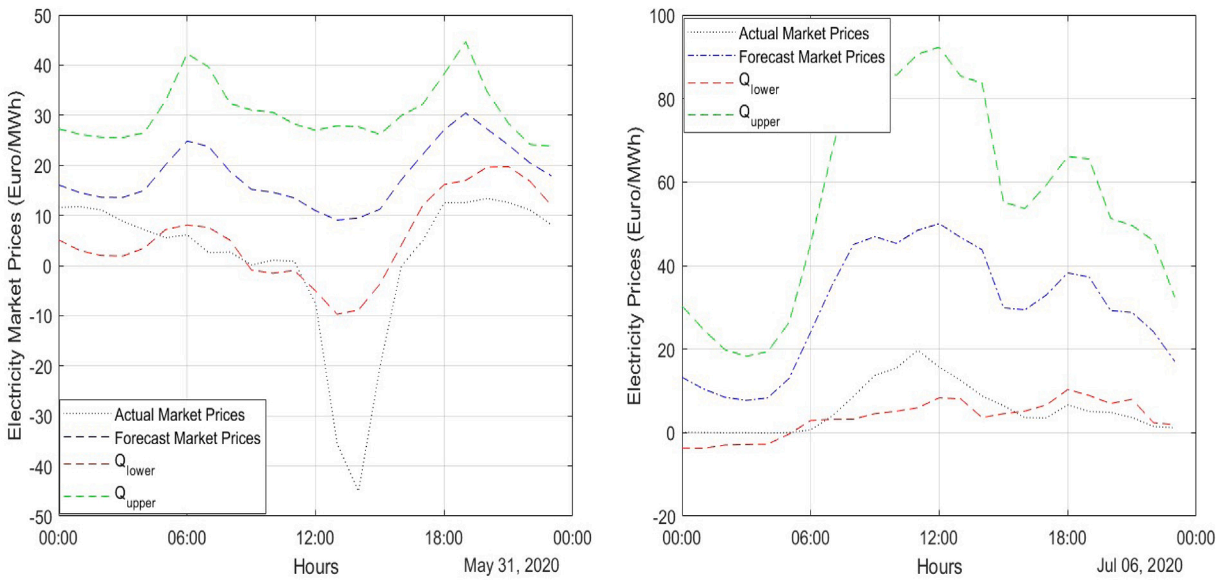
Fig. 16. Days with extremely high prices.

4.5. Time and error performance comparison with Artificial Neural Network (ANN) of the same architecture

This section presents the results of comparing the ELM-based methodology with the same methodology that utilizes an Artificial Neural Network (ANN) instead of the ELM. The comparison aims to assess the time and error performance differences between the two approaches as was done in another of our work [97]. The selection of an ANN for comparison was motivated by its widespread usage in the literature. Additionally, the choice of ANN was influenced by its architecture and the gradient descent algorithm that it uses during learning. Specifically, the execution time and resulting mean square errors of the two methods were compared. This analysis provides insights into the computational efficiency and accuracy of both the ELM-based methodology and the methodology employing an ANN.

An ANN with three layers was employed, where the hidden layer consisted of 20 neurons. This ANN architecture is similar to that of the ELM. The average results for normal, extremely high, and negative price values in the year 2019, which was considered a normal year without significant price fluctuations due to unforeseen health and socio-political events, are presented in Table 10.

The execution time for normal price days in a year is longer compared to days with extremely high and negative prices. This can be attributed to the higher number of normal price days compared to the occurrence of extremely high or negative price events within a year. As shown in Table 10, the execution time and performance of the ANN



(a) German market.

(b) Finnish market.

Fig. 17. Days with negative prices.

Table 10
Average Yearly RMSE (€/MWh) for ELM and ANN based methodologies.

German wholesale Day Ahead market, 2019				
	ELM		ANN	
	Time (min, ss)	Average RMSE (€/MWh)	Time (min, ss)	Average RMSE (€/MWh)
Normal prices	135,00	7.01	210,00	10.15
Extremely high prices	4,15	16.62	16,30	16.94
Negative prices	14,40	21.02	54,32	26.41

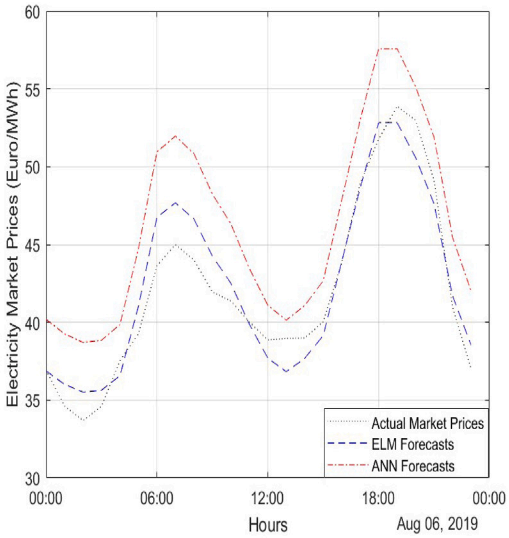
are inferior to those of the ELM method. These results validate the rationale behind selecting the ELM-based methodology as the preferred approach.

