Energy Reports 12 (2024) 3830–3842

Contents lists available at ScienceDirect

Energy Reports

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

Research paper

Quantifying and modeling price volatility in the Dutch intraday electricity market

Dane Birkeland [a,b], Tarek AlSkaif [a,*], Steven Duivenvoorden [b], Marvin Meeng [b], Joost M.E. Pennings [c,d]

[a] Information Technology Group, Wageningen University, Hollandseweg 1, 6706 KN, Wageningen, The Netherlands
[b] Authority for Consumers and Markets, Muzenstraat 41, 2511 WB, The Hague, The Netherlands
[c] Marketing and Consumer Behavior Group, Wageningen University, Hollandseweg 1, 6706 KN, Wageningen, The Netherlands
[d] Department of Marketing and Supply Chain Management, Maastricht University, Tongersestraat 53, 6211 LM, Maastricht, The Netherlands

ARTICLE INFO

Keywords:
Intraday electricity markets
Multivariate linear regression
Price volatility
Price realized volatility
Market fundamentals

ABSTRACT

This paper aims to provide a solid basis for the quantification and modeling of price volatility in the Dutch intraday electricity market. It analyzes price volatility through realized volatility, which is adapted from foundations in quadratic variation theory. Realized volatility is then estimated using differing multivariate linear regression and random forest regression models. We build these models around features pulled from quadratic variation theory, market fundamentals, liquidity, and information asymmetry. Furthermore, we assess the impact of features within the models using permutation feature importance and recursive feature elimination. The models leverage a multi-year dataset from EPEX SPOT containing completed trades of hourly products as well as other complementary data sources. The results of the paper include recommendations for future price volatility research within intraday electricity markets, mainly: (i) strive to utilize order book data to have a clearer idea of how prices settle and true bid–ask spreads, and (ii) increase model robustness by combining modeling efforts to assess DA, ID and balancing market impacts on price. This paper aims to benefit multiple stakeholders namely, academic researchers, industry participants, and European regulators, by providing a structured view on price volatility quantification and estimation for internationalized intraday electricity markets.

1. Introduction

Electricity trading takes place in multi-settlement markets, allowing to trade energy products with different temporal granularities. In the spot markets, intra-day (ID) electricity markets are gaining more importance as they allow for energy trading until shortly before the delivery period (Scharff and Amelin, 2016a; Le et al., 2019; Birkeland and AlSkaif, 2024). This can be profitable for market participants and beneficial for energy system operators to reduce imbalances from day-ahead planning, due to forecasting errors, since higher accuracy on (renewable) energy forecasts can be achieved closer to the physical delivery time. For instance, from 2010 to 2016, the volume of energy traded on the continuous German intraday (ID) market rose by about 400% (Martin and Otterson, 2018). Notably, one of the most pivotal recent developments in the ID market was the successful implementation of the Single IntraDay Coupling (SIDC) project. This ambitious initiative aimed to connect and harmonize intraday electricity trading across 23 countries (All NEMOs Committee, 2023). Starting in 2018, the SIDC connected all relevant geographic areas in 2023. Within this integrated framework, trading activity, *i.e.*, XBID, skyrocketed from 3.5 million trades in mid-2018 to a substantial 20.2 million trades in early 2023 (All NEMOs Committee, 2023).

The rise in ID trading activity within electricity markets is primarily driven by the growing presence of variable Renewable Energy Sources (vRES) (*e.g.*, wind and solar energy installations). This increase in vRES production can be partially attributed to the decreasing costs of essential technological components, including solar panels, as well as to subsidized investment capital (Uyterlinde et al., 2007). Furthermore, government mandates (*e.g.*, the European Directive 2009/28/EC which stated that, by 2020, 20% of the EU final energy consumption should be produced through vRES) increased vRES production across the EU (Council of the European Union European Parliament, 2009). Unlike fossil fuel-based generation, vRES energy generation is inherently intermittent, non-dispatchable and unpredictable due to its dependence on natural factors, such as solar irradiance and wind speed.

* Corresponding author.
E-mail address: tarek.alskaif@wur.nl (T. AlSkaif).

https://doi.org/10.1016/j.egyr.2024.09.031
Received 9 April 2024; Received in revised form 6 September 2024; Accepted 15 September 2024
Available online 1 October 2024

vRES shares have grown significantly across major EU states since the turn of the century, with Germany's vRES share growing from 8.79% to 19.45% between 2010 and 2021 (Our World in Data, 2023). The unpredictability of vRES generation exposes electricity market participants to risks of supply/demand shocks and significant short-term price movement (Tanaka et al., 2022). Within the context of price volatility, price movement and variation are crucial components. Price movement refers to changes in the price of a market product over time (*e.g.*, the H9–H10 ID hourly product), whereas price variation denotes the movement of prices on an ID market within a given day.

Price movement and variation in electricity markets have been explored through various modeling approaches in the literature. Pape et al. (2016) analyze price variation in the German day-ahead (DA) and ID markets by considering market fundamentals of marginal cost, demand for fossil-based operators and market supply states, employing a multiple linear regression approach (Pape et al., 2016). Other research focuses on forecasting prices, with volatility as a feature within a larger model. For example, Andrade et al. (2017) utilize probabilistic methods to forecast auction closing prices in the Spanish MIBEL market (Andrade et al., 2017), while Maciejowska et al. (2019) examine DA to ID price spreads per delivery hour using estimated vRES production and published price forecasts to capture agency estimated price movement (Maciejowska et al., 2019). Our study builds upon previous research by Chan et al. (2008), Ullrich et al. (2012), Haugom et al. (2011), and Frommel et al. (2014), focusing exclusively on the ID market and extracting volatility signals from realized measures such as realized variance (Chan et al., 2008; Haugom and Ullrich, 2012; Haugom, 2011; Frömmel et al., 2014). Qu et al. (2018) expand on the decomposition of realized volatility and inclusion of jump components in Australian energy markets by developing HAR-GARCH models that outperform benchmark models proposed by Chan et al. Qu et al. (2018). Similarly, our study addresses model combinations based on the benchmark HAR models and incorporates a broader collection of market features. Furthermore, our approach builds on the work of Karakatsani and Bunn (2010), who decomposed volatility into observable features in the UK spot ID market using GARCH methods, highlighting the importance of incorporating time-varying price responses for increased model accuracy (Karakatsani and Bunn, 2010).

Understanding ID price volatility is essential to comprehending the historical, present and future ID market status for participants and stakeholders alike. This underscores the need for quantification and modeling of price volatility. Researchers have proven quadratic variation theory to be a useful approach for analyzing the daily realized volatility (RV) of prices under such circumstances (Barndorff-Nielsen and Shepard, 2004, 2006). Key applications to electricity markets are in work both within and outside the European market (Chan et al., 2008; Haugom and Ullrich, 2012; Ciarreta et al., 2017). However, none of this research on price volatility – whether using quadratic variation or other methods – considers the Dutch ID electricity market, focusing instead on the UK, German, and Iberian markets. There are other approaches that adopt more traditional market returns for analyzing price volatility such as Lee and Mykland (2007), which is also based on the mathematical basis found in the work of Barndorff-Nielsen and Shepard (2004, 2006). The procedure denoted by Lee and Mykland (2008) is accessible and the transaction level data available for this study would make their procedure applicable on these data quality requirements (Lee and Mykland, 2007). However, the Lee and Mykland (2008) construction assumes that each product has no more than one price volatility jump per day since their analyses are focused on individual security pricing of a sample S&P 500 companies (Lee and Mykland, 2007). The Dutch ID market is actually made up of 24 distinct hourly products, where each could have more than 1 price volatility jump per day since they are all traded along differing timelines and address differing actor needs. Therefore, this paper uses the price jump decomposition techniques presented in Chan et al. (2008), Haugom and Ullrich (2012), Ciarreta et al. (2017) to measure the Dutch ID price volatility instead of methods centered around standard securities, such as those presented in Lee and Mykland (2007).

This paper explores the ID price movements in the Dutch bidding zone of EPEX SPOT in recent years and develops a regression model to better explain RV of continuous ID hourly delivery products in this zone. We achieve this by breaking down price volatility into separate components and showing the importance of time variation. To explain RV, we develop four distinct regression models, leveraging existing price forecasting and price premia literature. These four disparate models are supplemented with separate benchmarking models to add robustness to the results. The paper provides insights on price movements in the Dutch continuous ID electricity market. These insights have societal relevance for industry participants and regulators. Industry participants' understanding of volatility and price trends directly impact their day-to-day operations. For regulators, understanding volatility in their market of interest increases insights into extreme market behaviors and market design impacts. Additionally, research and deeper understanding of the Dutch electricity market has connections to the Title Transfer Facility (TTF) market, which is the largest European gas market and based in the Netherlands. Hence, this analysis can contribute to the development of economic and financial theories related to energy markets, price dynamics, and market behavior across these disparate markets.

