Forecasting price spikes in day-ahead electricity markets: techniques, challenges, and the road ahead

BY Samaneh Sheybanivaziri, Jérôme Le Dréau and Hussain Kazmi

DISCUSSION PAPER

NHH

Institutt for foretaksøkonomi
Department of Business and Management Science

FOR 1/2024

ISSN: 2387-3000

January 2024

Forecasting price spikes in day-ahead electricity markets: techniques, challenges, and the road ahead

Samaneh Sheybanivaziri* Jérôme Le Dréau † Hussain Kazmi*‡

Abstract

Due to the increase in renewable energy production and global socioeconomic turmoil, the volatility in electricity prices has considerably increased in recent years, leading to extreme positive and negative price spikes in many electricity markets. Forecasting (the risk of) these prices accurately in advance can enable risk-informed decision-making by both consumers and generators, as well as by the grid operators. In this work, focusing on day-ahead markets, we review recent developments in how price spikes are defined, as well as which explanatory factors and methodologies have been used to forecast them. The paper identifies seven categories of influencing factors, which come with over 30 sub-classifications that can cause price spikes. In terms of methodologies, probabilistic models are being increasingly utilized to capture uncertainty in the price forecast. The review uncovers a wide range in all of these choices as well as others, which makes it difficult to compare methods and select best practices for predicting price spikes.

Keywords— Spikes, Electricity markets, Day-ahead market, Point forecast, Probabilistic forecasts

1 Introduction

Electricity markets provide vital information to generators and consumers in liberalized settings, both on short- and long-term horizons. In the short-term, producers and consumers modulate their electricity demand or generation in response to price fluctuations [120]. These markets also provide the signals needed by grid operators to ensure the grid's resilience to unforeseen disturbances [61]. In the long-term, electricity prices help guide investment decisions on the construction of new power plants, transmission and distribution infrastructure, storage facilities, etc. [10].

With a greater proliferation of renewable energy sources (RES) and demand electrification (heat, transport, industry), electricity prices have become increasingly more volatile [66]. This has led to a greater coupling between electricity prices and the weather - which formerly influenced prices primarily via demand fluctuations [85, 60]. These price variations have been further exacerbated by unprecedented supply-chain disruptions caused by the COVID-19 pandemic as well as geopolitical tensions, including the war in Ukraine. The increased exposure of consumers and producers to variations in electricity prices has led to an enormous socioeconomic

*NHH, Dept. of Business and Management Science, samaneh.sheybanivaziri@nhh.no, Norway

†La Rochelle University, LaSIE UMR CNRS 7356, jledreau@univ-lr.fr, France

‡Dept. of Electrical Engineering, KU Leuven, hussain.kazmi@kuleuven.be, Belgium

burden in recent years, as well as a resurgence in energy poverty. This is true both for end users, including households [4] which have to pay considerably more to purchase energy, as well as electricity generators, including renewable energy producers, which have had to shelve projects because of increasing price volatility. This volatility has underscored the need for a deeper understanding of the various electricity markets that play a pivotal role in shaping the cost and availability of electricity.

In liberalized markets, several products (or sub-markets) exist, differing in terms of time or duration between the agreement made in the contract and the actual delivery in real-time as illuminated by [48]:

1. The long-term futures market is used to trade energy far in advance (years) and for long periods;
2. The day-ahead (DA) market, the focus of this paper, is primarily a financial market for the exchange of forward contracts for energy, with a gate closure time approximately one day in advance (i.e. bids for energy demand or generation cannot be updated after this time);
3. Intraday markets are forward markets for the exchange of electricity, which open after the gate closing time of the DA market, and can be used by generators or consumers to rebalance their portfolios;
4. The imbalance market is used to penalize market actors (e.g. producers) for deviating from their stated bids (nominations) on the DA or intraday markets. The objective is to minimize the imbalance as much as possible;
5. The imbalances that occur in the system are met by different frequency reserves ranging from very slow tertiary to almost instantaneous primary reserves in terms of response times. Even though there is considerable diversity in the naming of these reserves across different countries, the core idea remains the same;
6. In addition to these markets, several other markets, such as peer-to-peer markets to enable local energy exchanges and communities, have been explored recently as well [38].

Together these markets cover most of the electricity traded in the system. For instance, in the United States, approximately 95% of energy transactions are attributed to day-ahead markets [33] and in Nordpool (one of the largest power exchanges in the EU), more than 90% of all energy transactions occur on the day-ahead market, with the remaining being handled by the real-time markets[1].

In this paper, focusing on the DA market, we consider how accurate price forecasts can guide generation or demand nominations, and the risks that extreme price events pose to this exercise. This is motivated by the fact that while price forecasting algorithms often achieve good performance on average, they exhibit certain characteristics that warrant attention. Most notably they frequently (1) only produce point forecasts, and (2) lack reproducibility and report results only on a brief test period [28]. Furthermore, most studies do not consider the forecast models' performance during periods of extreme prices, which can expose market parties to elevated risk. The objective of this work is therefore three-fold: (1) it reviews the different definitions of electricity price spikes posed by researchers; (2) it identifies the most important factors and algorithms in forecasting electricity price spikes accurately; and (3) it discusses advances and core challenges in forecasting spikes in electricity market prices. In the remainder of this section, we describe key motivating factors for forecasting price spikes accurately, before discussing existing relevant work and the methodology for carrying out the literature review.

[1] https://www.nordpoolgroup.com/en/message-center-container/newsroom/exchange-message-list/2023/q1/nord-pool-announces-2022-trading-figures/: :text=Nord%20Pool%20is%20a%20Nominated,of%201077.35%20TWh%20traded%20power

1.1 Why is accurate spike prediction important?

Electricity generators as well as large industrial consumers and aggregators of energy demand use price forecasts to optimize their bidding schedules and asset allocation. For instance, many retailers buy a large portion of their predictable base-load demand in advance (on the futures market). Subsequently, closer to the time of delivery, they procure the remainder of the actual demand on the DA and intraday market. The fraction of electricity demand procured on the DA market depends heavily on the (perceived) confidence in the price forecast. This is especially relevant when these market players have to also act as balance responsible parties (BRPs), which is increasingly the case. Combined with possible flexibility from energy storage, this allows for new avenues of optimal operation of grid-connected assets on several aggregation scales [111, 63].

Consequently, failure to predict spikes accurately is harmful to both the generators as well as retailers (or large industrial users). Very low or negative electricity prices, for instance, pose a grave risk to generators (i.e. they might be forced to pay consumers for electricity off-take). Similarly, according to Eichler et al., [88], the occurrence of price surges might present a significant challenge for retailers as they typically procure electricity at a fluctuating cost while vending it to customers at a constant rate (e.g. in the Australian electricity market). This has been considered elsewhere too. For instance, researchers have shown that a retail company, which is trying to acquire power at the lowest possible price, can potentially lose a yearly dividend due to extreme fluctuations [106]. Furthermore, researchers have also highlighted the significance of considering the occurrence of negative prices and spikes when evaluating the profitability of purchasing or selling energy contracts during periods of low demand [89]. Thus, erroneous spike forecasts undermine risk management [93].

Unlike large consumers (and parties that act as BRPs), small consumers, such as households, are typically shielded from day-ahead markets by fixed-price contracts. Only a few European countries are proposing dynamic pricing that follows the spot markets, including Estonia, Sweden, Spain, Finland, and Denmark [23]. These dynamic pricing schemes have been promoted by the EU Directive 2019/944 on the rules of the internal market for electricity, in order to encourage end-user engagement. As highlighted by Numminen et al. [17], the limited adaptation rate of such contracts can be explained by the perceived risk and limited benefits. Large price fluctuations, as evidenced in Fig. 1, may be perceived as a disadvantage, despite the possibility of contracting a price cap option to avoid price spikes. With greater proliferation of distributed generation and storage, these types of programs are likely to gain in popularity. As before, mis-forecasts can expose end users to adverse outcomes.

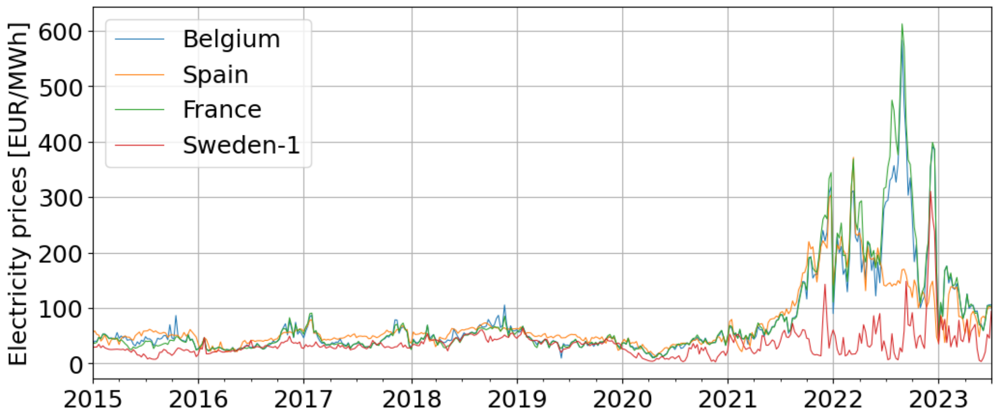

Figure 1: Weekly-averaged DA electricity prices in four different bidding zones (BE, ES, FR, SE-1) in the European Union (EU)

1.2 Related work and reviews

In recent years, research on energy forecasting has developed substantially. This has led to a sizable literature on probabilistic wind and solar power forecasting [87, 22, 19, 7]. Likewise, Hong and Fan [73] examine probabilistic electric load forecasting based on four alternative forecast horizons in their instructional review. Another recent review on energy forecasting as a whole field brings these topics together [36].

In terms of electricity markets, Aggarwal et al. [99] found, upon reviewing and classifying electricity price forecasting publications across 11 different power markets, that spike prediction was still a developing field. Subsequently, Weron et al. [86] reviewed the methodologies used to forecast electricity prices between 1998 and 2013, covering the main model types (multi-agent models, fundamental models, reduced-form models, statistical models, computational intelligence models), with a special emphasis on their forecasting capabilities. Regarding spikes, reduced-form models (such as jump diffusion and Markov-regime Switching) were discussed and reported to perform well. Additionally, probabilistic forecasts were considered as well, but this was not the main focus of the paper. A follow-up review, specifically on probabilistic electricity price forecasting, can be seen in [59], in which the lack of research on probabilistic electricity price forecasting compared to wind power is addressed. Using existing literature, the authors illustrate four approaches to make probabilistic forecasts, but this review does not focus on price spikes and does not cover the recent volatility in electricity markets. Finally, a more recent review by Lago et al. [28] covers best practices and algorithms in forecasting day-ahead electricity prices, but the focus of the article is on point forecasting and does not explicitly discuss spikes as well as not considering the recent high volatility.

Despite these reviews, to the best of our knowledge, no existing review has actually focused on forecasting price spikes. With the recent volatility in DA prices, this paper is meant to address this gap in the published literature.

1.3 Literature selection methodology

To provide a snapshot of the research studies focusing on price spikes in energy markets, we utilized the Scopus [12] database. Query 1 was focused on spikes in energy markets[2], resulting in a total of 450 documents. From

[2] We used the following Scopus query: ((TITLE-ABS-KEY(electricity) AND TITLE-ABS-KEY(spike) AND TITLE-ABS-KEY("energy market") OR TITLE-ABS-KEY("electricity market") OR TITLE-ABS-KEY("market")) AND PUBYEAR > 1998 AND PUBYEAR < 2024)

1999 until 2014, the number of publications has increased with a significant spike observed in 2014 (37), which represents a peak in research activity. There's a consistent high level of research activity from 2012 to 2018, with the number of publications staying above 20 per year. There are some fluctuations in the later years (2019-2023), but the overall trend remains relatively stable.