In addition, Figs. 18(a) to 18(c) represent the market price forecasts generated by both the ELM and ANN models. It can be seen that the ELM based methodology consistently outperforms the ANN based methodology across all three classes. The corresponding RMSE (€/MWh) for the ELM and ANN are detailed below: Normal prices: ELM: 1.62, ANN: 4.52. Extremely high prices: ELM: 7.94, ANN: 11.43 and Negative prices: ELM: 13.04, ANN: 15.28.

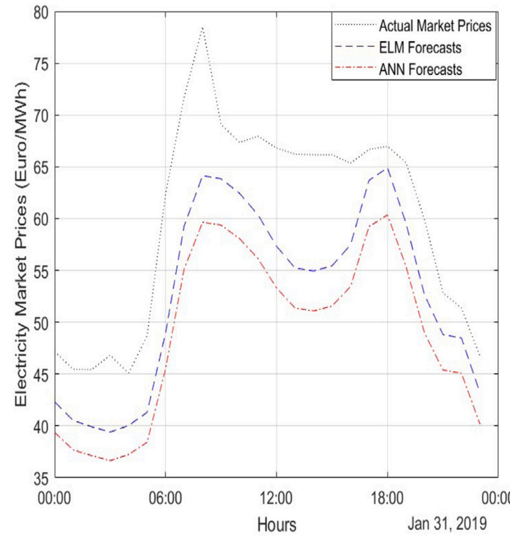
5. Conclusions

The Day-Ahead market is characterized by multiple and diverse sources of uncertainty, making price forecasting a challenging task. In this study, a forecasting methodology based on Extreme Learning Machine (ELM) and bootstrapping is proposed. The analysis begins by examining the German and Finnish Day Ahead market prices to identify the distinct price ranges associated with normal, extremely high, and negative prices. The developed forecasting methodology is then applied to these markets. The main conclusions drawn from the study are as follows:

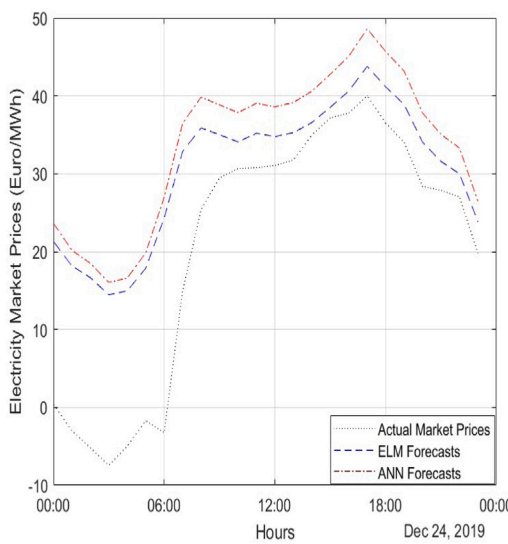
- The electricity energy prices exhibit unexpected changes which expose the market participants to high risks. This is due to the fact they are influenced by many interrelated factors such as fuel prices, demand, renewable energy production, unexpected socio-political events and health events and inability of electrical energy storage in high quantities.
- The prices are divided into three classes, normal, extremely high and negative. Normal price ranges are derived based on histograms. Extremely high prices are defined as prices that exceed the upper limit of the range of normal prices and negative prices are defined as prices that are smaller than zero.
- The basic learning element of the proposed methodology applied is the extreme learning machine. The methodology, due to ELM's low run-time complexity, instead of utilizing one ELM for the whole training sample it utilizes one ELM for each sample.
- The results for normal price days are excellent with very low errors across the years 2019, 2020, and 2021 for both the German and the Finnish markets. On the contrary the results for days with extremely high prices for the German market, the price forecasts of the actual spike prices differ from the actual ones. Usually, the methodology gives forecasts lower than the actual prices. In the Finnish market, the qualitative behaviour of the prices changes on days of extremely high prices causing the algorithm to perform slightly worse than in the German market. A similar observation is made with forecasting on days with negative price occurrences.
- The bootstrap intervals, which can be considered as indicators of the range of predictions, exhibited the ability to capture extremely high or negative prices. The upper limit is very close to the observed extremely high prices and similarly the lower limit is very close to negative prices. These observations could be utilized to improve forecasts of extremely high and negative prices.
- The proposed methodology was systematically compared with established models such as GARCH and ARFIMA, demonstrating consistently superior performance across all price classes.
- The ELM has lower execution time and better performance than the ANN. This is because the computation of the pseudo-inverse during ELM learning requires less computing time than the gradient descent based learning of ANN. The ELM based methodology performed better in terms of the MSE. Given that the Day Ahead market gate closes at 12:00, the ELM can be used in real-time scenarios to help market participants generate better bids.