The remainder of this paper is organized as follows: Section 2.1 presents the methods utilized to construct RV and targeted regression models, Section 3 describes the data sources which feed the regression models as well as the ID price time series, Section 4 illustrates the regression constructions and reports the results of the price volatility regression estimates, and Section 5 summarizes the price volatility study and discusses the implications of its results. Section 6 concludes the article and provides recommendations for future price analysis on the Dutch ID market.

2. Methods

2.1. Volatility measures

This study uses different methods to estimate price volatility. RV is a measure for the unobserved volatility of a high-frequency time series. Researchers traditionally estimate RV based on quadratic variation theory, which also enables total price variation to be decomposed nonparametrically into its continuous and jump components for equally spaced observations (Ciarreta et al., 2017). Next, this paper also seeks to understand whether RV can be explained using more common market metrics, such as trade volume or daily gas price. RV is constructed for each day using the following equation construction. Assume that prices are set T times for each day d. The RV for day d can then be estimated as:

$$RV_d = \sum_{t=1}^{T} r_t^2, \tag{1}$$

where $r_t = P_t - P_{t-1}$ refers to the difference between sequential sampled ID prices of adjacent delivery hours (*e.g.*, at time t and t-1). Each P_t represents one completed trade price for a unique year, month, day, and delivery hour combination from our dataset of the Dutch EPEX SPOT trades from 2020–2023. A specific P_t is sampled from the many prices sharing the same temporal combination using the pandas library in Python.

This process ensures that, for each unique time slot, a single completed trade price is chosen, making the selection unbiased and representative of each delivery hour within the dataset. Since this paper focuses on hourly ID products, T equals 24 for each day d, and the calculated RV captures the daily ID price volatility by aggregating price differences between individual hourly ID products across all hours within a given day.

2.1.1. Realized volatility and price jump detection

While RV captures the unobserved volatility, integrated volatility captures the continuous, predictable component of RV and can be estimated using a concept called bipower variation (BV) utilized within the work of Barndorff-Nielsen and Shepard (2006), where BV is given as:

$$BV_d = 1.57 \frac{T}{(T-1)} \sum_{t=2}^{T} |r_t \parallel r_{t-1}|. \tag{2}$$

In order to detect a statistically significant day, d, in terms of price movement, which will be called a jump day, Huang and Tauchen (2005) suggest the following statistic (Z_d) which can be evaluated at differing levels of significance, α.

$$Z_d = \sqrt{T} \frac{(RV_d - BV_d)/RV_d}{\sqrt{0.61 \max\left[1, TQ_d/BV_d^2\right]}}. \tag{3}$$

The denominator variable TQ_d[1] is described as:

$$TQ_d = 1.74 \frac{T^2}{(T-2)} \sum_{t=3}^{T} (|r_t||r_{t-1}||r_{t-2}|)^{4/3}. \tag{4}$$

Prior research proves the statistic Z_d converges to a normal distribution as $T \to \infty$ (Barndorff-Nielsen and Shepard, 2004; Ciarreta et al., 2017). Therefore, if the statistic Z_d exceeds the critical value $\phi_{(1-\alpha)}$, where α is the chosen significance level, day d is classified as a price jump day. The jump component of volatility (JV) at day d is obtained through:

$$JV_d = I_{Z>\phi_{(1-\alpha)}}(RV_d - BV_d), \tag{5}$$

where $I_{Z>\phi_{(1-\alpha)}}$ is 1 when day d is a jump day and 0 otherwise.

Once total RV and JV are estimated, the continuous component of the total variation, CV, is given by the difference between the two quantities, as:

$$CV_d = RV_d - JV_d. \tag{6}$$

CV_d is critical in understanding volatility days since it measures the price variation without taking the jump component.

2.1.2. Liquidity

According to Lybek and Sarr (2002) liquidity in any market is generally perceived as desirable because it increases allocation and information efficiency. Over the last six years, there has been a substantial increase in liquidity within the Dutch continuous ID market, as evidenced by the significant growth in both the number of completed trades and the quantity of traded volume measured in MWh. Data for the Dutch continuous ID market from the EPEX SPOT European power exchange shows that the trade count increased from around 57,000 trades in 2016 to around 2,400,000 in 2022, while the traded volume increased from around 1500 to 6500 GWh over the same time span (EPEX Spot, 2023). Earlier studies stress the importance of increasing liquidity to efficiently integrate wind energy into the larger DA and ID markets (Weber, 2010). We adopt commonly used indicators of liquidity (*e.g.*, the bid–ask spread, price high-to-low difference, price variance, trade volume in terms of unique trade count, and trade volume in terms of energy traded). We arrived at this specific list by following the work from Hagemann and Weber (2013). How each of these indicators is constructed from the price time series will be outlined in the following sub-sections.

Bid–ask spread. The bid–ask spread (BAS) is typically measured via the effective or quoted spread between orders on opposite sides of the order book. This calculation of the BAS would require order book data, which was not available for this research. Alternatively, the BAS can be calculated using the ID market transaction data. Hagemann et al. (2013) laid out an approach to approximate classical BAS using this type of market transaction data (Hagemann and Weber, 2013). First, we sort all trades in the yearly transaction list in a chronological row according to their execution timestamp. We then label the EPEX SPOT data with “buy” or “sell” indicators, which can be compared if they pertain to the same product, volume and delivery hour. For every pair of chronologically subsequent “buy” and “sell” trades, we calculate the difference between both trade prices and generate one BAS data point. Finally, we utilize all calculated BAS data points for a single product at a specific delivery hour to calculate the average BAS for that delivery hour. We then average these hourly BAS averages across all 24 delivery products yielding a daily BAS average for a day d. Note, however, that according to Hagemann and Weber (2013), this procedure tends to underestimate the magnitude of the actual BAS at a delivery hour. In addition, it may not correctly reflect the transactions corresponding to the order book “buy” and “sell” spread at the time of execution.

High-to-low difference. We define the high-to-low price difference variable as the difference between the highest and the lowest trade price for a delivery hour. The average high-to-low difference for each delivery hour t during day d can be calculated as:

$$\tau_d = \frac{\sum_{t=1}^{T} \max_{i=1,\ldots,N_t}(P_{i,t,d}) - \min_{i=1,\ldots,N_t}(P_{i,t,d})}{T}, \tag{7}$$

where N_t is the number of trades registered for the ID product for delivery hour t, $P_{i,t,d}$ is the price of trade i for delivery hour t in day d, and T is the number of times prices are set per day (*i.e.*,$T = 24$ in this work). The maximum and minimum are taken over the same hour t in day d, so that the maximum transaction price for a specific ID product is compared to the minimum transaction price for the same ID product. Thus, this indicator (τ_d) reflects the maximum price spread for a delivery product and captures some of the volatility present in RV once the high-to-low difference is averaged over all delivery products in a day.

Price variance. Furthermore, we calculate price variance as the average price variance of all trades for one delivery hour. This delivery hour variance is then averaged over all delivery hours, t, to obtain the average price variance for day d. The average price variance for a day d variable can be described as:

$$\sigma_d^2 = \frac{\sum_{t=1}^{T} \left(\frac{1}{N_t} \sum_{i=1}^{N_t} (P_{i,t,d} - \bar{x}_{t,d}) \right)^2}{T}, \tag{8}$$

where N_t is the number of trades registered for the delivery hour t product, $P_{i,t,d}$ is the price of trade i for delivery hour t in day d, and $\bar{x}_t$ is the mean price for the product deliverable at hour t. All of this calculation occurs across the transaction within a single day d. Thus, this indicator (σ_d^2) reflects the average price variance across all trades within each given ID product at time t and then made into a daily quantity by averaging over T. This is distinct, yet similar, to the key volatility measure RV since RV captures the volatility between sequential ID products before aggregating to a daily value, while σ_d^2 captures the variance within each ID product in reference to the daily mean ($\bar{x}_t$) before aggregating to a daily value.

Trade volumes. Finally, trading activity is measured via two variables: trade volume in terms of energy and trade volume in terms of unique trades. The energy trade volume is calculated as the total sum in MWh, while the number of unique trades is measured as the total number of trades during the delivery hours on day d.

[1] For a more extensive discussion on the rationale behind using TQ_d and jump detection in quadratic variation theory, the reader might refer to Huang and Tauchen (2005).

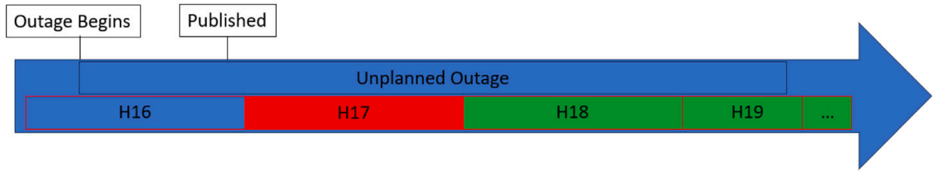


Fig. 1. Illustration of potential inside information using a 30-minute trading closure market design.