Narrowing down our query to probabilistic electricity price spikes[3] resulted in 64 publications, while supporting an upward trend in recent years, shown by query 2 in Fig. 2. The plot highlights a clear gap in this field, indicating a need for further research. With increasing electricity prices in global markets since 2021, addressing this gap becomes crucial for future analysis. We ran the queries initially in January 2023, and then again in August 2023 to ensure our research remained up-to-date and comprehensive. Fig. 2 summarizes the overall trend throughout the years, revealing increasing attention being paid to the topic of spikes in the electricity and energy markets.

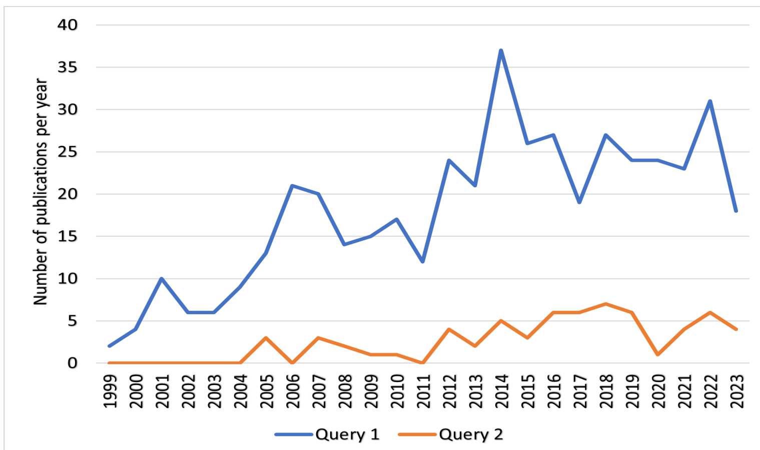

Figure 2: Scopus-indexed publications from 1999 to 2023 studying price spikes in energy markets

Figure 3 offers a condensed overview of our review study spanning from 2006 to 2023. It visually summarizes the breadth of our research, encompassing geographical regions, timeframes, forecast methodologies, models employed, and the number of external factors, providing a holistic view of our study's coverage.

[3]((TITLE-ABS-KEY(spikes) OR TITLE-ABS-KEY("extreme price") OR TITLE-ABS-KEY("extreme observations") OR TITLE-ABS-KEY("price jumps") AND TITLE-ABS-KEY("price forecasting") OR TITLE-ABS-KEY("probabilistic forecasting") AND TITLE-ABS-KEY("energy market") OR TITLE-ABS-KEY("electricity market")) AND PUBYEAR > 2004 AND PUBYEAR < 2024)

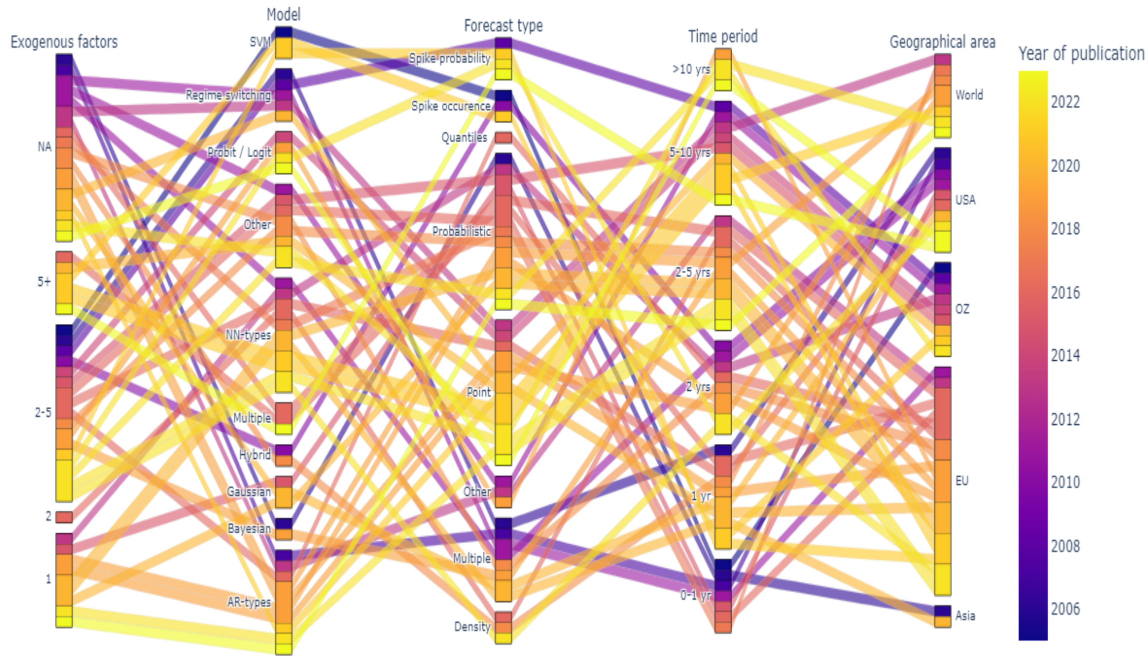

Figure 3: An overview of different dimensions of consideration for the papers reviewed in this study.

1.4 Structure of this review

Using this selection methodology, we will review the resulting articles on price spikes and high volatility in the day-ahead electricity market. The following is an overview of the remainder of the paper. Section 2 provides an overview of the different methods used to define electricity price spikes. Section 3 presents an overview of the factors that can lead to extreme electricity prices. Section 4 covers forecast models that have been used in the literature to forecast electricity price spikes. Section 5 describes how price spike forecasts are evaluated in practice, while Sections 6 and 7 conclude the review with a discussion on open challenges and an outlook on the future.

2 How are spikes defined?

Before discussing price spike forecast methods, we consider different techniques that researchers have employed to detect price spikes. This detection often forms the first step in the subsequent prediction task. Towards this end, researchers have proposed many different algorithms to estimate price ranges (either static or dynamic) beyond which market behavior is considered abnormal. We distinguish these techniques into those that work directly with the price time series and those that preprocess it using some transformation.

2.1 Price spike identification using thresholds and data distributions

2.1.1 Fixed thresholding-based methods

Thresholding-based methods rely solely on the historical price time series data to decide whether a specific price at a given time step can be considered a spike or not and are arguably the most common strategy in the published literature on electricity price spike forecasting [3, 11, 21, 32, 31, 30, 45, 52]. These methods typically disregard the context in which the market is operating under and often focus on positive price spikes rather than negative ones. This is due, at least in part, to many markets not experiencing negative electricity prices until very recently or at all.

In many such proposed methods, defining the threshold therefore takes central importance. As a case in point, [93] classifies prices beyond a certain level as "extreme pricing occurrences", but the thresholds, in general, remain open to interpretation depending on the market and the researcher. For instance, noting conditions for the Australian electricity market, Christensen et al. classify mild spikes as ranging from \$100/MWh to \$300/MWh, and severe spikes as price events ranging from \$300/MWh to \$10,000/MWh [93]. Eichler et al.[88] also follow this approach outlined by categorizing prices above \$ 100 and \$ 300/MWh as spikes and extreme spikes respectively. Likewise, for European markets, researchers have often used thresholds of 100 EUR/MWh or 200 EUR/MWh to identify positive price spikes, depending on the bidding zone.

These fixed thresholds assume that prices remain stationary over time - an assumption that can lead to issues with too many (or too few) spikes being identified at different periods in time. Fig. 4 shows the price in Belgium for five years between 2018 and 2022 (inclusive). It is obvious that the time series under consideration is non-stationary, with several price spikes over the years. Applying a price spike threshold of 200 EUR/MWh leads to very different price spikes over the years i.e. almost none in 2019 and 2020, over 10% in 2021, and almost 60% of the year in 2022. Consequently, according to this thresholding scheme, the electricity price spike lasted more than half of the entire year. Notwithstanding the extreme conditions the European electricity market went through in 2022, this still does not line up with the notion that a spike is meant to signify extraordinary conditions.

2.1.2 Adaptive thresholding-based methods

While the fixed threshold-based definition of electricity price spikes is widely used due to its simplicity, it is fundamentally limited by its non-adaptive nature. To address this issue, researchers have turned to dynamic or adaptive thresholds. For instance, Muninan and Weron identified extremely low and high electricity prices, by analyzing histograms and density functions of observed prices [43]. These visualizations showed that there were infrequent but significant deviations from typical price values, creating tails in the distribution of price changes. This idea has also been explored by Janczura et al. in setting the thresholds at the 2.5% and 10% percentiles of observed prices [90], and Jin et al. [13] considering the top 1% highest prices and the lowest 2.5% prices to identify spikes in the German electricity market. Different time windows ranging from one month to one year are typically used to estimate the percentiles or distributions, and the results have been demonstrated in several markets ranging from France and Italy to Hungary. Likewise, using the Australian NEM market as a motivating example, Lu et al. [106] show that the entire probability distribution of observed electricity prices can be used to define the threshold for abnormal market operation.

By establishing a linear correlation between predictors and log-odds of spikes, Bajai's study [11] on the Hungarian DA Market, employs a logit model to pinpoint price spikes beyond the hour-specific 30-day moving average of $-/+30$ euros/MWh. This method adeptly accommodates intraday fluctuations by calculating moving averages for each of the 24 hours, ensuring spike identification in relation to the specific hourly moving average. This methodology has been further extended by Zhang et al. to consider thresholds that depend on the time of day in several markets (PJM, MISO, NYISO, ISO-NE, Spain), i.e. the method leads to a vector of 24 price thresholds for any given day, and then adapts it using the fixed calibration window [21]. This accounts

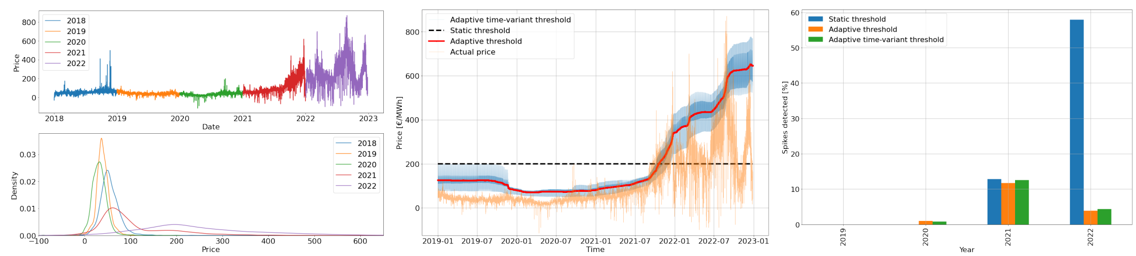

Figure 4: Spike detection methods applied to electricity day-ahead price data from Belgium: (Left top) Electricity prices in Belgium between 2018 and 2022; (Left bottom) Kernel density estimate of the electricity price distribution; (Middle) Electricity prices along with thresholds calculated using three different methods; (Right) Histogram showing frequency of spikes detected between 2019 and 2022 using three different methods

for the seasonal (daily, weekly, and yearly) trends in electricity prices. Another example of such a threshold includes defining prices that are 3 standard deviations away from the historic data mean as a spike [52, 31, 11, 53]. As an example, Sandhu et al. have identified price spikes in the Ontario electricity market by comparing prices to a threshold level computed from the mean and standard deviation of historical prices [78].

Fig. 4 shows the results of applying three different methods to the Belgian electricity prices between 2019 and 2022 (static threshold of 200 euros/MWh, adaptive three sigma threshold over the past twelve months, adaptive three sigma threshold conditioned on individual hours of the day). As expected, the amount of price spikes identified until 2021 remains roughly the same with all three methods, but changes dramatically in 2022 when the static threshold flags an unrealistically high percentage of spikes.