(a) Normal Prices: ELM: RMSE=1.62 (€/MWh), ANN: RMSE=4.52.



(b) Extremely high prices: ELM: RMSE=7.94 (€/MWh), ANN: RMSE=11.43 (€/MWh).



(c) Negative prices: ELM: RMSE=13.04 (€/MWh), ANN: RMSE=15.28 (€/MWh).

Fig. 18. Performance comparison between ELM and ANN architecture.

CRediT authorship contribution statement

Stylianos Loizidis: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Conceptualization. **Andreas Kyprianou:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **George E. Georghiou:** Writing – review & editing, Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

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).

References

[1] Loizidis Stylianos, Theocharides Spyros, Venizelou Venizelos, Evagorou Demetris, Makrides Giorgos, Kyprianou Andreas, Georghiou George E. Extreme supervised algorithm for day ahead market price forecasting. In: 2023 IEEE international smart cities conference (ISC2), Bucharest, Romania, 24-27 September. 2023, p. 1–7. http://dx.doi.org/10.1109/ISC257844.2023.10293566.

[2] Loizidis Stylianos, Konstantinidis Georgios, Theocharidis Spyros, Kyprianou Andreas, Georghiou George E. Market conditions and their effect on market clearing price forecasting. In: 2022 2nd international conference on energy transition in