2.1.3. Market fundamentals

Since fossil generation is still a large component of energy generation in the Netherlands and in Europe, it is reasonable to assume that the prices of fossil fuels will have some impact on ID electricity market prices and their volatility. The prices of gas, coal and CO2-emissions are controlled for as they affect the marginal costs at which a gas or coal-fired power plant can produce electricity in the Netherlands (Mulder and Scholtens, 2013). Thus, gas and coal prices control for the raw materials needed for production and for whether fossil-fuel generation operates based on marginal cost economics. Meanwhile, the CO2-emission price controls for the penalty of emitting CO2 during production. Overall, it is reasonable to assume that higher fossil prices translate into higher electricity prices. In addition, it makes sense to assume that national level forecasting errors between DA and ID windows for load and vRES generation will impact ID price formation. These errors may contribute to volatility within the ID market, as market participants adapt to either upward or downward gaps in the load-generation balance (Valitov and Maier, 2020). All these factors are to be combined with the net physical transport of electricity taking place between the grids of major trading partners, the Netherlands and Belgium, which serves as a proxy measure for the number of XBID orders that require significant transmission capacity (Scharff and Amelin, 2016b).

2.1.4. Private vs. public capacity outage

Power plants often need to undergo scheduled maintenance which will leave gaps in the supply of electricity to the market. These supply dips could impact prices and must be communicated to the market in advance (ACER, 2023). However, unplanned shutdowns also occur at power plants. While these outages must be reported, there have been instances where the outage and reporting were not simultaneous and the unexpected outage of a significant generation source impacted market behavior before the relevant message was published. Additionally, outages and late messages can occur close to trading window closures. These timing dynamics are illustrated in Fig. 1 for a market design whereby trade closes 30 min prior to delivery and the publication of an unplanned outage does not reach stakeholders until 10 min after gate closure. In this figure, the “Outage Begins” tag at 15:20, indicates the beginning of the unplanned outage, followed by “H17 Close”, which indicates the closure of trade for the H17 delivery window. This left parties aware of the unplanned outage with 10 min to act on insider information regarding the H17 delivery window. This insider information was not publicized until 15:40, indicated by the “Published” tag. From this point on, there is no longer any insider information regarding the unplanned outage that might impact the trading of energy slated for delivery in the H18 window and beyond.

2.2. Regression methods

We implement regression methods in this study to estimate the daily realized ID price volatility RV (see Section 2.1). We adopt two common and effective regression models: multivariate linear regression (MLR) and random forest (RF) regression. MLR is suitable for simpler, linear relationships, while RF regression is more versatile, capable of handling nonlinear relationships and robust to outliers and high-dimensional data. The relationship between most of the variables in the previously discussed sections and RV is unknown; both methods can be used to take advantage of their respective strengths and give more insight into modeling price volatility.

2.2.1. Multivariate linear regression

MLR is a simple yet effective regression model that is widely adopted to analyze ID electricity market prices (Chan et al., 2008; Ciarreta et al., 2017; Pape et al., 2016). Based on training data, the MLR model uses a loss function to determine the coefficients that explain a linear relation between the predictor variables and the target variable. The MLR model estimates the target variable RV as:

$$\mathbf{RV}^{MLR} = \beta_0 + \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \cdots + \cdots + \beta_k \mathbf{x}_k, \tag{9}$$

where the βs are the regression coefficients and k is the number of predictor variables.

2.2.2. Random forests regression

RF regression is a tree-based regression model that has proven its value for time-series forecasting and regression applications (Grinsztajn et al., 2022; Jain et al., 2022; Visser et al., 2022). It will also be used here to estimate RV and compare its performance with MLR. RF is an ensemble based model that consists of a number of trees, each made up of n layers and 2^n decision nodes, with $n = 0$ at the first layer. The decision trees are created independently and are built by considering bootstrap samples of the training dataset. Next, for each tree a random subset of the predictor variables is considered to construct the decision nodes by optimizing on a loss function, *e.g.*, least squares (Breiman, 2001). The output of an RF model is equal to the conditional mean of all constructed trees.

2.3. Feature importance and selection

Several external factors can influence ID market prices and their volatility. In this study, we establish feature importance and selection methods using permutation feature importance (PFI), backward feature selection, and recursive feature elimination (RFE). The PFI is performed separately for the MLR and RF models, for the purpose of ID market price volatility analysis (scikit-learn developers, 2023c). The selection is based on the (lowest) obtained R^2 and mean squared error (MSE) scores, when a single feature value is randomly shuffled. This procedure breaks the relationship between a feature and the target variable, thus the drop in the model score is indicative of how much the model depends on that feature. PFI is used for this scenario because the technique is model-agnostic and can be applied to both the MLR and RF regressor. Furthermore, it provides a global view of feature importance since the same features are present across both models.

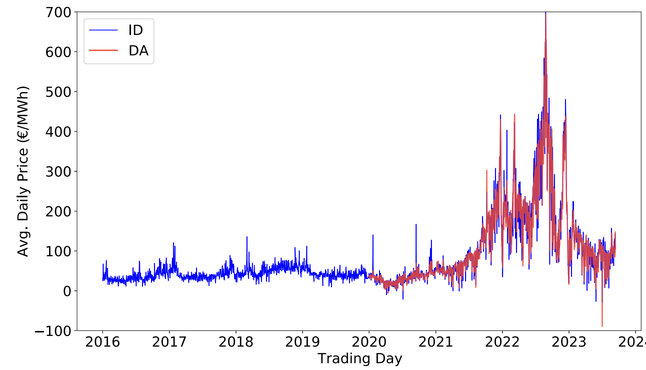
Backward feature selection involves an iterative procedure wherein the model is initially trained with each individual feature present. The feature, or predictor variable, that contributes least to the best-performing model is removed from the initial selected feature list. This value is established in this study using change in R^2. All R^2 values are calculated on the training set. Subsequently, the model is trained with the remaining features. The second feature that reduces the model's performance is then removed from the selected features list. This process is reiterated until all feature sets have been tested into the model and an optimal set of features is found. The specific backward selection method used for feature set creation is RFE utilizing the `sklearn:RecursiveFeatureElimination` package (scikit-learn developers, 2023b). It recursively fits a model and removes the least important features at each iteration based on the training set using R^2. RFE starts with a given number of features and repeatedly eliminates the least important one until the desired number of features remains. This technique can be done iteratively over the feature space to optimal features within a specific subset of the total features as well as finding the optimal amount of features as described above. This technique is distinct from PFI since it is model-specific.

3. Data and simulation setup

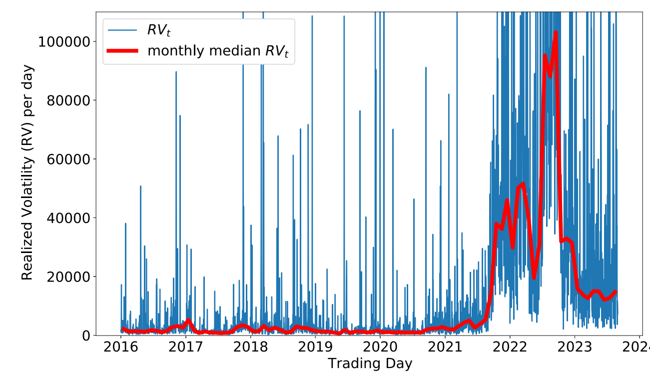
3.1. Price time series

This paper focuses on the Dutch continuous ID market price between 2020 and 2023. Specifically we focus on hourly ID products over this time frame. We obtained the DA market clearing prices and continuous ID trades data for the Dutch market from the EPEX SPOT European power exchange (EPEX Spot, 2023). Fig. 2 shows why the 2020–2023 period has become the focus for a deep analysis of the price movements. We can observe from Fig. 2(a) that the average daily prices (*i.e.*, the average price over all hours and transactions for a given day) are unevenly distributed within a single day and across days. Prices exhibited minor volatility, but had very little upward or downward trending until 2020. Fig. 2(b) displays the monthly median RV from 2016 to 2024, with significant spikes in volatility observed starting in late 2021. Year 2020 saw both ID (in blue) and DA (in red) (on Fig. 2(a)) prices fall as industrial demand and overall economic output dropped during the COVID-19 pandemic. After a period of lockdowns in the Netherlands and globally, prices began to surge upward as the Dutch economy sought to rebound from the period of economic contraction (CBS, 2023). This trend remained steady in the first half of 2021 across the ID and DA markets until the full reopening of European economies and the scramble to regain lost production began to increase RV in late 2021. RV reached its highest level in the timeframe once the European gas shortage caused by the Russia–Ukraine war sent prices into full volatility in mid 2022. Year 2022 also experienced record breaking high summer temperatures which contributed to pushing prices to upward extremes. An example of the impact the extreme weather had in 2022 was on water levels in Germany and France which dropped extremely low, inhibiting nuclear reactor cooling in France and coal shipping in Germany (Shiryaevskaya et al., 2022; Kollewe, 2022).