2.1.3 The specific case of negative prices

Various methodologies have been also proposed by researchers to set fixed thresholds for extremely low (or even negative) prices in electricity markets. Lu et al. use a threshold of 0 EUR/MWh, which is arguably not a very well-motivated choice, given that prices can fall significantly below this level and, as such, suffers from all the same limitations of fixed threshold-based methods highlighted above [106]. Some researchers have also employed the same three sigma standard deviations from the mean value as the dynamic threshold as described above.

2.2 Transformations and pre-processing

Alternatively, several researchers have proposed to transform the price time series in some way before applying the thresholding function. A simple example of this method is calculating the difference in price between two adjacent time periods and comparing it to the maximum difference within the normal price range. If the difference exceeds the threshold, the current time price is classified as an abnormal price jump. For instance, [3] propose using a stochastic process to capture the fluctuations in prices over time. The first step of this strategy is selecting a suitable stochastic model that can accurately represent the price data. After that, statistical filtering techniques, such as Sequential Monte Carlo algorithms, can be used to differentiate between sudden price spikes and normal price movements. Curiously enough, the authors chose to use a fixed threshold for detecting negative spikes, which allows them to bypass the technical difficulties involved in the statistical filtering technique. Another related approach, as exemplified in Bello's study [69], involves utilizing a cut-off value to differentiate prices between normal and unusually low, depending on the probability predicted by a model.

A different approach has been proposed by Fanone et al. who utilized hourly EPEX DA price data from 2007 to 2010 to identify price spikes after removing the annual mean [89]. They suggested identifying and eliminating extreme prices before estimating other components as the presence of price spikes in the demeaned time series can affect the analysis of seasonality and mean reversion. To do so, they use the "Peak Over Threshold (POT)" method to detect price spikes and choose a threshold value based on the mean excess function. The identified threshold values were then used to filter out the distributional properties of the spike processes. Consequently, a price of -37.99 EUR/MWh was established as the threshold for the lower tail, while prices exceeding 54.21 EUR/MWh were classified as notably positive. This definition would obviously need to be tweaked over time with the latest price data. This method can also be seen as a special case of dynamic thresholding.

Several extensions and generalizations of this method exist as well, for instance using time series decompositions such as STL or time-frequency methods (including the wavelet transform). The extracted seasonal and trend components can then be used to deseasonalize and detrend the time series before price spike detection thresholds or techniques are applied. For instance, using data from New York, Karakoyun et al. deseasonalize prices before labeling the highest five percent and the lowest five percent of the time series as spikes [6]. Markov-Regime-Switching models are also sometimes used on deseasonalized and/or detrended data and threshold probabilities are defined to identify regime switching [39].

3 Factors affecting day-ahead market prices

In this section, we begin with an overview of factors that affect day-ahead market prices, before turning our attention to the different causes of price spikes in practice - both in the positive and negative direction.

3.1 Factors affecting electricity prices

Several factors influence electricity day-ahead market prices. These include electricity supply and demand, prevailing weather conditions, interconnection capacity and availability, and fuel and carbon prices among many others [64, 93]. The supply mix determines the merit order dispatch, which defines the order in which generating units are activated. Renewable energy sources, such as wind and solar, have very low operational production costs, and therefore play an important role in defining electricity market prices, specially when they are widely adopted and integrated into the electricity grid to a significant extent [64]. Likewise, fuel prices for a given generation mix (e.g. coal or gas) also directly influence electricity prices. The demand side, on the other hand, dictates how much generation will be required. During peak demand times, this can therefore lead to the activation of peaking plants, such as gas-fired plants, often leading to substantial increases in electricity prices [57].

Weather is one factor that has historically been closely tied to peak demand, but it also increasingly affects the supply side in renewable-dominated systems. In fact, in such systems, weather-related phenomena such as Dunkelflaute (i.e. the prolonged absence of both solar and wind energy), will have a profound effect on market operation as well as security of supply [18]. Likewise, exogenous factors including calendar events, such as holidays and sporting events, also have the potential to influence market prices, mediated by higher or lower demand [86].

3.2 Factors triggering upward price spikes

In this section, we begin with describing early experiences of price spikes in electricity markets, followed by more recent findings pertaining to both positive and negative price spikes.

3.2.1 Early experiences in North America

California presented an early example of an electricity market failure around the turn of the century. Woo [118] pinpointed price surges as a significant factor contributing to this malfunction. These spikes were not isolated issues; they occurred in conjunction with recurring capacity deficiencies, power outages, and increased price fluctuations. The root causes revolved around deficient market structuring, manipulation of the market by influential players, financial precariousness, imbalances between supply and demand, and mounting marginal costs [118]. These findings were echoed in [114], while investigating price fluctuations in the UK, Alberta, and Norway.

In this context, [116] contended that energy providers' and customers' bidding behavior play a role in price spikes. They specifically proposed "opportunistic tacit collusion" as an explanation for how bidding patterns may lead to spikes in prices. A subsequent study by [107] extends this notion and authors hypothesize that collusion occurs when firms coordinate their behavior tacitly, taking advantage of weak enforcement structures, a lack of transparency, or market uncertainty. Even if there is no explicit communication or coordination among enterprises, their activities might be consistent with collusion and impair customer welfare. Likewise, [119] identify "market rules, cost recovery, transmission congestion, lack of dispatchable demands, unbalanced demand and supply, market power, volatility, and uncertainty" as the variables that can cause price surges, and propose various monitoring indicators, such as clearing price and transmission congestion, in order to maintain a dynamic market.

Besides market shortcomings and failures, extreme fluctuations in prices across different sources of production can directly influence and contribute to higher electricity costs, highlighting the interconnected nature of energy markets. An early example of this is found in [115], where the authors explore how the transition to natural gas as a key power source in the United States may be properly managed if unexpected gas price fluctuations occur. In the years since then, this has been extended to other sources of generation.

3.2.2 Early experiences in Europe and Australia

In many aspects, early European experiences with market reform mirrored those in North America. To shed light on market design limitations and conditions, such as weather variability and power commodity storage limitations, leading to its potential failure [110], [117] examined the UK deregulation model and compared it with the Nordic approach. The authors claimed that since the Nordic model contains both massive hydropower sources and high-quality spot prices, modifications could be made quickly to prevent price spikes and improve financial contract pricing. Consequently, in order to create an effective financial power market in the UK, the authors suggest that several factors must be taken into consideration, including the type of power pool model utilized, the implementation of price indexes to ensure transparency within the market, the use of simple derivative contracts, and the ability to adjust the power supply source, such as through the utilization of a hydro-based system [117]. Additionally, to bolster market flexibility, Burger et al. advocate optional contracts and enabling fixed-rate energy purchases for defined periods [110].

A few years later, highlighting issues with the Nordic approach, a different set of researchers used market deregulation failures in Sweden to identify several factors that contribute to the long-term increase in energy prices [109]. These include unclear management strategies for hydroelectric resources, inadequate investment in domestic power infrastructure, excessive energy taxes, and a policy of phasing out nuclear power. However, this long-term increase in price is different than increasing volatility or spikiness in prices. In the years since the EU has also made remarkable progress towards a single market (also known as the single day-ahead coupling or SDAC). An example of this market coupling can be seen in the very similar prices in Belgium, France, and Germany, as seen in Figs. 1 and 5.

Focusing on the Australian market, which is much smaller compared to the European and American markets and consequently, the factors for price spikes are unsurprisingly a bit different such as market manipulation or exercising power during high-demand/low-supply periods that cause supply curve adjustments, through withholding the reserved capacity or bidding strategies aimed at pushing the prices up [106, 81, 84]. Additionally, extreme weather events [106, 84], underestimation of demand, inventory constraints [106] and strong inelastic demand [84] have been among spike triggers in this market as well.

3.2.3 Recent perspectives

In recent years, researchers have considered additional elements influencing price surges, for instance the history of spikes leading to a high level of persistence for future spikes [93], the time interval between prior spikes impacting the likelihood of future severe observations [93, 88], and duration and magnitude of the spikes [88]. Thus, peaks occur in bursts, indicating that values above the threshold tend to occur in succession [88]. Building on previous research, Mauritzen et al. [91] further include unexpected power disruptions and shortages, the inability to store electricity, retailers' inflexible demand, and the bidding strategies of auctioneers as possible factors causing price spikes.

Maryniak et al. [42] suggest that low reserve margins are a significant driver of electricity price hikes. Reserve margins refer to the excess generating capacity available in the system above and beyond what is necessary to meet demand under normal operating conditions. When reserve margins are low, unexpected events such as power plant failures or extreme weather can cause a shortage in supply, leading to a surge in prices. However, the availability of reserve margin data varies among markets, and in many cases, this information may not be easily accessible. A somewhat related measure is the "load-to-capacity ratio" (LCR), which is used by [24] to show its correlation with the magnitude of price surges. Notably, the study identifies that extreme situations, characterized by significant wind generation losses during periods of peak demand, serve as the primary catalysts behind sudden price spikes. This finding emphasizes the critical role of wind power and the limited influence of nuclear generation and natural gas prices in the occurrence of extreme price events [24]. Likewise, a recent study by Bajai et al. [11] reveals that certain variables, such as liquidity ratio [4] and high net exports also have a significant impact on the higher conditional probability of spikes. Very recently, a study by Galarneau et al. [3] suggests that the occurrence of price spikes can also be attributed to the delayed reaction of major low-cost electricity producers.

Furthermore, validating earlier concerns on market operation, high gas, and fuel price spikes have contributed greatly to the electricity price spikes seen in Europe. As a result, average electricity prices in most European countries soared by more than five times on average in 2022 (see Fig. 1 for an overview). In this

[4] Bid-ask spread

figure, France and Belgium stand out in particular as prices surged by almost an order of magnitude from 32 euros/MWh on average in 2020 to almost 250 euros/MWh in 2022. In the coming years, the Belgian transmission system operator Elia expects average prices to fall once again depending on several factors, contingent primarily on gas and CO2 prices. In case of high gas and CO2 prices, electricity prices might converge at over 100 euros/MWh, which is two- to three-fold higher than its historical value. Only in the case of low fuel and CO2 prices will electricity prices return to their long-term historic average of less than 50 euros/MWh [2]. The same holds in many other countries which still rely on fossil-based generation to meet either the base load or peak demand.

Table 1 summarizes the different factors we have identified in this section and categorizes them in broader groups. This categorization reveals a total of 34 factors that contribute to extreme price increases. The majority of these factors are associated with groups G2 (market features, i.e. the characteristics or dynamics inherent to the electricity market itself) and G3 (market participants' actions or decisions). In addition to enabling researchers to better understand price spikes, many of these features also serve as the foundation for building forecasting models, a topic we will return to in section 4 in greater detail.

Table 1: Contributing factors to positive upward spikes

Group	**Factor**
G1: commodity features	Non-storability [91, 110], length of spikes [88], volatility [119]
G2: market features	Market rules [119], lack of transparency [107], transmission congestion [119], improper market design [118, 119], supply-demand imbalance [119, 118], liquidity [11], inventory constraint [106], low reserve margin [42], load-to-capacity ratio [24], weak enforcement structure [107]
G3: participant actions / decisions	Bidding strategies [116, 91, 81], market power [119, 118], unclear hydro-power management strategies [109], insufficient investment in domestic energy infrastructure [109], large reserve capacity [106], delayed reaction of low-cost producers [3], pricing strategies [113], financial insolvency [119, 118], underestimation of demand [106], supply curve adjustment [84], retailer inflexible demand [91]
G4: political considerations	Excessive energy taxes [109], nuclear disengagement [109]
G5: temporal effects	Inelastic demand [114, 84], lack of dispatchable demand [119], history of spikes [93], time interval between prior spikes [93, 88]
G6: unexpected incidents	Extreme weather conditions [110, 106, 42, 84], power outages [91, 42], uncertainty [119]
G7: external factors	Gas prices [115], electricity trade disparity[11]

3.3 Factors triggering downward price spikes

Historically, negative prices have been a rare occurrence in most markets. However, many of the conditions discussed above, which lead to positive price spikes, can also lead to negative price spikes when inverted. For instance, an ongoing surplus of renewable sources, such as wind and solar, can lead to sustained periods of very low electricity prices. This has been shown in research by several authors. Notably, Nicolosi et al. show that a significant amount of intermittent renewable energy sources being fed into the system during periods of low demand leads to negative or suppressed prices [98]. Additionally, limited flexibility of the power system components and a lack of alignment between power markets can also exacerbate this. Similarly, [64] highlights the connection between renewable energy integration and extremely low prices, as well as the need for the power system to become more flexible to tackle these challenges (these were also identified as challenges in the previous section).