the mediterranean area (SyNERGY MED), Thessaloniki, Greece, 17–19 October. 2022, p. 1–6. http://dx.doi.org/10.1109/SyNERGYMED55767.2022.9941425.
[3] N. Van Aarde C Masinga. Quarterly report on European electricity markets, with focus on developments in annual wholesale prices. 2013;14:1–11.
[4] Zhong Haiwang, Tan Zhenfei, He Yiliu, Xie Le, Kang Chongqing. Implications of COVID-19 for the electricity industry: A comprehensive review. CSEE J Power Energy Syst 2020;6(3):489–95. http://dx.doi.org/10.17775/CSEEJPES.2020.02500.
[5] Navon Aviad, Machlev Ram, Carmon David, Onile Abiodun Emmanuel, Velikov Juri, levron Yoash. Effects of the COVID-19 pandemic on energy systems and electric power grids—A review of the challenges ahead. Energies 2021;14:1–14. http://dx.doi.org/10.3390/en14041056.
[6] Kyritsis Evangelos, Andersson Jonas, Serletis Apostolos. Electricity prices, large-scale renewable integration, and policy implications. Energy Policy 2017;101:550–60. http://dx.doi.org/10.1016/j.enpol.2016.11.014.
[7] Caves Douglas, Eakin Kelly, Faruqui Ahmad. Mitigating price spikes in wholesale markets through market-based pricing in retail markets. Electr J 2000;13(3):13–23. http://dx.doi.org/10.1016/S1040-6190(00)00092-0.
[8] Frondel Manuel, Ritter Nolan, Schmidt Christoph M, Vance Colin. Economic impacts from the promotion of renewable energy technologies: The German experience. Energy Policy 2010;38(8):4048–56. http://dx.doi.org/10.1016/j.enpol.2010.03.029, URL https://www.sciencedirect.com/science/article/pii/S0301421510001928.
[9] Clò Stefano, Cataldi Alessandra, Zoppoli Pietro. The merit-order effect in the Italian power market: The impact of solar and wind generation on national wholesale electricity prices. Energy Policy 2015;77:79–88. http://dx.doi.org/10.1016/j.enpol.2014.11.038, URL https://www.sciencedirect.com/science/article/pii/S0301421514006661.
[10] Tveten Åsa Grytli, Bolkesjø Torjus Folsland, Martinsen Thomas, Hvarnes Håvard. Solar feed-in tariffs and the merit order effect: A study of the German electricity market. Energy Policy 2013;61:761–70. http://dx.doi.org/10.1016/j.enpol.2013.05.060, URL https://www.sciencedirect.com/science/article/pii/S0301421513003960.
[11] Bao Minglei, Ding Yi, Zhou Xiaoxin, Guo Chao, Shao Changzheng. Risk assessment and management of electricity markets: A review with suggestions. CSEE J Power Energy Syst 2021;7(6):1322–33. http://dx.doi.org/10.17775/CSEEJPES.2020.04250.
[12] Lagarto João, de Sousa Jorge, Martins Álvaro, Ferrão Paulo. Price forecasting in the day-ahead iberian electricity market using a conjectural variations ARIMA model. In: 9th international conference on the European energy market, Florence, Italy, 10-12 May. 2012, p. 1–7.
[13] Jakaša Tina, Andročec Ivan, Sprčić Petar. Electricity price forecasting — ARIMA model approach. In: 2011 8th international conference on the European energy market (EEM), Zagreb, Croatia, 25-27 May. 2011, p. 222–5.
[14] Karabiber Orhan Altuğ, Xydis George. Electricity price forecasting in the danish day-ahead market using the TBATS, ANN and ARIMA methods. Energies 2019;12(5):1–29.
[15] Conejo AJ, Plazas MA, Espinola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 2005;20(2):1035–42.
[16] Maciejowska Katarzyna, Nitka Weronika, Weron Tomasz. Enhancing load, wind and solar generation for day-ahead forecasting of electricity prices. Energy Econ 2021;99:105273. http://dx.doi.org/10.1016/j.eneco.2021.105273, URL https://www.sciencedirect.com/science/article/pii/S014098832100178X.
[17] Demir Sumeyra, Mincev Krystof, Kok Koen, Paterakis Nikolaos G. Data augmentation for time series regression: Applying transformations, autoencoders and adversarial networks to electricity price forecasting. Appl Energy 2021;304:117695. http://dx.doi.org/10.1016/j.apenergy.2021.117695, URL https://www.sciencedirect.com/science/article/pii/S0306261921010527.
[18] Garcia RC, Contreras J, van Akkeren M, Garcia JBC. A GARCH forecasting model to predict day-ahead electricity prices. IEEE Trans Power Syst 2005;20(2):867–74.
[19] Pindoriya NM, Singh SN, Singh SK. An adaptive wavelet neural network-based energy price forecasting in electricity markets. IEEE Trans Power Syst 2008;23(3):1423–32.
[20] Bento PMR, Pombo JAN, Calado MRA, Mariano SJPS. A bat optimized neural network and wavelet transform approach for short-term price forecasting. Appl Energy 2018;210:88–97. http://dx.doi.org/10.1016/j.apenergy.2017.10.058.
[21] Keles Dogan, Scelle Jonathan, Paraschiv Florentina, Fichtner Wolf. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. Appl Energy 2016;162:218–30. http://dx.doi.org/10.1016/j.apenergy.2015.09.087.
[22] Marcjasz Grzegorz, Narajewski Michał, Weron Rafał, Ziel Florian. Distributional neural networks for electricity price forecasting. Energy Econ 2023;125:106843. http://dx.doi.org/10.1016/j.eneco.2023.106843, URL https://www.sciencedirect.com/science/article/pii/S0140988323003419.