Quantifying price volatility is essential for all market participants to assess risk, make informed decisions, and develop strategies to support energy generation and investment in economically feasible ways. Accurate measurements and representations of volatility provides valuable insights into the behavior of ID electricity market prices and the broader market. Due to daily variability in vRES generation and energy consumption, prices for different ID delivery products (DPs), which correspond numerically to the hour of delivery, can exhibit different behavior within a day and across seasons. This can be visually represented through candlestick charts, whereby each candlestick shows the opening, closing, minimum and maximum prices for a given day. The opening price is the price of the first trade on a given day, and the closing price is the price of the last trade on the same day. If the opening price is lower than the closing price on a given day, the body of the candle is red-colored; in the opposite case, it is green. The lowest end of the wick shows the minimum price in the ID market on a specific day, while the highest shows the maximum price. The motivation for this study is the 2020–2022 period where price and volatility have increased dramatically. Significant variability is observed in the trade prices for specific DPs (14, 5), as indicated in Figs. 3. The first subfigure shows the candlestick chart for DP = 14 – *i.e.*, during peak demand hours – for the trade period from 1 July 2022 to 30 September 2022. The maximum price for DP = 14 was reached on 25 August 2022, when the price reached over 600 €/MWh. On most days between 1 July 2022 and 30 September 2022, the ID prices remained around 250 €/MWh for DP = 14. However, this period also saw many days with price spikes beyond 400 €/MWh, as well as some with negative prices,



(a) ID and DA average daily prices.



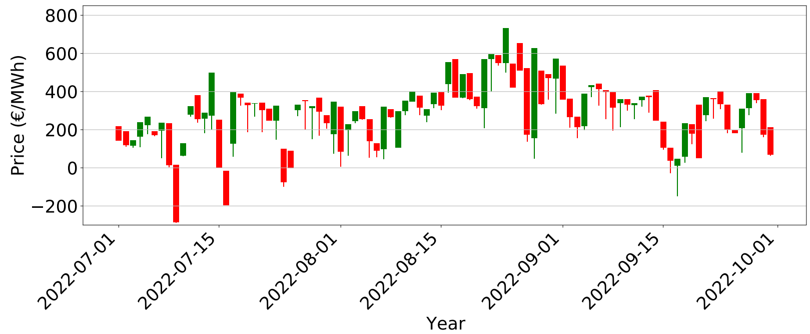
(b) ID prices' realized volatility.

Fig. 2. ID average daily prices and their realized volatility over years.

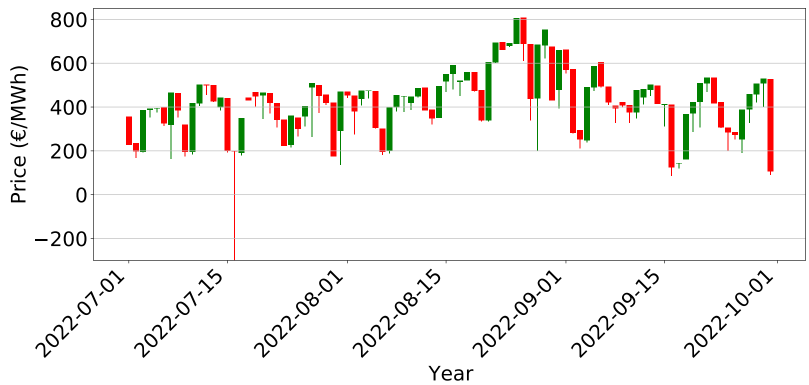
reaching a period low of less than −200 €/MWh. In this same period, DP = 5 – *i.e.*, the off-peak demand hours – saw a maximum price of 800 €/MWh, reached on two sequential days (25 and 26 of August 2022). In this period, DP = 5 saw consistently high prices compared to DP = 14, with highs of over 400 €/MWh on most days. Additionally compared to DP = 14, DP = 5 saw only one day with a negative price; 16 July 2022.

This level of volatility also occurred in the winter months of 2022 prior to the beginning of the Russia–Ukraine war (on 24 February 2022), as shown in Figs. 4. The first figure shows the candlestick chart for DP = 14, during the trading period from 1 January 2022 to 31 March 2022. The maximum price for DP = 14 reached 400 €/MWh on 16 March 2022. On most days, however, the ID prices for DP = 14 remained around 250 €/MWh. However, this period also saw many days with price spikes beyond 300 €/MWh, as well as some with negative prices, reaching a period low of −300 €/MWh on 20 March 2022. In this same period, DP = 5 – representing the off-peak demand hours – saw a maximum price of over 400 €/MWh on three sequential days (8, 9, and 10 March 2022). This period saw consistently high prices for DP = 5, averaging around 250 €/MWh, (*i.e.*, similar to those for DP = 14).

We provide summary statistics of the hourly ID market price time series across differing time aggregations in Table 1. Table 1 shows that mean prices are higher in autumn and summer in the Dutch ID market. Also, day-of-the-week seasonality is observed, in that reduced economic activity during weekends – due to fewer hours worked nationally – results in lower prices. According to the Jarque–Bera test, the price distribution of probability is not normal. Multiple literature record the same type of statistical patterns in electricity prices (Qu et al., 2018; Haugom, 2011; Chanatásig-Niza et al., 2022; Frömmel et al., 2014).

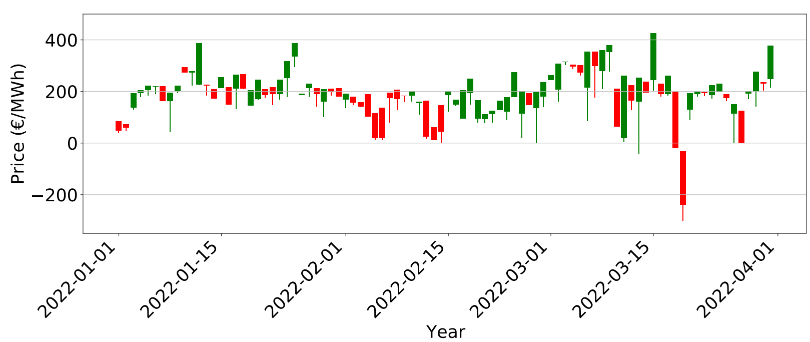


(a) Quarter 3; DP=14.

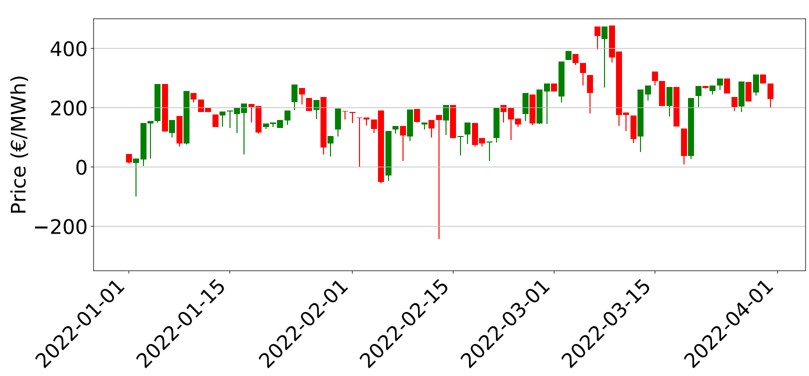


(b) Quarter 3; DP=5.

Fig. 3. Candlestick charts of ID market prices for Q3 2022.



(a) Quarter 1; DP=14.



(b) Quarter 1; DP=5.

Fig. 4. Candlestick charts of ID market prices for Q1 2022.

Table 1
Summary statistics for daily aggregated ID prices (€/MWh) per quarter and day of week (2020–2022).

Time	Mean	Std. Dev.	Minimum	Maximum	Skewness	Kurtosis	Jarque–Bera
Q1	117.84	104.92	−400.00	850.12	1.26	1.55	$6.84 \cdot 10^5$
Q2	126.24	100.36	−349.99	1,993.91	0.60	0.09	$1.51 \cdot 10^5$
Q3	225.98	189.63	−662.00	1,994.00	0.79	−0.27	$3.16 \cdot 10^5$
Q4	178.26	125.98	−997.90	1,994.00	0.86	0.48	$4.26 \cdot 10^5$
Monday	178.22	149.23	−430.61	950.00	1.16	1.14	$4.40 \cdot 10^5$
Tuesday	182.23	149.40	−400.00	1,145.64	1.16	1.05	$4.31 \cdot 10^5$
Wednesday	183.62	151.94	−997.90	1,088.59	1.13	0.99	$3.94 \cdot 10^5$
Thursday	180.77	151.47	−349.39	1,199.99	1.13	0.96	$4.07 \cdot 10^5$
Friday	173.71	146.98	−346.57	1,994.00	1.24	1.74	$5.94 \cdot 10^5$
Saturday	140.45	127.15	−400.00	994.99	1.34	2.56	$7.54 \cdot 10^5$
Sunday	128.47	120.27	−662.90	983.88	1.23	2.04	$5.55 \cdot 10^5$

Table 2
Missing hours of aggregated ID price data per year of focus.