One example of this phenomenon has been in the low prices observed in the Spanish electricity market during the last decade (i.e. prior to COVID-19 disruptions). To understand the factors contributing to this trend, researchers have identified several factors, including the rise of low-cost renewable energy, the retirement of thermal power plants, and the limited capacity for interconnection to handle surplus renewable energy [69]. Additionally, the authors note that price-taking technologies such as zero-variable-cost run-of-river hydro and inflexible nuclear power plants are able to produce electricity at a very low marginal cost, which puts downward pressure on prices. In a similar vein, [89] claims that the expansion of renewable electricity production is the primary cause of negative price fluctuations in the German electricity market besides other factors, such as low demand meeting high supply, sudden declines in industrial activity, limited flexibility in power plant operations, and constrained transmission capacities. However, the research findings imply that when wind generation increases, it tends to lower the average spot prices, affecting both low and high-price components [24].

More recently, in the aftermath of large-scale demand destruction due to COVID-19, electricity market prices fell precipitously in many countries across the world. In the US, average electricity market prices fell to \$21/MWh, the lowest in the twenty-first century [29]. In fact, negative prices were observed in 4% of all hours, either due to transmission interconnection limits leading to local supply-demand imbalance or due to system-wide oversupply. Very low fossil fuel prices along with a rapidly expanding renewable build-up contributed greatly as well. Similar observations have been reported for many European markets with the German day-ahead market prices dropping by over a half in 2020 when compared to 2019 [25]. Many of the factors identified in this section have meant that, even during the unusually high prices witnessed in 2022, electricity prices were negative for a considerable amount of time (this is true, especially on sunny and windy days with limited demand).

To summarize, our analysis has identified 10 elements that contribute to significant downward price movements, which are summarized in Table 2. Notably, none of these factors can be attributed to the decisions made by market participants (G3).

Table 2: Contributing factors to downward price spikes

Group	**Factor**
G1: commodity features	Non-storability [91, 110]
G2: market features	Low demand [89], insufficient interconnection infrastructure to accommodate excess renewable energy/ constrained transmission capacity [69, 89], low-cost renewable energy [69, 64, 89, 98, 24]
G4: political considerations	Retirement of thermal power plants [69]
G5: temporal effects	Limited flexibility of power system components/operations [89, 98], inflexible nuclear power plants [69]
G6: unexpected incidents	Sudden declines in industrial activity [89]
G7: external factors	Price taking technologies [69]

3.3.1 Evidence from the European DA market

Despite increasing price harmonization across European DA markets, considerable price differences can exist due to the factors discussed previously (excess or shortfall of renewable generation coupled with interconnection bottlenecks invariably play a large role here). Fig. 5 provides an illustrative example by comparing the distribution of hourly DA prices in four different EU countries, France, Spain, Sweden (BZ1), and Belgium, for the years 2020 and 2022 [5]. The plots show that while the different markets are coupled to an extent, there is no guarantee that a price spike in Belgium will also translate to a spike in Spain or Sweden. In fact, we see clear bimodal distributions in many markets such as Spain and France in 2022, where prices are either harmonized or much higher in France. The same behavior is also evident between Sweden and Belgium, and Sweden and France. If we consider price data from 2020, it is immediately evident that prices in all markets were lower and less spiky, and the correlation between different markets was also much higher.

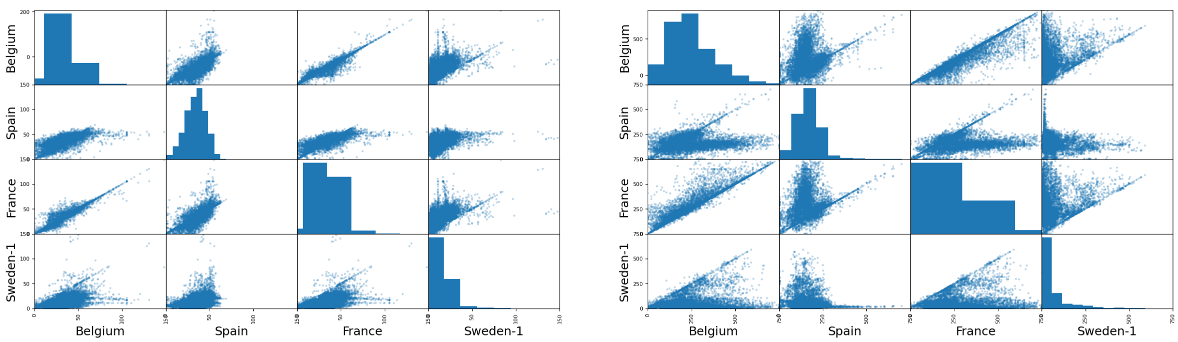

Figure 5: Distribution and association of hourly day-ahead prices for four EU countries in 2020 (left) and 2022 (right)

4 Forecast methodologies

We classify electricity price spike forecasting methods into four different categories, depending on their output. These include (1) point forecasts, (2) interval, quantile, and scenario forecasts, (3) spike probability forecasts, and (4) threshold forecasts. In doing so, we choose to focus on time series related methods, rather than game theoretic or simulation-based approaches [105].

4.1 Point forecast methods

Methods which produce a point forecast for day-ahead electricity prices have arguably been the most common in published literature [20, 30, 46, 45, 78, 83, 28]. Historically, these models typically belong to the general SARIMAX (seasonal autoregressive moving average with exogenous factors) family of algorithms. The simplest and most common variants of such techniques include the AR (autoregressive) and ARX (autoregressive with exogenous factors) model types. AR models rely solely on the time series' history when making forecasts, while ARX models additionally consider exogenous variables such as those discussed in the previous section as well (e.g. renewable generation, load, etc.). Some examples of such ARX models include [79] and [58]. It is straightforward to extend linear function approximation in SARIMAX models to utilize non-linear function approximators such as tree-based methods and neural networks. The latter, often referred to as NARX models [49, 56, 78, 95, 76, 26], have seen rapid growth in published papers in the last decade.

[5] The data used in this analysis was sourced from the website https://transparency.entsoe.eu/.

In addition to automatic feature selection in ARX models [28], researchers have also investigated several explicit feature selection algorithms. In this vein, the Heckman selection model provides a unified framework to investigate the factors influencing the frequency and magnitude of price spikes in electricity markets [83]. Using four regional electricity markets in Australia, the authors show that load, relative air temperature, and reserve margins can be used in such a scheme for forecasting price spikes [83]. Likewise, defining price spikes as above \$150/MWh and \$200/MWh, Amjady et al. show that their "MR-IG" feature selection method can choose a small subset of highly relevant characteristics for electricity price spike predictions in conjunction with a probabilistic neural network (PNN) [94]. This allows for a more accurate prediction of both regular and price spikes (when compared with eight frequently used methods on the MAPE metric). A related matter to feature selection is feature transformation. In this regard, researchers have explored several different methods ranging from simple min-max scaling to variance stabilizing transforms [68, 31]. While more complex transformations have shown promise to improve forecast accuracy in general, their effect on accurately predicting price spikes remains unclear.

Another technique that has been employed by several researchers is to first decompose the price signal into its constituent components, and then use the components directly in the prediction pipeline, followed by a recombination step. In this vein, the so-called "seasonal component autoregressive with exogenous factors (SCARX)" model has been proposed by Afanasyev et al. to forecast day-ahead electricity prices, using the day-ahead load forecast as the exogenous variable in their study [45]. The authors apply their method to four distinct power markets and show that forecast accuracy can be improved by filtering the input time series data [6]. Likewise, using data from the UK, Zafar et al. propose decomposing the price time series into different trend series of varying frequencies, which are then forecast using a pure AR model [20]. A similar scheme is adopted by [21], albeit in a two-step approach, and will thus be covered in a later section.

Such decompositions, even though they may improve average price prediction performance, should be used with caution as they often end up filtering out high-frequency data, which can be critical in forecasting price spikes or extreme events. In fact, most of these studies do not explicitly consider price spikes. Instead, they focus on predicting an 'expected' outcome, often by minimizing the mean squared or absolute error. In fact, as highlighted by Lu et al. [106], many point forecasting models only perform at a competitive level in predicting prices when extreme observations are removed from the data. Consequently, researchers have increasingly turned to quantile and interval price forecasts, which we discuss next.

4.2 Interval, quantile, and scenario forecast methods

As opposed to point forecasts, these methods tend to provide more information on the actual distribution of price forecasts, which can enable their use in risk-based planning. In practice, these methods can produce different outputs. One option is to produce the entire probability distribution. Another common approach is to produce only a few specific quantiles, say the 10th and 90th quantiles. One final approach is to produce a set of scenarios, along with their probability of occurrence. We discuss these methods in greater detail in this section.

[6] The detected extreme values were substituted with the sample mean of prices with seasonal adjustments.

4.2.1 Quantile forecasting

In an early study, Lu et al. propose a Bayesian model to forecast price spikes, based on data from the Queensland electricity market [106]. This method classifies prices into five categories of spikes[7], and utilizes a wavelet and neural network-based model to make forecasts. This results in a price spike prediction along with the level of spike and the associated confidence level. In the ensuing years, the combination of (wavelet-based) decomposition and neural networks has been reincarnated several times to make probabilistic forecasts with varying degrees of rigor [77, 65, 67]. While these studies show promising results, they are often conducted on limited test periods lasting well short of the minimum recommended one-year period and are not benchmarked against other modern methods, such as the ones we highlight next.

In the context of the Global Energy Forecasting Competition 2014 (GEFCom2014) probabilistic price forecasting track, Dudek et al.[71] proposed a feedforward neural network to predict day-ahead electricity prices, incorporating historical price data, zonal load, and system load as predictive factors. Notably, this approach, combined with Bayesian regularization, demonstrated the most promising performance, particularly in accurately forecasting price spikes. Likewise, using the GEFCom 2014 dataset, Gaillard et al. examine three distinct quantile forecasting approaches [72]: a quantile generalized additive model, a combination of 13 predictors including autoregressive models, and a kernel-based quantile regression with Lasso penalty customized for winter conditions. Using the average pinball score, the results indicated that the first approach showed higher robustness to price spikes, suggesting its superiority in capturing extreme price movements and providing reliable quantile forecasts. In a subsequent study, Dudek et al. [55] introduce a method for forecasting 99 quantiles of day-ahead and imbalance electricity prices in Poland, employing the Nadaraya-Watson estimator, a non-parametric kernel regression technique. Initially, the method generates point predictions and then extends them to quantile forecasts using historical price residual distribution. The proposed model outperforms exponential smoothing, naive, and ARIMA models in the evaluation.

Ziel and Steinert [80, 62] proposed a method to directly model the sale and purchase curves to forecast electricity prices, and apply it to the German and Austrian day-ahead prices. However, despite the innovative approach, the model underestimates the occurrence of spikes. Very recently, an interesting method using "conditional time series generative adversarial networks" (CTSGANs) has been investigated by [15] to obtain prediction intervals for Australian spot prices. Unfortunately, the results provide no indication of whether the proposed method is better at capturing spikes than more established methods highlighted elsewhere in the paper.