[23] Gunduz Salih, Ugurlu Umut, Oksuz Ilkay. Transfer learning for electricity price forecasting. Sustain Energy Grids Netw 2023;34:100996. http://dx.doi.org/10.1016/j.segan.2023.100996, URL https://www.sciencedirect.com/science/article/pii/S2352467723000048.
[24] Shrivastava Nitin Anand, Panigrahi Bijaya Ketan. A hybrid wavelet-ELM based short term price forecasting for electricity markets. Int J Electr Power Energy Syst 2014;55:41–50.
[25] Zhang Y, Li C, Li L. Wavelet transform and kernel-based extreme learning machine for electricity price forecasting. Energy Syst 2018;9:113–34.
[26] Zhang Yangrui, Tao Peng, Wu Xiangming, Yang Chenguang, Han Guang, Zhou Hui, Hu Yinlong. Hourly electricity price prediction for electricity market with high proportion of wind and solar power. Energies 2022;15(4). http://dx.doi.org/10.3390/en15041345.
[27] Anbazhagan S, Kumarappan N. Day-ahead deregulated electricity market price forecasting using recurrent neural network. IEEE Syst J 2013;7(4):866–72.
[28] Amjady N. Day-ahead price forecasting of electricity markets by a new fuzzy neural network. IEEE Trans Power Syst 2006;21(2):887–96.
[29] Amjady Nima, Daraeepour Ali. Design of input vector for day-ahead price forecasting of electricity markets. Expert Syst Appl 2009;36(10):12281–94. http://dx.doi.org/10.1016/j.eswa.2009.04.059.
[30] Xu Keqi, Sun Beibei, Wang Peng, Zhu Zhizhong, Tang Huidi. Electricity market price forecasting for a high renewable penetrated power system via random forest. In: 2022 IEEE/IAS 58th industrial and commercial power systems technical conference (I&CPS), Las Vegas, USA, 2-5 May. 2022, p. 1–7. http://dx.doi.org/10.1109/ICPS54075.2022.9773839.
[31] Luo Shuman, Weng Yang. A two-stage supervised learning approach for electricity price forecasting by leveraging different data sources. Appl Energy 2019;242:1497–512. http://dx.doi.org/10.1016/j.apenergy.2019.03.129.
[32] Meng Anbo, Wang Peng, Zhai Guangsong, Zeng Cong, Chen Shun, Yang Xiaoyi, Yin Hao. Electricity price forecasting with high penetration of renewable energy using attention-based LSTM network trained by crisscross optimization. Energy 2022;254:124212. http://dx.doi.org/10.1016/j.energy.2022.124212.
[33] Yang Zhang, Ce Li, Lian Li. Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods. Appl Energy 2017;190:291–305.
[34] Zhang Tingting, Tang Zhenpeng, Wu Junchuan, Du Xiaoxu, Chen Kaijie. Short term electricity price forecasting using a new hybrid model based on two-layer decomposition technique and ensemble learning. Electr Power Syst Res 2022;205:107762. http://dx.doi.org/10.1016/j.epsr.2021.107762.
[35] Yang Wendong, Sun Shaolong, Hao Yan, Wang Shouyang. A novel machine learning-based electricity price forecasting model based on optimal model selection strategy. Energy 2022;238:121989. http://dx.doi.org/10.1016/j.energy.2021.121989.
[36] Wang Deyun, Luo Hongyuan, Grunder Olivier, Lin Yanbing, Guo Haixiang. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. Appl Energy 2017;190:390–407. http://dx.doi.org/10.1016/j.apenergy.2016.12.134.
[37] Xiong Xiaoping, Qing Guohua. A hybrid day-ahead electricity price forecasting framework based on time series. Energy 2023;264:126099. http://dx.doi.org/10.1016/j.energy.2022.126099.
[38] Heidarpanah Mohammadreza, Hooshyaripor Farhad, Fazeli Meysam. Daily electricity price forecasting using artificial intelligence models in the Iranian electricity market. Energy 2023;263:126011. http://dx.doi.org/10.1016/j.energy.2022.126011.
[39] Wang Keke, Yu Min, Niu Dongxiao, Liang Yi, Peng Sha, Xu Xiaomin. Short-term electricity price forecasting based on similarity day screening, two-layer decomposition technique and Bi-LSTM neural network. Appl Soft Comput 2023;136:110018. http://dx.doi.org/10.1016/j.asoc.2023.110018, URL https://www.sciencedirect.com/science/article/pii/S1568494623000364.
[40] Plakas Konstantinos, Karampinis Ioannis, Alefragis Panayiotis, Birbas Alexios, Birbas Michael, Papalexopoulos Alex. A predictive fuzzy logic model for forecasting electricity day-ahead market prices for scheduling industrial applications. Energies 2023;16(10). http://dx.doi.org/10.3390/en16104085, URL https://www.mdpi.com/1996-1073/16/10/4085.
[41] Magalhães Bianca G, Bento Pedro MR, Pombo José AN, Calado Maria RA, Mariano Sílvio JPS. Spot price forecasting for best trading strategy decision support in the iberian electricity market. Expert Syst Appl 2023;224:120059. http://dx.doi.org/10.1016/j.eswa.2023.120059, URL https://www.sciencedirect.com/science/article/pii/S0957417423005614.
[42] Memarzadeh Gholamreza, Keynia Farshid. Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm. Electr Power Syst Res 2021;192:106995. http://dx.doi.org/10.1016/j.epsr.2020.106995.
[43] Jiang Ping, Nie Ying, Wang Jianzhou, Huang Xiaojia. Multivariable short-term electricity price forecasting using artificial intelligence and multi-input multi-output scheme. Energy Econ 2023;117:106471. http://dx.doi.org/10.1016/j.eneco.2022.106471.
[44] Shao Zhen, Zheng Qingru, Liu Chen, Gao Shuangyan, Wang Gang, Chu Yan. 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. http://dx.doi.org/10.1016/j.epsr.2021.107453.