Year	Empty Hours	% Missing
2020	0	0.00
2021	1	0.01
2022	2	0.02
2023	8	0.14

Table 3
Identified jump days per year and quarter (2020–2023).

Time	Total	Q1	Q2	Q3	Q4	% of Days
2020	108	27	35	27	19	29.5
2021	79	29	19	20	11	21.6
2022	23	5	7	3	8	6.3
2023 (Jan–Aug)	4	1	2	1	N/A	1.7

Finally, we observe seasonal behavior of the various delivery products as shown in Figs 4 and 3 and the distinctly different patterns shown for DP = 5 and DP = 14. Following a standard procedure that was also used for spot prices seasonality in the Pennsylvania–New Jersey–Maryland (PJM) market in Haugom and Ullrich (2012), and reaffirmed with energy market experts at Autoriteit Consument en Markt (ACM, *i.e.*, the Dutch market regulator), the medians of the ID price differences over a day d are subtracted from the price differences for each month of the year, day of the week, and hour of the day. This new metric is deployed to account for seasonality when performing the estimation of RV using the differing regression models. The resulting adjusted ID price differences are calculated as:

$$r_t^* = r_t - median(r_{m,d,t}), \tag{10}$$

where $r_{m,d,t}$ is the median of the month-over-month (m), day-of-the-week (d), and hour-over-hour price differences (t). We replace the ID price differences with the adjusted difference, r^*, in all RV calculations described in Section 2 Methods. Using a seasonality robust version of RV_d and using the jump statistic described in Section 2.1.1 with $\alpha = 0.001$, results in Table 3 counting the number of jump days per year and quarter. By pre-processing the data set composition we found the data set to contain missing values. We report these missing values in Table 2. However, the overall missing values represent less than 0.1% of the total number of hours in the assessment period. These data gaps mean that there have been delivery hours when no electricity was traded on the Dutch ID market. The data is missing for one assumed reason; liquidity efficiencies. A liquidity inefficiency is when there is no trade for each delivery hour in the continuous Dutch ID market across all years. However, market liquidity, inferred via trade volume, has been increasing over time. This explains the decreasing percentage of missing aggregated hourly prices per year, excluding year 2019. However, to study the explanatory power of the different models regarding ID prices and their volatility, it is essential to transform the ID trade prices into a complete price time series. We estimate and fill-in the time series for analysis via the padding method, as proposed by Shinde et al. (2021). This padding method entails filling the missing data points with either the most recent or the next data point in the timeline.

As established by Barndorff-Nielsen and Shepard (2006) and reaffirmed by Ciarreta et al. (2017) for the case of EPEX ID prices with jumps present, the decomposition approach of RV is upward biased in contexts with finite samples BV_d. As such, the JV_d component of RV is underestimated. However, in the context of near-zero returns on variability BV_d is affected, thus reducing its value, to the detection of more jumps. We affirm the presence of this phenomenon for the Dutch ID market and illustrate its impact in Table 3. The table shows that more jumps were detected in 2020 and 2021 when prices were more closely bunched but jumps decreased significantly in 2022 and 2023 when variability became more extreme. Finally, we divide the price time series into training and test sets. The constructed time series between 2020 and 2022 is the training set, while the partial 2023 time series is reserved as the test set to evaluate the MLR models.

3.2. ENTSO-E platform

The work in Hagemann and Weber (2013) discusses in depth the determinants for German ID price formation. Among these determinants are load and vRES forecast errors as well as net cross-border physical flows between major electricity trading partners. Similarly, Kiesel and Paraschiv (2017) uses vRES forecasts in their econometric analysis to explain 15-minute German ID electricity pricing. Furthermore, Hu et al. (2021) and Karanfil and Li (2017) also include cross-border ID exports to understand price deviation within the Swedish power market and were able to show their significant impacts on price premia in markets with large exposure to wind generation. The ideas behind these explanatory variables is that a trade on the ID market is influenced by actual demand/supply differing from the DA forecasted demand/supply. Thus, market participants use the ID market to balance these differing values. To assess their importance in explaining price volatility, we assemble data from ENTSO-E, by taking the difference between the expected and actual load as well as expected and actual vRES generation. We calculate cross-border net physical exports as the difference between electricity sent to Belgium and received from Belgium over the ID period, while load and vRES forecast errors are calculated as the aggregation of quarter-hourly differences between the day-ahead forecasts of load or vRES generation and the actual load or vRES generation for that same quarter-hour. It is important to note that the numbers below are aggregated at a daily level. This explains how a seemingly large −3.5GWh forecast error is reported since it is only about 1% of the total electricity GWh load for a single day in the Netherlands.

Table 4 shows the summary statistics for these variables.

Table 4
Summary statistics for daily aggregated ENTSO-E variables (MWh) (2020–2022).

Variable	Mean	Std. Dev.	Minimum	Maximum	Skewness	Kurtosis	Jarque–Bera
Net Export	8,884.26	22,882.17	−53,640.00	69,078.00	−0.06	−0.56	14.77
Load Forecast	−3,457.07	24,739.83	−92,161.00	60,303.50	−0.76	0.76	131.54
vRES Forecast	27.61	133.15	−14,508.50	24,290.25	7.28	172.73	$1.37 \cdot 10^6$

Table 5
Summary statistics for daily aggregated fossil variables (€/MWh) (2020–2022).

Variable	Mean	Std. Dev.	Minimum	Maximum	Skewness	Kurtosis	Jarque–Bera
Gas Price	32.22	60.60	2.70	318.88	1.31	1.48	415.64
Coal Price	152.99	113.11	38.63	473.33	0.77	−0.81	137.78
CO_2 Price	52.95	24.58	15.20	97.59	0.09	−1.49	102.59

Table 6
Summary statistics for included UMM (MW) (2020–2023).

Reporting time	Count	Mean	Std. Dev.	Minimum	Maximum
Peak	3,344	322.20	256.38	100.00	1,304.00
Off-peak	3,765	332.83	254.04	100.00	1,304.00
All	7,109	327.83	255.18	100.00	1,304.00

3.3. Fossil pricing

We obtained all fossil data used in this study from the London Energy Brokers Association (LEBA). For the gas prices, we utilize the Amsterdam based TTF DA gas price for week and weekend days. For coal prices, we use the one-month delivery coal prices on the Rotterdam coal index. For carbon prices, we use the EUA spot data. For coal and carbon, trade only takes place on weekdays. Therefore, we set the price for the weekend days as equal to the price on the preceding Friday. Table 5 shows the summary statistics for the fossil-based variables.

3.4. Urgent market messages

The EEX transparency platform is an inside information platform approved by The EU Agency for the Cooperation of Energy Regulators (ACER) which is commonly used by participants in the Dutch market. From the EEX platform, we received a dataset containing all planned and unplanned outages of power plants within the Dutch bidding zone published on the EEX platform during the years of interest. These messages are called Urgent Market Messages (UMM). However, from the entire dataset, we exclude messages about outages of less than 100 MW and less than one hour in duration as a simplifying assumption and for completeness, since outages of 100 MW or more must always be reported due to their impact, for lower volumes this depends on market conditions (ACER, 2019). Whether the MWs reported in an UMM are classified as planned or unplanned outages depends on the time lag between the start of the outage and its reporting, following the procedure established by Valitov and Maier (2020). If an UMM is issued before or simultaneously with the outage, we classify it as a planned outage. Conversely, we classify an UMM as an unplanned outage if the publication timestamp is from after the beginning of the outage.

Table 6 shows the summary statistics for the remaining UMMs from 2020 to 2023. The mean capacity loss caused by an outage is about 327.18 MW. The majority of UMM messages between 2020 and 2023 were published during off-peak hours (*i.e.*, 1–8 and 21–24).

Finally, we assign the MWs within the UMM to two distinct explanatory variables, described in Section 2.1.4 Private vs. public capacity outage: Private information and Public information. We classify private information as the sum of capacity loss in UMMs from the beginning of the outage until its publication on the relevant platform. Consequently, the variable Public Information aggregates total capacity loss in UMMs from the publication timestamp until the expected end of the outage. Table 7 shows the summary statistics for the *Private* and *Public* variables using the method described in this section.