A method that has gained in popularity recently is to combine several point forecasts to obtain the quantiles for the predicted price [92, 74, 50, 41, 37]. The point forecast models are typically created by utilizing a set of ARX (point forecast) models trained on similar input features (historical lags, load and renewable generation, etc.), but applying different calibration window lengths (ranging from several weeks to years). The different window lengths capture different information, and combining them to obtain a single-point forecast can significantly improve predictive accuracy [28]. However, it is also possible to use the outputs of these models to obtain quantiles using either Quantile Regression Averaging (QRA) or Quantile Regression Machine (QRM). QRA uses the point forecasts directly to obtain the quantiles, while QRM first averages them, and then applies quantile regression to the combined forecast. Several other variations exist as well: for instance, [74] computes the principal components of the point forecasts and uses these as input to the QRA

[7] Ranges lie between [75-100], [100-150], [150-250], [250-500], [500-2000] and greater than 2000

or QRM, as opposed to the raw point forecasts. Other mechanisms, that replace the learning algorithm (e.g. to obtain Quantile Regression Boosting or Quantile Regression Neural Networks), have been proposed as well [75]. Uniejewski et al. [51, 8] evaluated the ability of different quantile evaluation techniques to predict extreme quantiles (and thus spikes), and showed that QRA-based prediction techniques significantly outperformed other techniques.

Such methods offer two key benefits: (1) they do not require much additional machinery beyond the requirements for point forecast (beyond a separate calibration window for the probabilistic forecasts), and (2) they have shown promising results in terms of producing well-calibrated and sharp forecasts (these terms are explained in greater detail in a subsequent section). However, few, if any studies, have looked at using these methods for the recent volatility in the electricity markets.

4.2.2 Probability density function forecasting

In addition to predicting specific quantiles, researchers have also proposed methods which predict the entire probability distribution of day-ahead prices. This can be used to better understand the risk posed by exceptionally high or low prices - which may be missed by individual quantiles. While some researchers have taken this to mean predicting several, finely-spaced quantiles (e.g. the first quantile all the way to the 99th quantile), this approach often suffers from overlapping quantiles, which has to be corrected in a post-processing step. Other researchers have used predictive methods, which can directly output estimates of uncertainty. For instance, Gao et al. utilize "autoregressive Gaussian process regression", employing various sampling techniques to generate multiple scenarios of probabilistic prediction of California day-ahead electricity prices, with Monte Carlo Sampling [34]. The authors show that Contour Sampling prioritized extreme price scenarios by focusing on the tail of the Gaussian distribution.

Other approaches have been studied as well. For instance, in their day-ahead electricity price study for the Austrian-German market, Muniain and Zeil used autoregressive models with exogenous variables and jump-diffusion models [43]. Likewise, Kath and Zeil have applied conformal prediction (CP) to determine interval forecasts for the day-ahead markets, and have compared their results against previously mentioned techniques such as QRA [27]. They show how symmetric prediction intervals can be computed by exploiting the errors made based on point forecasts, and conclude that CP can achieve comparable performance to other, more established methods in yielding well-calibrated predictions that can be used to predict extreme spikes.

Some studies indicate that parsimonious stochastic models struggle to capture the complex and non-linear nature of electricity price time series [9]. To address this limitation and ensure appropriate management of data variability before modeling [47], there is a growing adoption of deep neural network-based research for probabilistic electricity price forecasting. Notable studies have investigated real data from diverse markets, including California ISO, Italian, Belgian, and Singapore electricity markets [9, 47, 35]. These studies utilize various approaches, such as "deep Gabor convolutional neural networks" that have been further developed into a mixture density network [9], "Bayesian deep learning" [47], and "deep convolutional neural networks" to extract features which were then fed into a decision tree to forecast the probability density function of electricity prices [35]. Bello et al., focusing on the month-ahead probabilistic forecasting of prices on the Spanish market, compare multilayer perceptrons with decision trees and a market equilibrium model with logistic regression [69]. Likewise, Campos et.al utilize a hybrid neural network model to generate probabilistic week-ahead forecasts for both the Spanish and PJM markets [54].

4.3 Spike probability forecasting

Besides predicting the price distribution or quantiles, some researchers have adopted a two-step process: the first step is to predict if a spike will occur, followed by the price prediction (possibly using a different model conditioned on the spike probability). In this direction, early work by Zhao et al. [108] and Wu et al. [104] proposed probability classification techniques to predict the occurrence of spikes using support vector machines. The method did not misclassify non-spikes but failed to detect about half of them.

Imbalanced data (i.e. non-spikes far outnumber spikes) is a recurring challenge in predicting price spikes. Different researchers have tried to address this using different methods. For instance, Amjadi and Keynia proposed a hybrid data model that combines time and frequency domain information [97]. The selected inputs from the feature selection technique, along with calendar indicators and an "Existence" feature were then input to a Probabilistic Neural Network (PNN) model [97], which was used to predict price spike occurrences for both the Queensland and PJM markets. Recently, borrowing from the broader field of machine learning, researchers have addressed this imbalance issue using data augmentation techniques such as SMOTE and borderline-SMOTE [21, 31]. The idea behind these methods is to generate additional samples in the vicinity of observed spikes in the training data to reduce class imbalance.

To underscore the significance of load and capacity as key drivers of exceptionally high prices, in line with prior studies, Maryniak and Weron [42] proposed utilizing the demand-to-capacity ratio to forecast the probability of price spike occurrence in the UK market based on the curvature of the hyperbolic tangent, a methodology developed by [100]. They concluded that the likelihood of observing a price spike increases with an increase in the demand-to-capacity ratio, regardless of the method employed. Likewise, other researchers have also explored the use of load to predict the likelihood and severity of spikes [93, 81]. In [81], a multivariate self-exciting point process can accurately predict the probability of extreme prices with the benefit of including the impact of excess capacity of interconnected regions, while in [93] an ARMA process in combination with autoregressive conditional hazard can explain the persistence of spikes and intensity. This model improved performance when compared against a simple benchmark logit model, including only load and temperature variables.

Analysis in Hungary by Bajai et al. [11] concurs with Clements et al. [81] in that power exchanges between markets are beneficial in predicting extreme prices and they have applied a logit model which takes into account market liquidity, renewable generation, and net interchanges. Likewise, Lu et al. focus on forecasting normal regional reference prices for Queensland using a wavelet neural network, while predicting spike occurrence through Bayesian classification and k-closest neighbor estimation, considering demand and supply relationships [106]. Finally, Datta et al. also account for seasonal impacts (e.g. monthly and weekly events in the Victoria region), meter uncertainty, and past spike influences when predicting price spikes [70]. Their findings show that incorporating meter data and using the Naive-Bayes classifier significantly improves spike prediction accuracy, outperforming Random Forest Classifiers.

Finally, researchers have also considered the effect of transmission and distribution bottlenecks in forecast models on day-ahead prices. For instance, Galarneau et al. investigate the influence of transmission capacity constraints between cost-effective power plants and the Long Island peninsula in the NYISO market on market prices [3]. To do so, they employed four different techniques to forecast the probability of day-ahead minus real-time price spikes, i.e., logistic regression, random forests, gradient boosting trees, and "feed-forward deep neural networks" [3]. Among these, the second technique demonstrated a slight advantage. Additionally, trading strategies based on model-generated spike probabilities resulted in higher profits and lower risk, indicating the economic significance of the generated signal. Likewise, researchers have explored ideas from multi-task learning, i.e. by using the same forecast model to predict the prices in several markets simultaneously [58]. This induces cross-learning, which appears to be beneficial to the forecasting task while enabling the system to learn about underlying bottlenecks in the system. However, the performance of such models has not been investigated in great detail on spikes or extreme events.

4.4 Threshold and regime switching forecast methods

To effectively represent different states of electricity prices and capture short-lived spikes, several studies advocate for the implementation of regime-switching models [112, 101]. In a way, these methods are a natural extension to the methods discussed previously on forecasting the probability of price spike occurrence. For instance, Huisman and Mahieu propose a three-regime model using day-ahead baseload prices from the Dutch APX market, the German LPX market, and the UK market through Kalman filter methodology [112], while Becker et al. develop a two-state regime-switching model with time-varying probability specifically in the Queensland electricity market [101], and find that meteorological conditions and electricity demand factors significantly influence market regime tenacity and help to predict transitions between the two regimes. Furthermore, Doering et al. highlight that the load-to-capacity ratio impacts the mean price of regimes and the frequency of transitioning between them in a novel regression-based mixture model, enabling the analysis of low- and high-price regimes' average and frequency within the ERCOT ISO [8] [24]. These findings underscore the multifaceted nature of the factors driving price dynamics and transitions, as discussed in a previous section.

On the contrary, Eichler et al. assert that parametric Markov regime-switching models cannot account for the time-varying central moments of the spike distribution and argue that spikes are predictable and dependent on past data, occurring in blocks with persistence [88]. This stands in contrast to previous studies like [106], suggesting electricity price spikes are random. The authors propose and evaluate several models, including Benchmark logit, Dynamic Hawkes logit/scobit, and Regime Switching logit/scobit, using various metrics such as predictive log-likelihood and regression loss functions to assess predictive power over AEMO data and conclude that certain logit models perform better in capturing the blocking and clustering characteristics of peaks as "spikes do not only occur in clusters, i.e., that there are times of high probabilities of spikes and times of low probabilities, but that in fact, spikes occur in blocks of varying length, i.e., that there are consecutive prices above the threshold" [88].

5 Evaluation metrics

Electricity price (spike) forecasting algorithms have been evaluated according to several different metrics in the literature. One fundamental difference in these metrics arises from whether the problem is posed as a regression or a classification problem, and whether the output is a point, quantile, or entire distribution. In this section, we describe some commonly used evaluation methods for the different forecast methodologies highlighted previously.

[8] Electric Reliability Council of Texas system

5.1 Metrics for point forecasts

By far, the most common strategy to evaluate electricity price forecasting models remains via some variation of comparing observed against predicted price. In fact, many studies including ones that emphasize accurate price spike forecasting do not evaluate performance on spike prediction separately. Consequently, the most popular choices in the regression-based category include the mean error (ME) or bias, the mean absolute error (MAE), the mean squared error (MSE), and the root mean squared error (RMSE) [30, 46, 103, 78, 83, 13, 6, 28, 40]. Some researchers have also explored the use of variants of MAE such as weekly-weighted MAE [20, 45].

While all of these metrics compare observed and predicted prices, it is important to note the individual idiosyncrasies of these evaluation metrics. For instance, ME provides an indication of whether the forecast is biased towards making over- or under-predictions, while the other metrics do not. Likewise, metrics utilizing a squared term (e.g. MSE and RMSE) give a larger weight to large errors, which can occur when prices take on extreme values, as compared to absolute metrics (e.g. MAE). All of these metrics also depend on the magnitude of the time series, which makes it difficult to contextualize the errors, i.e. a forecast error of 20 EUR/MWh might be considered a good result when evaluating a very high price spike, but not in regular operating conditions. This property also makes it difficult to compare predictive accuracy across time and/or different markets, where prices may inherently be different due to different demand or supply mix etc.

Often, to address these issues, scale-independent metrics such as mean absolute percentage error (MAPE), R^2, and adjusted R^2 [96, 30, 46] are adopted as well. These metrics, however, come with their own limitations. Most notable amongst these is that MAPE can take on undefined or very high values for prices close to 0. Researchers have traditionally relied on weighted variants of MAPE to address this issue, but there has been an increasing shift towards scaled and relative metrics, which instead normalize prediction accuracy by the accuracy of a baseline method, often a seasonal naive model [5]. This has been seen for price forecasting in several recent studies [28, 21]. Unfortunately, most of these evaluation metrics do not focus specifically on spike forecasting. More concretely, they are averaged metrics, which do not necessarily account for rare, extreme conditions.