[45] Pourdaryaei Alireza, Mohammadi Mohammad, Karimi Mazaher, Mokhlis Hazlie, Illias Hazlee A, Kaboli Seyed Hamidreza Aghay, Ahmad Shameem. Recent development in electricity price forecasting based on computational intelligence techniques in deregulated power market. Energies 2021;14(19). http://dx.doi.org/10.3390/en14196104.

[46] Pourdaryaei Alireza, Mokhlis Hazlie, Illias Hazlee Azil, Kaboli SHr Aghay, Ahmad Shameem. Short-term electricity price forecasting via hybrid backtracking search algorithm and ANFIS approach. IEEE Access 2019;7:77674–91. http://dx.doi.org/10.1109/ACCESS.2019.2922420.

[47] Agrawal Rahul Kumar, Muchahary Frankle, Tripathi Madan Mohan. Ensemble of relevance vector machines and boosted trees for electricity price forecasting. Appl Energy 2019;250:540–8. http://dx.doi.org/10.1016/j.apenergy.2019.05.062.

[48] Zhang Xiaobo, Wang Jianzhou, Gao Yuyang. A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM. Energy Econ 2019;81:899–913. http://dx.doi.org/10.1016/j.eneco.2019.05.026.

[49] Sandhu Harmanjot Singh, Fang Liping, Guan Ling. Forecasting day-ahead price spikes for the ontario electricity market. Electr Power Syst Res 2016;141:450–9.

[50] Fragkioudaki Anna, Marinakis Adamantios, Cherkaoui Rachid. Forecasting price spikes in European day-ahead electricity markets using decision trees. In: 2015 12th international conference on the European energy market (EEM), Lisbon, Portugal, 19-22 May. p. 1–5.

[51] Voronin Sergey, Partanen Jarmo. Price forecasting in the day-ahead energy market by an iterative method with separate normal price and price spike frameworks. Energies 2013;6(11):5897–920.

[52] Shi Wei, Wang Yufeng, Chen Yiyuan, Ma Jianhua. An effective two-stage electricity price forecasting scheme. Electr Power Syst Res 2021;199:107416. http://dx.doi.org/10.1016/j.epsr.2021.107416.

[53] Shrivastava Nitin Anand, Panigraphi Bijaya Ketan, Lim Meng-Hiot. Electricity price classification using extreme learning machines. Neural Comput Appl 2016;27:9–18. http://dx.doi.org/10.1007/s00521-013-1537-1.

[54] Zhao Jun Hua, Dong Zhao Yang, Li Xue, Wong Kit Po. A framework for electricity price spike analysis with advanced data mining methods. IEEE Trans Power Syst 2007;22(1):376–85. http://dx.doi.org/10.1109/TPWRS.2006.889139.

[55] Vu Dao H, Muttaqi Kashem M, Agalgaonkar Ashish P, Bouzerdoum Abdesselam. Short-term forecasting of electricity spot prices containing random spikes using a time-varying autoregressive model combined with kernel regression. IEEE Trans Ind Inf 2019;15(9):5378–88. http://dx.doi.org/10.1109/TII.2019.2911700.

[56] Amjady Nima, Keynia Farshid. Electricity market price spike analysis by a hybrid data model and feature selection technique. Electr Power Syst Res 2010;80(3):318–27. http://dx.doi.org/10.1016/j.epsr.2009.09.015.

[57] Tan Yong Qiang, Shen Yan Xia, Yu Xin Yan, Lu Xin. Day-ahead electricity price forecasting employing a novel hybrid frame of deep learning methods: A case study in NSW, Australia. Electr Power Syst Res 2023;220:109300. http://dx.doi.org/10.1016/j.epsr.2023.109300, URL https://www.sciencedirect.com/science/article/pii/S037877962300189X.

[58] Box George EP, Jenkins Gwilym M, Reinsel Gregory C, Ljung Greta M. Time series analysis, forecasting and control. 5th ed.. USA: Wiley; 2015.

[59] Weron Rafael. Electricity price forecasting: A review of the state-of-the-art with a look into the future. Int J Forecast 2014;30:1030–811.

[60] Lehna Malte, Scheller Fabian, Herwartz Helmut. Forecasting day-ahead electricity prices: A comparison of time series and neural network models taking external regressors into account. Energy Econ 2022;106:105742. http://dx.doi.org/10.1016/j.eneco.2021.105742, URL https://www.sciencedirect.com/science/article/pii/S0140988321005879.

[61] Vega-Márquez Belén, Rubio-Escudero Cristina, Nepomuceno-Chamorro Isabel A, Arcos-Vargas Ángel. Use of deep learning architectures for day-ahead electricity price forecasting over different time periods in the spanish electricity market. Appl Sci 2021;11(13). http://dx.doi.org/10.3390/app11136097.

[62] Fraunholz Christoph, Kraft Emil, Keles Dogan, Fichtner Wolf. Advanced price forecasting in agent-based electricity market simulation. Appl Energy 2021;290:116688. http://dx.doi.org/10.1016/j.apenergy.2021.116688.