4. Results

4.1. Regression models constructions

Once RV is separated into jump and continuous components using the procedure described prior in Section 2.1.1 Realized volatility and price jump detection, the HAR-CV-JV model, introduced by Chan et al. (2008) and shown to be the best performing in explaining RV_d by Ciarreta et al. (2017). We also consider a log transformation of RV_d and its decomposition as:

$$\begin{aligned} logRV_d^h &= \text{constant} + \beta_1 logCV_{d-1} + \beta_2 logCV_{w,d-1} + \beta_3 logCV_{m,d-1} \\ &\quad + \beta_4 logJV_{d-1} + \beta_5 logJV_{w,d-1} + \beta_6 logJV_{m,d-1}, \end{aligned} \tag{11}$$

where $logCV_{d-1}$ and $logJV_{d-1}$ are the CV and JV for the previous day, respectively. $CV_{w,d-1}$ and $JV_{w,d-1}$ are the average CV and JV, respectively, over the previous week and $CV_{m,d-1}$ and $JV_{m,d-1}$ are the average CV and JV, respectively, over the previous month. This model construction was the best performing in the work by Ciarreta et al. (2017) when used to forecast RV at a quarter-hourly DP level. When there is a no jump day such that $JV = 0$, which would result in $\log(0)$, then a 0 value is assigned. We estimate the HAR-CV-JV model using ordinary least squares (OLS), where heteroskedasticity and autocorrelation-corrected (HAR) consistent standard errors are utilized in coefficient estimation. We apply the log transformation for RV_d because the concavity of this nonlinear form means that high values in the time series decrease more than low values, making the time series smoother and the jumps less pronounced.

Presenting a distinctly different view on describing price volatility is the market fundamental model utilizing the ENTSO-E data and fossil fuel data described prior in Section 2.1.3 Market fundamentals. We regress the identified variables against the $log RV_d$ as shown in Eq. (12).

$$\begin{aligned} logRV_d^f &= \text{constant} + \beta_1 \text{NetExports}_d + \beta_2 \text{LoadForecastErr}_d + \\ &\beta_3 \text{vRESForecastErr}_d + \beta_4 \text{GasPrice}_d + \beta_5 \text{CoalPrice}_d + \beta_6 \text{CO2Price}_d \end{aligned} \tag{12}$$

We then extend this model to include the *Private* and *Public* variables constructed in Section 3.4 Urgent market messages. This extension of the market fundamentals model results in Eq. (13).

$$\begin{aligned} logRV_d^u &= \text{constant} + \beta_1 \text{NetExports}_d + \beta_2 \text{LoadForecastErr}_d + \\ &\beta_3 \text{vRESForecastErr}_d + \beta_4 \text{GasPrice}_d + \beta_5 \text{CoalPrice}_d + \\ &\beta_6 \text{CO2Price}_d + \beta_7 \text{Private}_d + \beta_8 \text{Public}_d \end{aligned} \tag{13}$$

We also consider a final baseline model using the liquidity variables discussed in Section 2.1.2 Liquidity. When used to describe $log RV_d$ with a linear function, these variables result in Eq. (14).

$$\begin{aligned} logRV_d^l &= \text{constant} + \beta_1 \text{BAS}_d + \beta_2 \text{HighLowDiff}_d + \\ &\beta_3 \text{PriceVar}_d + \beta_4 \text{EnergyVolume}_d + \beta_5 \text{TradeCount}_d, \end{aligned} \tag{14}$$

where *BAS*, *HighLowDiff* and *PriceVar* are the averages of their respective hourly variables over day *d*.

Table 7
Summary statistics for Private and Public variables (MW) (2020–2022).

Variable	Mean	Std. Dev.	Minimum	Maximum	Skewness	Kurtosis	Jarque–Bera
Private	171,577.06	190,936.49	0.00	1,030,424.00	2.63	7.54	3,863.45
Public	51,254.10	36,580.34	0.00	214,030.00	1.37	2.40	607.65

Table 8
Parameter estimates for HAR-CV-JV model $logRV_d^h$.

Parameter	Coefficient	Std. Error	*P*-value
Intercept	0.748	0.297	0.012
$logCV_{d-1}$	0.419	0.034	0.000
$logCV_{w,d-1}$	0.269	0.045	0.000
$logCV_{m,d-1}$	0.219	0.035	0.000
$logJV_{d-1}$	0.086	0.013	0.000
$logJV_{w,d-1}$	−0.018	0.010	0.090
$logJV_{m,d-1}$	−0.002	0.013	0.887
Adj. R^2		58.7%	

Table 9
Parameter estimates for Market Fundamentals model on $logRV_d^f$.

Parameter	Coefficient	Std. Error	P-value
Intercept	7.487	0.099	0.000
NetExports$_d$	$1.058 \cdot 10^{-5}$	$1.640 \cdot 10^{-6}$	0.000
LoadForecastErr$_d$	$1.488 \cdot 10^{-6}$	$1.740 \cdot 10^{-6}$	0.393
vRESForecastErr$_d$	$-6.531 \cdot 10^{-6}$	$1.240 \cdot 10^{-6}$	0.599
GasPrice$_d$	0.013	0.001	0.000
CoalPrice$_d$	0.002	0.001	0.002
CO2Price$_d$	0.007	0.003	0.017
Adj. R^2		53.8%	

As in the case of the HAR-CV-JV model, we estimate these additional base models using OLS, where HAR consistent standard errors are utilized in coefficient estimation.

In order to increase result robustness, we compare the proposed models above against a composite regression model made up of all the described variables. This model will undergo RFE and feature importance variable reduction techniques whose results are reported in Section 4.3 Feature importance and benchmark models. Furthermore, we construct a hyper-parameter tuned RF regression model using the whole variable space to provide an unsupervised benchmarking for the targeted regression models as well as another view on feature importance. We introduce the RF regressor in Section 2.2.2 Random forests regression. Each of the decision trees is trained on a random subset of the RV time series and tuned number of features, such randomness helps to reduce overfitting. The RF regression is generated using the `sklearn:RandomForestRegessor` package (scikit-learn developers, 2023a). We discuss the selection of variables for the RF model in further detail in Section 4.3 Feature importance and benchmark models.

4.2. Regression outputs

Table 8 reports the parameter estimates of the HAR-CV-JV model per Eq. (11).

The continuous component of volatility, CV appears to have a significant positive impact on RV. This impact is common across all three time lag periods showing that price volatility exhibits an impact that exceeds a single day. This longer time horizon of price volatility is evidenced by the positive significant impact of JV only for the 1-day lag variable at 5%, but not for the weekly or monthly lag variables. It is important to note that all of these significant variables impact $logRV_d^h$ positively which is in line with the upward pricing trend over the 2020–2022 period, as shown in Fig. 2.

Table 9 reports the parameter estimates of the market fundamentals model as in Eq. (12).

Table 10
Parameter estimates for Market Fundamentals with UMM model on $logRV_d^u$.

Parameter	Coefficient	Std. Error	P-value
Intercept	7.432	0.092	0.000
NetExports$_d$	$7.839 \cdot 10^{-6}$	$1.500 \cdot 10^{-6}$	0.000
LoadForecastErr$_d$	$1.835 \cdot 10^{-6}$	$1.350 \cdot 10^{-6}$	0.174
vRESForecastErr$_d$	$-1.600 \cdot 10^{-5}$	$2.450 \cdot 10^{-5}$	0.514
GasPrice$_d$	0.013	0.001	0.000
CoalPrice$_d$	0.002	0.001	0.008
CO2Price$_d$	0.011	0.003	0.000
Private$_d$	$-1.285 \cdot 10^{-6}$	$2.060 \cdot 10^{-7}$	0.000
Public$_d$	$3.697 \cdot 10^{-6}$	$1.010 \cdot 10^{-6}$	0.000
Adj. R^2		55.5%	

Net exports have a significant positive impact on price volatility. Thus, as the Netherlands exports more electricity to Belgium than it receives, $logRV_d^f$ increases by 1.058 for every 1 TW. This significance does not carry over to load forecast errors and vRES forecast errors. This may be due to the effectiveness of the forecasting by grid operators. However, as expected, both gas and coal prices score significantly positive on $logRV_d^f$ since these are still two of the largest sources of electricity generation in Europe. This is particularly true for gas generation which can be turned on and off quickly, so it can adapt to ID changes in grid balance. The $GasPrice$ value will be a large factor in determining ID market participation for gas generation actors. Since electricity producers using gas may not participate in the ID market if the spread between their operating cost driven heavily by gas price and the ID price is not favorable due to volatility impacting their risk assessment or simply ID prices being too low. This is for instance in opposition to operators of other power generators (*i.e.*, coal and nuclear) who have less flexibility in switching on/off and thus less impact on the ID price volatility.

Table 10 reports the parameter estimates of the market fundamentals including the UMM model per Eq. (13).