5.2 Metrics for quantile and probabilistic forecasts

Evaluating interval forecasts is inherently more challenging than evaluating point forecasts. This is because even though we observe the realized market price, we do not observe the entire distribution at each time step. Consequently, researchers have developed alternative methods to evaluate quantile and interval forecasts. For these forecasts, we can represent the forecast quantiles of y_t, to be exceeded by probability $1-p$, as $Q_t(p)$. If $G(p)$, defined as the proportion of times y_t is less than $Q_t(p)$, is 'reasonably close' to p, then $Q_t(p)$ is an accurate forecast distribution [1]. This similarity or closeness can, in turn, be evaluated using a variety of techniques ranging from the Kolmogorov-Smirnov test to the mean absolute excess probability. The same method can be extended to individual quantiles which can be evaluated separately in a hits-and-misses framework, which determines whether the observed data distribution is identical to a specific quantile prediction. These metrics however focus on reliability or calibration, i.e. whether the predicted intervals actually match observations, using conditional and unconditional tests. Nowotarski et al. also identify sharpness as a key concern when making interval forecasts [59], which refers to the width of the confidence intervals. Narrower, but equally well-calibrated, intervals should be preferred to wider ones.

In recent years, researchers have used several concrete metrics to evaluate these forecast methods, e.g. the quantile or pinball score has been used to evaluate individual quantiles, the Winkler score has been used to evaluate prediction intervals, as opposed to individual quantiles [121], and proper scoring rules like continuous ranked probability score (CRPS) have been used for the whole forecast distribution, rather than particular quantiles or prediction intervals [82]. For instance, in the context of electricity prices, the pinball loss has been used by Marcjasz et al. and Uniejewski et al. for the Nordpool market [41, 51], Uniejewski et al. for the German, Spanish, Nordic and PJM markets [8], Muniain and Zeil for Austrian-German market [43], and Dudek et al. for the Polish market [55]. Likewise, Lu et al. used the Winkler score for the Australian market [15], and Maciejowska et al. used it for the British market [74]. Finally, the CRPS is used by Janke et al. for the German-Austrian market [37] and in the form of the Energy Score[9] by [43] which is recommended due to the strong correlation between peak and off-peak time series, distinguishing double-spike events, Lu et al. for the Australian market [15], Moreira et al. for the Iberian market [75], Brusaferri et al. for the Italian and Belgian markets [47], and Afrasiabi et al. for the Californian market [9]. Other methods, which are often variations on these metrics, have been employed by researchers as well. He et al., for instance, examine the width of prediction intervals to evaluate the performance of their forecasts [35]. Likewise, Tahmasebifar et al. and Toubeau et al. use related metrics such as the average coverage error (ACE), coverage width-based criterion (CWC), and the prediction interval normalized average width (PINAW) to evaluate their forecasts [67, 44].

While the adoption and evaluation of probabilistic forecasts in the electricity price forecasting community is still rather limited, it is on an upward trajectory [59]. A key limitation here remains that evaluating the entire distribution of the forecasts is not straightforward in non-stationary settings. Furthermore, even well-calibrated quantile or distribution forecasts are not necessarily sufficient to fully capture the risk posed by extremely high or low price spikes.

5.3 Metrics for threshold and occurrence forecasts

Threshold forecasting models are trained to output discrete categories and are evaluated accordingly. In the limit, where only spike occurrence is being forecast, the threshold forecasting problem reduces to a simple binary classification task. As described earlier in the paper, this workflow typically follows a two-step approach and must be evaluated as such as well. First, the accurate prediction of spike occurrence needs to be evaluated. Second, the accurate prediction of the price spike, given the spike occurs, also needs to be determined. Treating price spikes in this way has the potential to enable practitioners to define their risk tolerance levels when dealing with market uncertainties.

However, there are several challenges with the evaluation phase; i.e. it might be tempting to only consider the accuracy of spike occurrence prediction. However, this does not account for the trade-off between precision and recall [102], as well as the fact that price spikes are rare events by definition. Consequently, most researchers account for this in their evaluation of price forecasting algorithms, whether they consider binary categorization or threshold forecasts. Some examples include the use of confusion matrices [103], F1 score [21, 32, 31], as well as the related AUC-ROC (Area Under the Curve and the Receiver Operator Characteristic) metric [3, 11, 32]. Another challenge when evaluating such methods is in the creation of the bins or thresholds themselves as

[9] A loss function in the context of multivariate probabilistic forecasting [37]

discussed in section 2. This is a concern, especially in non-stationary settings where the thresholds and bins identifying spikes during the evaluation will change over time, making temporal comparisons difficult.

5.4 From metrics to tests

Despite the existence of the metrics described above, evaluating electricity price forecasting models in a standardized manner remains a formidable challenge [28]. There are a number of reasons for this:

1. Different studies consider different evaluation metrics for different markets spanning different time periods; this makes it almost impossible to compare the reported evaluation metrics;
2. Many studies do not consider recommended techniques such as time series cross-validation [5], which raises data leakage concerns, and limits their applicability in real-world conditions;
3. The evaluation period is often limited to a few weeks, which is generally insufficient to capture seasonal trends;
4. Many studies test a very large number of forecast models on limited evaluation datasets. Owing to the very nature of the forecasting task, some models will perform better than others, but this will be an artifact of finite evaluation datasets, rather than any intrinsic skill of the winning forecast models.

In light of these concerns, a final post-processing step is increasingly applied to determine whether one forecast is better than a different one in a statistically significant manner, or if the difference is caused only due to random chance. Some tests that have been commonly used by the electricity price forecasting community include the Diebold-Mariano test (for comparison between two forecast models), and the Nemenyi test (for comparison across several forecast models) [21, 20, 28]. Unfortunately, due to the rarity of price spikes, it becomes even more challenging to draw statistically significant conclusions about the efficacy of methods. Furthermore, these tests are less commonly found in probabilistic or threshold forecasting studies.

6 Discussion

The field of electricity price (spike) forecasting has seen considerable attention and advances in the last few years. In this section, we discuss the most important insights drawn from the large-scale literature review carried out in this paper, both to identify promising future research directions as well as gaps in the current state of the art.

6.1 Defining spikes

Section 2 provided a detailed overview of the different ways in which researchers have approached defining price spikes. However, as discussed, these methods suffer from several shortcomings. In addition to the non-stationarity of the price time series, which was identified as a big challenge, the definitions are typically context-independent, i.e. a single numerical threshold cannot capture whether market prices are abnormally high, or if they are simply reflecting a well-functioning market. In this sense, we expect the field to move increasingly towards multi-variate, context-aware definitions for electricity price spikes in the future. One additional factor complicates the definition of thresholds in real-world markets. When applying thresholding techniques - either static or dynamic - an additional issue arises: in most liberalized markets, the maximum and minimum prices are not fixed, and may change over time. This directly affects not just the data percentiles, but also what can be realistically seen as an outlier. For instance, for the single day-ahead coupling (SDAC) European region, ACER (Agency for the Cooperation of Energy Regulators) had capped the maximum clearing price at +3,000 EUR/MWh and the minimum price at -500 EUR/MWh in Article 3 of the Annex I to the Decision No 04/2017 of 14 November 2017. This upper threshold of +3000 EUR/MWh was held until April 2022, when it was raised to 4,000 EUR/MWh due to market rules.

6.2 Choice of covariates

While section 3 identified a large number of (overlapping) covariates which are important for both positive and negative spike prediction, this does not necessarily translate to the actual studies we reviewed in this paper. In fact, in most studies, only a limited number of features are used for spike forecasting. In certain extreme cases, no covariates at all, except the history of the price time series, were considered. With increasing open-source datasets (e.g. on ENTSO-E transparency platform containing EU-wide forecasts for covariates of interest [14]), and computational resources available to train and test electricity price forecasters, this omission is not tenable anymore. This is especially true for creating forecasters, as blind spots in influencing factors can lead the model to fail at capturing the extreme tails of the price distribution. Several studies have already demonstrated that feature selection techniques can be used quite effectively with large covariate sets to create state-of-the-art forecasters.

6.3 Selection of pre-processing and forecasting methodology

As highlighted in the paper, researchers have investigated the usefulness of several forecasting technologies and pre-processing choices. Going against the grain in energy forecasting [16], neural architectures continue to boast state-of-the-art performance [28]. On the one hand, this is arguably due to the long tradition of using neural networks to predict prices (at the expense of other techniques such as gradient-boosted trees which have shown comparable or even better performance for time series forecasting). At the same time, it is unclear if these results hold for probabilistic forecasting or forecasting extreme spikes, where cross-comparisons of different techniques are more difficult to come by. Other open questions abound as well: do variance stabilizing transforms, which have been shown to work well for point forecasting, also provide improvements for predicting extreme quantiles? Likewise, detrending or deseasonalizing data before feeding it to a function approximator has a long but chequered history. In the coming years, these questions will need to be answered in much greater detail.

6.4 Evaluating forecasts

In addition to the metrics which have been adopted by the research community at large, there is a need for a broader assessment of the downstream risks misforecasts pose to system operation and the utility of different stakeholders. At the moment, only very few articles go into this direction, most limit their attention only to the production of the forecast (rather than its use). This is perhaps also influenced by the short evaluation periods and focus on singular markets chosen by many studies. These present the risk of spurious conclusions, which is reflected in different papers coming to contradictory conclusions. Some methodological limitations also persist here: there are few statistical tests for probabilistic forecasts especially in the presence of strong non-stationarity. This makes spike forecasting evaluation even more tricky than evaluating simple point forecasts for expected prices.

7 Conclusion

Electricity markets serve a pivotal role in the functioning of modern society - they provide appropriate signals for both consumers and generators to modulate their demand or generation, as well as guide long-term decisions on investments in the energy grid. These markets are typically split into many different products, based on the time of delivery. The focus of this paper has been on the day-ahead market, which is among the most actively traded markets in most places with liberalized electricity markets. In doing so, the paper has carried out an extensive review of the existing literature on electricity price spike forecasting. It shows increasing attention has been paid to the topic of spikes in recent years, and as evidenced by recent surges in prices (and their sticky nature), it appears that this will gain even more attention in the years to come.

Existing studies have considered electricity markets around the world, and early experiences and fears around the influence of the fuel mix on market prices seem to have come full circle in the last few years. At the same time, several limitations exist in published literature. The most obvious among these is the lack of an overarching framework showing how effectively electricity price spikes can be forecast. This means there is a lack of consensus on which forecasting techniques or pre-processing methods work best. Some researchers have proposed filtering spikes in the data fed to the forecasters, but this can have deleterious consequences for price spike forecasting accuracy. This lack of clarity is further exacerbated by the fact that many studies present their results for only limited time periods (lasting well short of the recommended minimum of at least one year), and specific quantiles without motivating this choice. Standardizing study design for price spike forecasting, in a similar vein to that seen for point forecasting in the EPFToolbox [28] can alleviate this situation.

Acknowledgments

Hussain Kazmi acknowledges support from the Research Foundation – Flanders (FWO), Belgium (research fellowship 1262921N). Samaneh Sheybanivaziri acknowledges that was awarded a scholarship by the Norwegian School of Economics to support research during the Ph.D. program. The authors would like to extend their sincere appreciation to Professor Jonas Andersson, for invaluable guidance and constructive comments that significantly improved this review study.

"**Statement: Declaration of generative AI and AI-assisted technologies in the writing process:** during the preparation of this work, the authors used ChatGPT to improve language and readability. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication."

References

[1] R. J. Hyndman, *Evaluating extreme quantile forecasts*, https://robjhyndman.com/talks/ISF11_RobHyndman.pdf, Accessed: [August 21, 2023].

[2] ELIA, "Adequacy and flexibility study for belgium (2024 - 2034)," 2023.

[3] R. Galarneau-Vincent, G. Gauthier, and F. Godin, "Foreseeing the worst: Forecasting electricity DART spikes," *Energy Econ.*, vol. 119, p. 106 521, Mar. 2023.