[63] Tschora Léonard, Pierre Erwan, Plantevit Marc, Robardet Céline. Electricity price forecasting on the day-ahead market using machine learning. Appl Energy 2022;313:118752. http://dx.doi.org/10.1016/j.apenergy.2022.118752.

[64] Huang Guang-Bin, Zhu Qin-Yu, Siew Chee-Kheong. Extreme learning machine: Theory and applications. Neurocomputing 2006;70(1):489–501.

[65] Zhao Jun Hua, Dong Zhao Yang, Xu Zhao, Wong Kit Po. A statistical approach for interval forecasting of the electricity price. IEEE Trans Power Syst 2008;23(2):267–76.

[66] Wan Can, Xu Zhao, Wang Yelei, Dong Zhao Yang, Wong Kit Po. A hybrid approach for probabilistic forecasting of electricity price. IEEE Trans Smart Grid 2014;5(1):463–70.

[67] Chen Xia, Dong Zhao Yang, Meng Ke, Xu Yan, Wong Kit Po, Ngan HW. Electricity price forecasting with extreme learning machine and bootstrapping. IEEE Trans Power Syst 2012;27(4):2055–62.

[68] Rafiei Mehdi, Niknam Taher, Khooban Mohammad-Hassan. Probabilistic forecasting of hourly electricity price by generalization of ELM for usage in improved wavelet neural network. IEEE Trans Ind Inf 2017;13(1):71–9.

[69] Zhang Chenxu, Fu Yong. Probabilistic electricity price forecast with optimal prediction interval. IEEE Trans Power Syst 2023;1–10. http://dx.doi.org/10.1109/TPWRS.2023.3235193.

[70] Brusaferri Alessandro, Matteucci Matteo, Portolani Pietro, Vitali Andrea. Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices. Appl Energy 2019;250:1158–75. http://dx.doi.org/10.1016/j.apenergy.2019.05.068.

[71] Yang Wendong, Wang Jianzhou, Niu Tong, Du Pei. A novel system for multi-step electricity price forecasting for electricity market management. Appl Soft Comput 2020;88:106029. http://dx.doi.org/10.1016/j.asoc.2019.106029.

[72] Grothe Oliver, Kächele Fabian, Krüger Fabian. From point forecasts to multivariate probabilistic forecasts: The schaake shuffle for day-ahead electricity price forecasting. Energy Econ 2023;120:106602. http://dx.doi.org/10.1016/j.eneco.2023.106602, URL https://www.sciencedirect.com/science/article/pii/S0140988323001007.

[73] Lu Xin, Dong Zhao Yang, Li Xue. Electricity market price spike forecast with data mining techniques. Electr Power Syst Res 2005;73(1):19–29.

[74] Christensen TM, Hurn AS, Lindsay KA. Forecasting spikes in electricity prices. Int J Forecast 2012;28(2):400–11.

[75] Shah Ismail, Akbar Sher, Saba Tanzila, Ali Sajid, Rehman Amjad. Short-term forecasting for the electricity spot prices with extreme values treatment. IEEE Access 2021;9:105451–62. http://dx.doi.org/10.1109/ACCESS.2021.3100076.

[76] Liu Luyao, Bai Feifei, Su Chenyu, Ma Cuiping, Yan Ruifeng, Li Hailong, Sun Qie, Wennersten Ronald. Forecasting the occurrence of extreme electricity prices using a multivariate logistic regression model. Energy 2022;247:123417. http://dx.doi.org/10.1016/j.energy.2022.123417.

[77] Hagfors Lars Ivar, Kamperud Hilde Hørthe, Paraschiv Florentina, Prokopczuk Marcel, Sator Alma, Westgaard Sjur. Prediction of extreme price occurrences in the German day-ahead electricity market. Quant Finance 2016;16(12):1929–48.

[78] Bello Antonio, Reneses Javier, Muñoz Antonio. Medium-term probabilistic forecasting of extremely low prices in electricity markets: Application to the Spanish case. Energies 2016;9(3). http://dx.doi.org/10.3390/en9030193.

[79] SMARD. German market data. Available, https://www.smard.de/en/downloadcenter/download-market-data.

[80] Elering. Estonian TSO. Available, https://dashboard.elering.ee/en/nps/price?interval=minute&period=days&start=2023-11-06T22:00:00.000Z&end=2023-11-07T21:59:59.999Z&show=table.

[81] Fingrid. Finnish market data. Available, https://data.fingrid.fi/open-data-forms/search/en/.

[82] Conroy Ronán M. The RCSI sample size handbook. 2018, http://dx.doi.org/10.13140/RG.2.2.30497.51043.