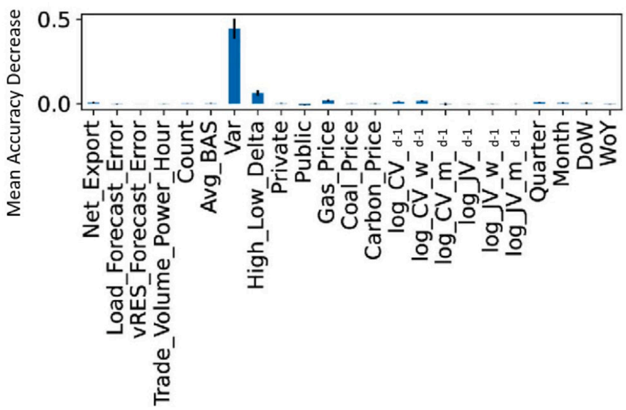
The addition of the *Private* and *Public* variables reveals opposite significant impacts on approximating $logRV_d^u$. Interestingly, increases in private capacity have a negative impact on $logRV_d^u$, whereas public capacity has a positive impact on $logRV_d^u$. We attribute the reason for this inverse impact to the fact that public outages have a more extensive influence on numerous market participants, consequently leading to increased price volatility. Private information capacity possibly impacts a smaller set of market participants who can react with clear knowledge and thus reduce price volatility within a day. Additionally, the non-zero coefficient for the *Private* variable could potentially indicate that market participants use the lapses in outage reporting as private information. In terms of overall model construction, these UMM-based variables are important additions to the model as the model's Adj. R^2 increases over the base market fundamentals model.

Table 11 reports the parameter estimates of the liquidity measure model per Eq. (14).

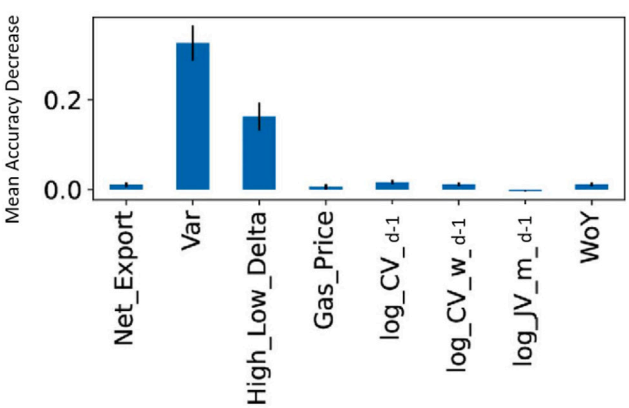
While the liquidity model has the highest Adj. R^2 of the set of models discussed in this section, it has only one feature that is significant at the 5% level (other than the intercept). This model may have the highest Adj. R^2 due to its good fit to RV time series since these features, especially *PriceVar* and *HighLowDiff*, are constructed from ID price differences and so is RV. The lone significant feature is *HighLowDiff* which shows that an average price gap of €1/MW between the high and low prices for DPs within a day leads to a 0.03 increase in $logRV_d^l$. The most basic liquidity measures, *BAS*, *EnergyVolume* and *TradeCount*, are distinctly nonsignificant possibly because *BAS* is an approximated

Table 11
Parameter estimates for Liquidity model on $logRV_d^l$.

Parameter	Coefficient	Std. Error	P-value
Intercept	7.521	0.100	0.000
BAS_d	−0.019	0.019	0.310
$HighLowDiff_d$	0.030	0.002	0.000
$PriceVar_d$	$-3.000 \cdot 10^{-4}$	0.000	0.061
$EnergyVolume_d$	$-7.485 \cdot 10^{-6}$	$6.820 \cdot 10^{-6}$	0.272
$TradeCount_d$	$-2.155 \cdot 10^{-5}$	$1.690 \cdot 10^{-5}$	0.201
Adj. R^2		63.9%	



(a) All Features by Mean Change in R^2.



(b) Final Features above 0.005 Mean Change in R^2.

Fig. 5. PFI results for the RF regressor.

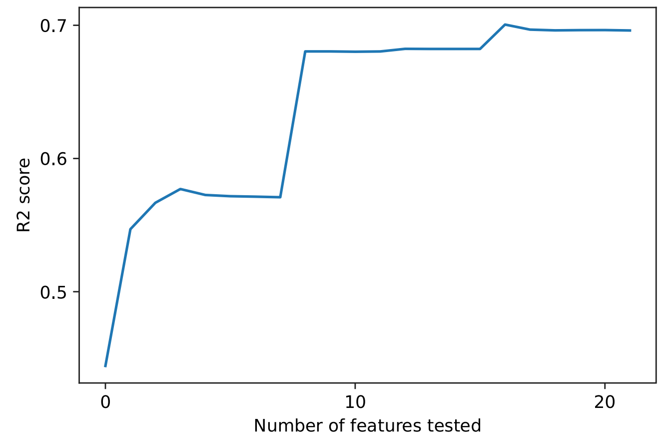


Fig. 6. Optimal feature count for full HAR OLS Model using iterative RFE process.

Table 12
Optimal feature set for the fulsome mixed HAR OLS model: parameter estimates.

Parameter	Coefficient	Std. Error	P-value
$TradeCount_d$	$-5.409 \cdot 10^{-5}$	$8.670 \cdot 10^{-6}$	0.000
BAS_d	−0.024	0.017	0.158
$PriceVar_d$	$-3.000 \cdot 10^{-4}$	0.000	0.005
$HighLowDiff_d$	0.013	0.002	0.000
$GasPrice_d$	−0.007	0.001	0.000
$CoalPrice_d$	0.001	0.001	0.056
$CO2Price_d$	0.007	0.002	0.004
$logCV_{d-1}$	0.271	0.031	0.000
$logCV_{w,d-1}$	0.218	0.041	0.000
$logCV_{m,d-1}$	0.322	0.032	0.000
$logJV_{d-1}$	0.072	0.011	0.000
$logJV_{w,d-1}$	0.011	0.009	0.223
$logJV_{m,d-1}$	0.074	0.012	0.000
Quarter	−0.119	0.114	0.295
Month	0.178	0.057	0.002
DayofWeek	0.020	0.015	0.171
WeekofYear	−0.020	0.010	0.047
Adj. R^2		99.0%	

measure using non-orderbook data, as described in Section 2.1.2 Liquidity, and this approximation may cause its ineffectiveness within the multiple linear regression format. Besides, an increased *EnergyVolume* and *TradeCount* does not necessarily lead to increased price volatility as prices may already reflect anticipated supply and demand changes.

4.3. Feature importance and benchmark models

We first consider the results of feature importance when performing PFI on the tuned RF regressor. We report the results of PFI on RF regressor in Fig. 5. These figures show the training set R^2 change for all features and then the final selection of features which contributed 0.005 or more change in the training set. We use this final set of features in performance evaluation for the RF regressor on the 2023 test set.

As discussed in Section 4.1 Regression models constructions, we consider additional HAR OLS models to aid in feature importance analysis as well to increase result robustness. The first of these models is a HAR OLS model which considers an optimal selection of all previous features as well as time dummy variables (*i.e.*, *Quarter*, *Month*, *Day of Week*, and *Week of Year*). We identify the optimal selection of features using RFE inputting an incrementally larger set of the total 23 features into the RFE algorithm to find both the optimal count and optimal set of features. The progression of R^2 as features are added is shown in Fig. 6.

The iterative RFE process finds that a certain subset of 17 features is the optimal number of features when using R^2 as the scoring metric. We use this optimal subset of 17 features to fit an HAR OLS model. We report this subset of features and their regression metrics in Table 12.

The optimal feature model still presents a fair number of features which are insignificant at the 0.005 level and excludes the intercept feature which is present in the four baseline models. Furthermore, the Adj. R^2 is very close to 100% which indicates that the model is overfitting and in need of feature pruning. Note this R^2 value differs from that shown in Fig. 6 since the latter shows the unadjusted R^2 obtained from the inclusion of a count of features in the RFE process, while Table 12 reports the Adj. R^2 after the optimal count of features have been utilized in a HAR OLS process. Thus, we implement an PFI algorithm to establish which of the selected features present significant importance to the model. Fig. 7 shows the features (post implementation of RFE and PFI), which exhibit importance above the 0.005 level during the PFI using MSE instead of training set R^2 as the scoring metric since overfitting is already present.

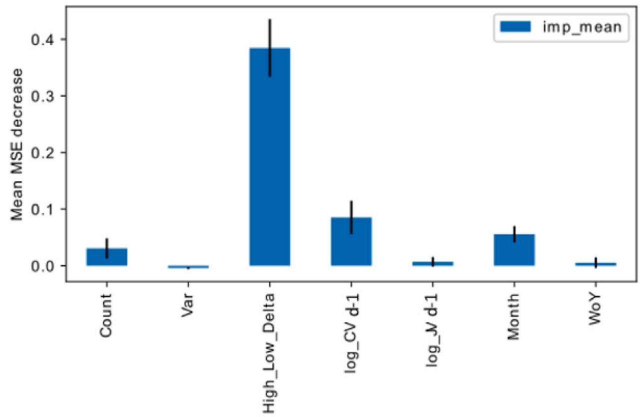


Fig. 7. Features from full HAR OLS model with score above 0.005 after PFI process.

Table 13
Parameter estimates for the post-PFI mixed HAR OLS model.

Parameter	Coefficient	Std. Error	P-value
Intercept	5.137	0.198	0.000
$PriceVar_d$	$8.773 \cdot 10^{-5}$	0.000	0.473
$HighLowDiff_d$	0.018	0.001	0.000
$logCV_{d-1}$	0.295	0.026	0.000
$logJV_{d-1}$	0.063	0.012	0.000
Adj. R^2		67.7%	

Table 14
Performance metrics for estimated models on 2023 test set using MAE, RMSE, and MAPE.