[4] Y. Guan *et al.*, "Burden of the global energy price crisis on households," *Nature Energy*, vol. 8, no. 3, pp. 304–316, 2023.

[5] H. Hewamalage, K. Ackermann, and C. Bergmeir, "Forecast evaluation for data scientists: Common pitfalls and best practices," *Data Mining and Knowledge Discovery*, vol. 37, no. 2, pp. 788–832, 2023.

[6] E. C. Karakoyun, H. Avci, A. S. Kocaman, and E. Nadar, "Deviations from commitments: Markov decision process formulations for the role of energy storage," *International Journal of Production Economics*, vol. 255, p. 108 711, 2023.

[7] M. L. Sørensen, P. Nystrup, M. B. Bjerregård, J. K. Møller, P. Bacher, and H. Madsen, "Recent developments in multivariate wind and solar power forecasting," *Wiley Interdisciplinary Reviews: Energy and Environment*, vol. 12, no. 2, e465, 2023.

[8] B. Uniejewski, "Smoothing quantile regression averaging: A new approach to probabilistic forecasting of electricity prices," *arXiv preprint arXiv:2302.00411*, 2023.

[9] M. Afrasiabi, J. Aghaei, S. Afrasiabi, and M. Mohammadi, "Probability density function forecasting of electricity price: Deep gabor convolutional mixture network," *Electric Power Systems Research*, vol. 213, p. 108 325, Dec. 2022.

[10] M. B. Anwar *et al.*, "Modeling investment decisions from heterogeneous firms under imperfect information and risk in wholesale electricity markets," *Applied Energy*, vol. 306, p. 117 908, 2022.

[11] M. Bajai, A. A. Vıg, and O. Hortay, "Electricity market liquidity and price spikes: Evidence from hungary," *Periodica Polytechnica Social and Management Sciences*, vol. 30, no. 1, pp. 49–56, 2022.

[12] Elsevier.com, *Analyze search results*, https://www.scopus.com/, Accessed on January 10th, 2023, 2022.

[13] K. Jin and C. Azuka, "Modeling electricity prices in the german energy market-with applications to renewables," 2022.

[14] H. Kazmi and Z. Tao, "How good are tso load and renewable generation forecasts: Learning curves, challenges, and the road ahead," *Applied Energy*, vol. 323, p. 119 565, 2022.

[15] X. Lu, J. Qiu, G. Lei, and J. Zhu, "Scenarios modelling for forecasting day-ahead electricity prices: Case studies in australia," *Applied Energy*, vol. 308, p. 118 296, 2022.

[16] C. Miller, L. Hao, and C. Fu, "Gradient boosting machines and careful pre-processing work best: Ashrae great energy predictor iii lessons learned," *arXiv preprint arXiv:2202.02898*, 2022.

[17] S. Numminen, S. Ruggiero, and M. Jalas, "Locked in flat tariffs? an analysis of electricity retailers' dynamic price offerings and attitudes to consumer engagement in demand response," *Applied Energy*, vol. 326, p. 120 002, 2022, ISSN: 0306-2619. DOI: https://doi.org/10.1016/j.apenergy.2022.120002. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261922012594.

[18] O. Ruhnau and S. Qvist, "Storage requirements in a 100% renewable electricity system: Extreme events and inter-annual variability," *Environmental Research Letters*, vol. 17, no. 4, p. 044 018, 2022.

[19] R. Tawn and J. Browell, "A review of very short-term wind and solar power forecasting," *Renewable and Sustainable Energy Reviews*, vol. 153, p. 111 758, 2022.

[20] U. Zafar, N. Kellard, and D. Vinogradov, "Multistage optimization filter for trend-based short-term forecasting," *Journal of Forecasting*, vol. 41, no. 2, pp. 345–360, 2022.

[21] C. Zhang, Y. Fu, and L. Gong, "Short-term electricity price forecast using frequency analysis and price spikes oversampling," *IEEE Transactions on Power Systems*, 2022.

[22] I. K. Bazionis and P. S. Georgilakis, "Review of deterministic and probabilistic wind power forecasting: Models, methods, and future research," *Electricity*, vol. 2, no. 1, pp. 13–47, 2021.

[23] S. Boeve, J. Cherkasky, H. Schult, and M. Bons, "Asset study on dynamic retail electricity prices," *URL: https://op. europa. eu/s/wSUQ. European Commission, Luxembourg*, 2021.

[24] K. Doering, L. Sendelbach, S. Steinschneider, and C. Lindsay Anderson, "The effects of wind generation and other market determinants on price spikes," *Appl. Energy*, vol. 300, p. 117 316, Oct. 2021.

[25] S. Halbrügge, P. Schott, M. Weibelzahl, H. U. Buhl, G. Fridgen, and M. Schöpf, "How did the german and other european electricity systems react to the covid-19 pandemic?" *Applied Energy*, vol. 285, p. 116 370, 2021.

[26] C.-J. Huang, Y. Shen, Y.-H. Chen, and H.-C. Chen, "A novel hybrid deep neural network model for short-term electricity price forecasting," *International Journal of Energy Research*, vol. 45, no. 2, pp. 2511–2532, 2021.

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

[28] J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Applied Energy*, vol. 293, p. 116 983, 2021.

[29] J. Seel, D. Millstein, A. Mills, M. Bolinger, and R. Wiser, "Plentiful electricity turns wholesale prices negative," *Advances in Applied Energy*, vol. 4, p. 100 073, 2021.

[30] I. Shah, S. Akbar, T. Saba, S. Ali, and A. Rehman, "Short-term forecasting for the electricity spot prices with extreme values treatment," *IEEE Access*, vol. 9, pp. 105 451–105 462, 2021.

[31] W. Shi, Y. Wang, Y. Chen, and J. Ma, "An effective two-stage electricity price forecasting scheme," *Electric Power Systems Research*, vol. 199, p. 107 416, 2021.

[32] D. H. Vu, K. M. Muttaqi, A. P. Agalgaonkar, and A. Bouzerdoum, "A multi-feature based approach incorporating variable thresholds for detecting price spikes in the national electricity market of australia," *IEEE Access*, vol. 9, pp. 13 960–13 969, 2021.

[33] K. Cleary and K. Palmer. "US electricity markets 101." Updated Mar. 17, 2022. Publication: Explainer. Accessed: 2023-09-03. (Mar. 2020), [Online]. Available: https://www.rff.org/publications/explainers/us-electricity-markets-101/.

[34] X. Gao and A. W. Dowling, "Making money in energy markets: Probabilistic forecasting and stochastic programming paradigms," in *2020 American Control Conference (ACC)*, Denver, CO, USA: IEEE, Jul. 2020.

[35] H. He, N. Lu, Y. Jiang, B. Chen, and R. Jiao, "End-to-end probabilistic forecasting of electricity price via convolutional neural network and label distribution learning," *Energy Reports*, vol. 6, pp. 1176–1183, Dec. 2020.

[36] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy forecasting: A review and outlook," *IEEE Open Access Journal of Power and Energy*, vol. 7, pp. 376–388, 2020.

[37] T. Janke and F. Steinke, "Probabilistic multivariate electricity price forecasting using implicit generative ensemble post-processing," in *2020 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, IEEE, 2020, pp. 1–6.

[38] X. Jin, Q. Wu, and H. Jia, "Local flexibility markets: Literature review on concepts, models and clearing methods," *Applied Energy*, vol. 261, p. 114 387, 2020.

[39] A. Lewis, N. Bean, and G. Nguyen, "Bayesian estimation of trend components within markovian regime-switching models for wholesale electricity prices: An application to the south australian wholesale electricity market," *arXiv preprint arXiv:2009.07471*, 2020.

[40] A. Lucas, K. Pegios, E. Kotsakis, and D. Clarke, "Price forecasting for the balancing energy market using machine-learning regression," *Energies*, vol. 13, no. 20, p. 5420, 2020.

[41] G. Marcjasz, B. Uniejewski, and R. Weron, "Probabilistic electricity price forecasting with narx networks: Combine point or probabilistic forecasts?" *International Journal of Forecasting*, vol. 36, no. 2, pp. 466–479, 2020.

[42] P. Maryniak and R. Weron, "What s the probability of an electricity price spike? evidence from the UK power market," 2020.

[43] P. Muniain and F. Ziel, "Probabilistic forecasting in day-ahead electricity markets: Simulating peak and off-peak prices," *Int. J. Forecast.*, vol. 36, no. 4, pp. 1193–1210, Oct. 2020.

[44] J.-F. Toubeau *et al.*, "Capturing spatio-temporal dependencies in the probabilistic forecasting of distribution locational marginal prices," *IEEE Transactions on Smart Grid*, vol. 12, no. 3, pp. 2663–2674, 2020.

[45] D. O. Afanasyev and E. A. Fedorova, "On the impact of outlier filtering on the electricity price forecasting accuracy," *Appl. Energy*, vol. 236, pp. 196–210, Feb. 2019.

[46] I. L. Amerise and A. Tarsitano, "A new method to detect outliers in high-frequency time series," *International Journal of Statistics and Probability*, vol. 8, no. 1, 2019.

[47] A. Brusaferri, M. Matteucci, P. Portolani, and A. Vitali, "Bayesian deep learning based method for probabilistic forecast of day-ahead electricity prices," *Appl. Energy*, vol. 250, no. C, pp. 1158–1175, 2019.

[48] A. Creti and F. Fontini, *Economics of Electricity: Markets, Competition and Rules*. Cambridge University Press, May 2019.

[49] G. Marcjasz, B. Uniejewski, and R. Weron, "On the importance of the long-term seasonal component in day-ahead electricity price forecasting with narx neural networks," *International Journal of Forecasting*, vol. 35, no. 4, pp. 1520–1532, 2019.

[50] T. Serafin, B. Uniejewski, and R. Weron, "Averaging predictive distributions across calibration windows for day-ahead electricity price forecasting," *Energies*, vol. 12, no. 13, p. 2561, 2019.

[51] B. Uniejewski, G. Marcjasz, and R. Weron, "On the importance of the long-term seasonal component in day-ahead electricity price forecasting: Part ii—probabilistic forecasting," *Energy Economics*, vol. 79, pp. 171–182, 2019.

[52] Zedda and Masala, "Price spikes in the electricity markets: How and why," in *3 May 2018 3rd HAEE Annual Conference: Energy Transition: European and Global Perspectives*, unknown, Jul. 2019.

[53] S. Zedda and G. B. Masala, "Price spikes in the italian electricity market: An economic analysis," en, *MODERN ECONOMY*, vol. 10, no. 09, pp. 2145–2156, 2019.

[54] V. Campos, G. Osorio, M. Shafie-khah, M. Lotfi, and J. P. S. Catalao, "Short-term hybrid probabilistic forecasting model for electricity market prices," in *2018 Twentieth International Middle East Power Systems Conference (MEPCON)*, Cairo, Egypt: IEEE, Dec. 2018.

[55] G. Dudek, "Probabilistic forecasting of electricity prices using kernel regression," in *2018 15th International Conference on the European Energy Market (EEM)*, ieeexplore.ieee.org, Jun. 2018, pp. 1–5.

[56] K. Hubicka, G. Marcjasz, and R. Weron, "A note on averaging day-ahead electricity price forecasts across calibration windows," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 321–323, 2018.

[57] D. S. Kirschen and G. Strbac, *Fundamentals of power system economics*. John Wiley & Sons, 2018.

[58] J. Lago, F. De Ridder, P. Vrancx, and B. De Schutter, "Forecasting day-ahead electricity prices in europe: The importance of considering market integration," *Applied energy*, vol. 211, pp. 890–903, 2018.

[59] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renewable Sustainable Energy Rev.*, vol. 81, pp. 1548–1568, Jan. 2018.