[83] Gullì Francesco, Balbo Antonio Lo. The impact of intermittently renewable energy on Italian wholesale electricity prices: Additional benefits or additional costs? Energy Policy 2015;83:123–37. http://dx.doi.org/10.1016/j.enpol.2015.04.001, URL https://www.sciencedirect.com/science/article/pii/S0301421515001433.

[84] De Vos Kristof. Negative wholesale electricity prices in the german, french and belgian day-ahead, intra-day and real-time markets. Electr J 2015;28(4):36–50. http://dx.doi.org/10.1016/j.tej.2015.04.001.

[85] Morales Juan M, Conejo Antonio J, Madsen Henrik, Pinson Pierre, Zugno Marco. Integrating renewables in electricity markets. Springer; 2014.

[86] Martin de Lagarde Cyril, Lantz Frédéric. How renewable production depresses electricity prices: Evidence from the german market. Energy Policy 2018;117:263–77. http://dx.doi.org/10.1016/j.enpol.2018.02.048, URL https://www.sciencedirect.com/science/article/pii/S0301421518301253.

[87] Borenstein Severin. Understanding competitive pricing and market power in wholesale electricity markets. Electr J 2000;13(6):49–57. http://dx.doi.org/10.1016/S1040-6190(00)00124-X, URL https://www.sciencedirect.com/science/article/pii/S104061900000124X.

[88] Schneider Stefan. Power spot price models with negative prices. Munich Pers Repec Arch 2010;(29958):1–23, URL https://mpra.ub.uni-muenchen.de/29958/.

[89] Tan Zhongfu, Zhang Jinliang, Wang Jianhui, Xu Jun. Day-ahead electricity price forecasting using wavelet transform combined with ARIMA and GARCH models. Appl Energy 2010;87(11):3606–10. http://dx.doi.org/10.1016/j.apenergy.2010.05.012, URL https://www.sciencedirect.com/science/article/pii/S0306261910001807.

[90] Billé Anna Gloria, Gianfreda Angelica, Del Grosso Filippo, Ravazzolo Francesco. Forecasting electricity prices with expert, linear, and nonlinear models. Int J Forecast 2023;39(2):570–86. http://dx.doi.org/10.1016/j.ijforecast.2022.01.003, URL https://www.sciencedirect.com/science/article/pii/S0169207022000036.

[91] Huurman Christian, Ravazzolo Francesco, Zhou Chen. The power of weather. Comput Statist Data Anal 2012;56(11):3793–807. http://dx.doi.org/10.1016/j.csda.2010.06.021, URL https://www.sciencedirect.com/science/article/pii/S0167947310002665. 1st issue of the Annals of Computational and Financial Econometrics Sixth Special Issue on Computational Econometrics.

[92] Chaâbane Najeh. A hybrid ARFIMA and neural network model for electricity price prediction. Int J Electr Power Energy Syst 2014;55:187–94. http://dx.doi.org/10.1016/j.ijepes.2013.09.004, URL https://www.sciencedirect.com/science/article/pii/S0142061513003773.

[93] Diongue Abdou Kâ, Guégan Dominique, Vignal Bertrand. Forecasting electricity spot market prices with a k-factor GIGARCH process. Appl Energy 2009;86(4):505–10. http://dx.doi.org/10.1016/j.apenergy.2008.07.005, URL https://www.sciencedirect.com/science/article/pii/S0306261908001748.

[94] Liu Heping, Shi Jing. Applying ARMA–GARCH approaches to forecasting short-term electricity prices. Energy Econ 2013;37:152–66. http://dx.doi.org/10.1016/j.eneco.2013.02.006, URL https://www.sciencedirect.com/science/article/pii/S0140988313000303.

[95] Boubaker Heni, Canarella Giorgio, Gupta Rangan, Miller Stephen M. A hybrid ARFIMA wavelet artificial neural network model for DJIA index forecasting. Comput Econ 2023;62:1801–43. http://dx.doi.org/10.1007/s10614-022-10320-z.

[96] Kang Sang Hoon, Yoon Seong-Min. Modeling and forecasting the volatility of petroleum futures prices. Energy Econ 2013;36:354–62. http://dx.doi.org/10.1016/j.eneco.2012.09.010, URL https://www.sciencedirect.com/science/article/pii/S0140988312002241.

[97] Loizidis Stylianos, Konstantinidis Georgios, Theocharides Spyros, Kyprianou Andreas, Georghiou George E. Electricity day-ahead market conditions and their effect on the different supervised algorithms for market price forecasting. Energies 2023;16(12). http://dx.doi.org/10.3390/en16124617, URL https://www.mdpi.com/1996-1073/16/12/4617.