[60] E. Trujillo-Baute, P. del Rio, and P. Mir-Artigues, "Analysing the impact of renewable energy regulation on retail electricity prices," *Energy Policy*, vol. 114, pp. 153–164, 2018.

[61] J. Villar, R. Bessa, and M. Matos, "Flexibility products and markets: Literature review," *Electric Power Systems Research*, vol. 154, pp. 329–340, 2018.

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

[63] A. Asadinejad and K. Tomsovic, "Optimal use of incentive and price based demand response to reduce costs and price volatility," *Electric Power Systems Research*, vol. 144, pp. 215–223, 2017.

[64] E. Kyritsis, J. Andersson, and A. Serletis, "Electricity prices, large-scale renewable integration, and policy implications," *Energy Policy*, vol. 101, pp. 550–560, Feb. 2017.

[65] M. Rafiei, T. Niknam, and M. H. Khooban, "Probabilistic electricity price forecasting by improved clonal selection algorithm and wavelet preprocessing," *Neural Computing and Applications*, vol. 28, pp. 3889–3901, 2017.

[66] T. Rintamaki, A. S. Siddiqui, and A. Salo, "Does renewable energy generation decrease the volatility of electricity prices? an analysis of denmark and germany," *Energy Economics*, vol. 62, pp. 270–282, 2017.

[67] R. Tahmasebifar, M. K. Sheikh-El-Eslami, and R. Kheirollahi, "Point and interval forecasting of real-time and day-ahead electricity prices by a novel hybrid approach," *IET Generation, Transmission & Distribution*, vol. 11, no. 9, pp. 2173–2183, 2017.

[68] B. Uniejewski, R. Weron, and F. Ziel, "Variance stabilizing transformations for electricity spot price forecasting," *IEEE Transactions on Power Systems*, vol. 33, no. 2, pp. 2219–2229, 2017.

[69] A. Bello, J. Reneses, and A. Muñoz, "Medium-Term probabilistic forecasting of extremely low prices in electricity markets: Application to the spanish case," en, *Energies*, vol. 9, no. 3, p. 193, Mar. 2016.

[70] A. R. Datta and S. Datta, "Electricity market price-spike classification in the smart grid," in *2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Sep. 2016, pp. 1–5.

[71] G. Dudek, "Multilayer perceptron for GEFCom2014 probabilistic electricity price forecasting," *Int. J. Forecast.*, vol. 32, no. 3, pp. 1057–1060, Jul. 2016.

[72] P. Gaillard, Y. Goude, and R. Nedellec, "Additive models and robust aggregation for GEF-Com2014 probabilistic electric load and electricity price forecasting," *Int. J. Forecast.*, vol. 32, no. 3, pp. 1038–1050, Jul. 2016.

[73] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecast.*, vol. 32, no. 3, pp. 914–938, Jul. 2016.

[74] K. Maciejowska, J. Nowotarski, and R. Weron, "Probabilistic forecasting of electricity spot prices using factor quantile regression averaging," *International Journal of Forecasting*, vol. 32, no. 3, pp. 957–965, 2016.

[75] R. Moreira, R. Bessa, and J. Gama, "Probabilistic forecasting of day-ahead electricity prices for the iberian electricity market," in *2016 13th International Conference on the European Energy Market (EEM)*, IEEE, 2016, pp. 1–5.

[76] I. P. Panapakidis and A. S. Dagoumas, "Day-ahead electricity price forecasting via the application of artificial neural network based models," *Applied Energy*, vol. 172, pp. 132–151, 2016.

[77] M. Rafiei, T. Niknam, and M.-H. Khooban, "Probabilistic forecasting of hourly electricity price by generalization of elm for usage in improved wavelet neural network," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 71–79, 2016.

[78] H. S. Sandhu, L. Fang, and L. Guan, "Forecasting day-ahead price spikes for the ontario electricity market," *Electric Power Systems Research*, vol. 141, pp. 450–459, Dec. 2016.

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

[80] F. Ziel and R. Steinert, "Electricity price forecasting using sale and purchase curves: The x-model," *Energy Economics*, vol. 59, pp. 435–454, 2016.

[81] A. E. Clements, R. Herrera, and A. S. Hurn, "Modelling interregional links in electricity price spikes," *Energy Econ.*, vol. 51, pp. 383–393, Sep. 2015.

[82] T. Gneiting and M. Katzfuss, "Probabilistic forecasting," *Annual Review of Statistics and Its Application*, vol. 1, pp. 125–151, 2014.

[83] R. Handika, C. Truong, S. Trueck, and R. Weron, "Modelling price spikes in electricity markets - the impact of load, weather and capacity," *HSC Research Reports*, May 2014.

[84] R. Herrera and N. Gonzalez, "The modeling and forecasting of extreme events in electricity spot markets," *Int. J. Forecast.*, vol. 30, no. 3, pp. 477–490, Jul. 2014.

[85] F. Paraschiv, D. Erni, and R. Pietsch, "The impact of renewable energies on eex day-ahead electricity prices," *Energy Policy*, vol. 73, pp. 196–210, 2014.

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

[87] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[88] M. Eichler, O. Grothe, H. Manner, and D. Türk, *(PDF) models for short-term forecasting of spike occurrences in australian electricity markets: A comparative study*, en, https://www.researchgate.net/profile/Tuerk-Dennis/publication/257947871_Models_for_short-term_forecasting_of_spike_occurrences_in_Australian_electricity_markets_A_comparative_study/links/0c960526778baa29c5000000/Models-for-short-term-forecasting-of-spike-occurrences-in-Australian-electricity-markets-A-comparative-study.pdf, Accessed: 2022-6-30, Oct. 2013.

[89] E. Fanone, A. Gamba, and M. Prokopczuk, "The case of negative day-ahead electricity prices," *Energy Econ.*, vol. 35, pp. 22–34, Jan. 2013.

[90] J. Janczura, S. Truck, R. Weron, and R. C. Wolff, "Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling," *Energy Econ.*, vol. 38, no. C, pp. 96–110, 2013.

[91] J. Mauritzen, "The silver lining of price spikes: How electricity price spikes can help overcome the energy efficiency gap," en, Aug. 2013.

[92] C. Wan, Z. Xu, Y. Wang, Z. Y. Dong, and K. P. Wong, "A hybrid approach for probabilistic forecasting of electricity price," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 463–470, 2013.

[93] T. M. Christensen, A. S. Hurn, and K. A. Lindsay, "Forecasting spikes in electricity prices," *Int. J. Forecast.*, vol. 28, no. 2, pp. 400–411, Apr. 2012.

[94] N. Amjady and F. Keynia, "A new prediction strategy for price spike forecasting of day-ahead electricity markets," *Appl. Soft Comput.*, vol. 11, no. 6, pp. 4246–4256, Sep. 2011.

[95] D. Singhal and K. Swarup, "Electricity price forecasting using artificial neural networks," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 3, pp. 550–555, 2011.

[96] S. Thomas, V. Ramiah, H. Mitchell, and R. Heaney, "Seasonal factors and outlier effects in rate of return on electricity spot prices in australia's national electricity market," *Applied Economics*, vol. 43, no. 3, pp. 355–369, 2011.

[97] N. Amjady and F. Keynia, “Electricity market price spike analysis by a hybrid data model and feature selection technique,” *Electric Power Systems Research*, vol. 80, no. 3, pp. 318–327, Mar. 2010.

[98] M. Nicolosi, “Wind power integration and power system flexibility–an empirical analysis of extreme events in germany under the new negative price regime,” *Energy Policy*, vol. 38, no. 11, pp. 7257–7268, Nov. 2010.

[99] S. K. Aggarwal, L. M. Saini, and A. Kumar, “Electricity price forecasting in deregulated markets: A review and evaluation,” *Int. J. Electr. Power Energy Syst.*, vol. 31, no. 1, pp. 13–22, Jan. 2009.

[100] C. L. Anderson and M. Davison, “A hybrid System-Econometric model for electricity spot prices: Considering spike sensitivity to forced outage distributions,” *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 927–937, Aug. 2008.

[101] R. Becker, S. Hurn, and V. Pavlov, “Modelling spikes in electricity prices,” en, *Econ. Rec.*, vol. 83, no. 263, pp. 371–382, Jan. 2008.

[102] J. Davis and M. Goadrich, “The relationship between precision-recall and roc curves,” in *Proceedings of the 23rd international conference on Machine learning*, 2006, pp. 233–240.

[103] T. D. Mount, Y. Ning, and X. Cai, “Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters,” *Energy Economics*, vol. 28, no. 1, pp. 62–80, 2006.

[104] W. Wu, J. Zhou, L. Mo, and C. Zhu, “Forecasting electricity market price spikes based on bayesian expert with support vector machines,” in *International Conference on Advanced Data Mining and Applications*, Springer, 2006, pp. 205–212.

[105] A. M. Gonzalez, A. M. S. Roque, and J. Garcia-Gonzalez, “Modeling and forecasting electricity prices with input/output hidden markov models,” *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 13–24, Feb. 2005.

[106] X. Lu, Z. Y. Dong, and X. Li, “Electricity market price spike forecast with data mining techniques,” *Electric Power Systems Research*, vol. 73, no. 1, pp. 19–29, Jan. 2005.

[107] D. H. C. Thai, “Analysing tacit collusion in oligopolistic electricity markets using a co-evolutionary approach,” en, Ph.D. dissertation, UNSW Sydney, 2005.

[108] J. Zhao, Z. Y. Dong, X. Li, and K. P. Wong, “A general method for electricity market price spike analysis,” in *IEEE Power Engineering Society General Meeting, 2005*, IEEE, 2005, pp. 286–293.

[109] F. E. Banks, “Economic theory and the failure of electricity deregulation in sweden,” *Energy and Environment*, vol. 15, no. 1, pp. 25–36, 2004.

[110] M. Burger, B. Klar, A. Müller, and G. Schindlmayr, “A spot market model for pricing derivatives in electricity markets,” *Quant. Finance*, vol. 4, no. 1, pp. 109–122, 2004.

[111] H. Yamin, S. Al-Agtash, and M. Shahidehpour, "Security-constrained optimal generation scheduling for gencos," *IEEE Transactions on power systems*, vol. 19, no. 3, pp. 1365–1372, 2004.

[112] R. Huisman and R. Mahieu, "Regime jumps in electricity prices," *Energy Econ.*, vol. 25, no. 5, pp. 425–434, 2003.

[113] S. Stoft, "The demand for operating reserves: Key to price spikes and investment," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 470–477, 2003.

[114] C.-K. Woo, D. Lloyd, and A. Tishler, "Electricity market reform failures: UK, norway, alberta and california," *Energy Policy*, vol. 31, no. 11, pp. 1103–1115, 2003.

[115] R. E. Oligney and M. J. Economides, "Natural gas: The excruciating transition," San Antonio, TX: Society of Petroleum Engineers, 2002.

[116] X. Guan, Y.-C. Ho, and D. L. Pepyne, "Gaming and price spikes in electric power markets," *IEEE Trans. Power Syst.*, vol. 16, no. 3, pp. 402–408, 2001.

[117] E. Mork, "Emergence of financial markets for electricity: A european perspective," *Energy Policy*, vol. 29, no. 1, pp. 7–15, 2001.

[118] C.-K. Woo, "What went wrong in california's electricity market?" *Energy*, vol. 26, no. 8, pp. 747–758, 2001.

[119] J. Yang, "A market monitoring system for the open electricity markets," in *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.01CH37262)*, vol. 1, Jul. 2001, 235–240 vol.1.

[120] D. S. Kirschen, G. Strbac, P. Cumperayot, and D. de Paiva Mendes, "Factoring the elasticity of demand in electricity prices," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 612–617, 2000.

[121] R. L. Winkler, "A decision-theoretic approach to interval estimation," *Journal of the American Statistical Association*, vol. 67, no. 337, pp. 187–191, 1972.

NHH

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no