Model	MAE	RMSE	MAPE
HAR-CV-JV	0.735	0.937	**0.076**
Market Fundamentals	0.954	1.226	0.094
Market Fundamentals & UMM	0.833	1.074	0.083
Liquidity	0.973	1.200	0.096
RF Regressor	0.722	0.945	0.095
Mixed HAR OLS (17 features)	1.052	1.285	0.110
Mixed HAR OLS (5 features)	**0.684**	**0.879**	0.086

We then refit an HAR OLS model using the PFI important variables, whereby we removed all features not significant at the 0.005 level. Table 13 shows the results of the final mixed HAR OLS model which began by considering the full feature space and ended with four main features (*i.e.*, without considering the time features and the intercept), which are all closely related to ID price.

4.4. Performance comparison on 2023 test set

We measure the descriptive power of the different models using various out-of-sample criteria. We considered a wide set of models to identify differences between the approaches in explaining RV, with the additional benefit of adding robustness to the results. To this end, observations from 1 January 2020 to 31 December 2022, covering 1095 days, are considered as in-sample data, whereas observations from 1 January to 31 August 2023, covering 242 days, are considered out-of-sample test data. We make comparisons between observed and predicted values on the test set using mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) as performance criteria. Table 14 lists the results of the MAE, RMSE, and MAPE criteria for all of the feature estimation models presented in the prior section.

The results of the performance metrics show that the mixed MLR after RFE, PFI and 5% p-value trimming is the best-performing model with two of the three metrics. This is an intuitive outcome as it combines the descriptive power of the HAR-CV-JV model with the market intelligence of the liquidity measures. It is interesting to note that none of the market fundamentals features are present in the final mixed model indicating that price volatility in the 2023 continuous Dutch ID market is a construct composed of underlying market trends around trading behavior and macroeconomics factors rather than of daily market events such as load forecast errors or gas price.

5. Discussion

In this paper we discuss the recent trends in price movement and volatility within the Dutch EPEX continuous ID market. Price volatility has become a major component of the market in recent years due to geopolitical and generation profile changes across Europe. This paper utilizes RV to describe price patterns quantitatively. To answer the research question of how to best estimate RV, we developed a set of MLR models and compared their performance on a test set composed of recent trade data. Relying on distinctly different features, each of these models seeks to model RV from a different angle. The modeling process is made more robust by the additions of a fulsome mixed HAR OLS model and a RF regression model. These additional models also offer insight into feature importance when describing RV. Feature importance testing revealed that *PriceVar*, *HighLowDiff* and $logCV_{d-1}$ contribute the most to model efficacy. *PriceVar* and $logCV_{d-1}$ are shared across the final benchmark models, as well as *HighLowDiff* which shows the largest impact during PFI in the fulsome model. The modeling study reveals that effective modeling of RV is possible when features are well targeted. The test results of the final mixed HAR OLS model, which shows a relatively low MAE and RMSE scores, attests to the results of the feature importance approaches. Additionally, the HAR-JV-CV model had the best MAPE score highlighting the importance of these features in modeling RV.

However, the lack of presence from UMM messages and gas price in the best-performing models is a surprise. These features impact market demand and supply balance as a whole and as such, would be expected to be of significance in RV modeling. On the other hand, the lack of importance of ENTSO-E features, such as *vRESForecastErr*, is not as surprising since these features are well-understood in the market and the market design allows for compensating for them within the ID market. For example, a significant *LoadForecastErr* in the Netherlands can be compensated for by selling load to – or buying load from – a market participant in a country with an opposite load balance. While this study provided valuable insights into the dynamics of the energy market in the Netherlands using ID EPEX SPOT data, several limitations and opportunities for future research should be acknowledged. One significant limitation is the exclusive use of ID EPEX SPOT data due to data availability constraints. To enhance the robustness of the analysis and increase generalization to other national markets, future research could consider incorporating data from other European Power Exchanges, such as Nord Pool, which also have a significant influence on the region's energy pricing dynamics. Furthermore, it is important to recognize that factors beyond the scope of this study, such as geopolitical events, changes in energy policy, and extreme weather conditions, could significantly impact ID price movements. Future research can explore the integration of these external factors into a separate analysis to provide a comprehensive understanding of market dynamics under different impacts. Additionally, an expansion of the public vs. private information classification to be linked to real-time regulator decisions following ACER guidance would be illuminating regarding the potential impacts of smaller than 100 MW outages during critical market states on price volatility.

6. Conclusion

We define price volatility quantitatively using a measure called realized volatility, RV. A stakeholder can calculate RV at different points of time granularity. The study shows the importance of decomposing RV into its continuous and jump components as well as considering more common market measures. The mixed multi-linear HAR OLS model exhibits the best evaluation results on the 2023 test set for estimating RV, which offers support for the importance of RV decomposition in volatility estimation. In this current ID market environment, it is key to consider the innate volatility in the market prices caused by the cumulative behavior of all market actors. Furthermore, more simplistic market metrics, such as high-to-low price difference, are important features for increasing model efficacy. This combination of complex features and classical features was the best case when describing price volatility with RV under the market scenario presented. However, there are several elements of this research that can be improved. As mentioned, the BAS variable was estimated using procedures adapted for non-orderbook data. These procedures can introduce biases and inaccuracies and should be replaced with true BAS calculations for more concrete work on the impacts of liquidity on price volatility. Another significant avenue for further research is to expand the range of products and markets included in this analysis. For example, an aftermarket product (*i.e.*, a complementary product to the ID market which allows participants to trade in the ex-post timeframe to possibly reduce participants' balancing costs) was introduced in the Dutch and Belgian bidding zones on EPEX SPOT in 2022. This marketplace could be assessed through its own descriptive analysis, or as an additional component when explaining price volatility within either the DA or ID markets. Furthermore, expansion of the market research base into major additional markets like the Dutch DA, German ID and DA, and the Dutch balancing market price information can significantly enhance the depth and breadth of modeling for the Dutch ID price volatility. On a more technical side, future work could test different seasonality-correlation techniques. Besides, since several explanatory variables were constructed from the ID prices themselves (as shown in Table 12), future work could also explore comprehensive time-series analyses, involving research into multiple lagged variables (*e.g.*, d-1, d-2, d-3) and convolution/smoothing techniques. These additional inclusions will provide a greater holistic understanding of price formation and how volatility emerges by considering the broader market environment, including regional influences and real-time balancing needs. Simply put, the body of work on the Dutch market is too small compared to the amount of economic output from the Dutch electricity markets.

CRediT authorship contribution statement

Dane Birkeland: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tarek AlSkaif:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Methodology, Funding acquisition, Formal analysis. **Steven Duivenvoorden:** Writing – review & editing, Supervision, Resources, Funding acquisition, Data curation. **Marvin Meeng:** Writing – review & editing, Supervision, Resources. **Joost M.E. Pennings:** Writing – review & editing, Resources, Project administration.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

Acknowledgments

This research is partly supported by the Project High Energy Physics Tools in Limit Order Book Analysis (HighLO) between Wageningen University & Research (WUR), Netherlands and the European Organization for Nuclear Research (CERN) that aims to analyze and identify market abuse in electricity markets. This research was performed in collaboration with the energy and data departments of the Dutch regulator: Autoriteit Consument en Markt (ACM).

References

ACER, 2019. Methodology for Assessing the Relevance of Assets for Outage Coordination. ACER, Ljubljana.

ACER, 2023. Market abuse. REMIT URL https://www.acer.europa.eu/remit/coordination-on-cases/market-abuse.

All NEMOs Committee, 2023. Single Intraday Coupling (SIDC). SIDC Joint Documentation. URL https://www.nemo-committee.eu/sidc.

Andrade, J., Filipe, J., Reis, M., Bessa, R., 2017. Probabilistic price forecasting for day-ahead and intraday markets: Beyond the statistical model. Sustainability (ISSN: 2071-1050) 9, 1990. http://dx.doi.org/10.3390/su9111990.

Barndorff-Nielsen, O.E., Shepard, N., 2004. Power and bipower variation with stochastic volatility and jumps. J. Financ. Econom. (ISSN: 1479-8409) 2, 1–37. http://dx.doi.org/10.1093/jjfinec/nbh001.

Barndorff-Nielsen, O.E., Shepard, N., 2006. Econometrics of testing for jumps in financial economics using bipower variation. J. Financ. Econom. (ISSN: 1479-8409) 4, 1–30. http://dx.doi.org/10.1093/jjfinec/nbi022.

Birkeland, D., AlSkaif, T., 2024. Research areas and methods of interest in European intraday electricity market research—A systematic literature review. Sustain. Energy Grids Netw. (ISSN: 2352-4677) 101368. http://dx.doi.org/10.1016/j.segan.2024.101368.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

CBS, 2023. Statistics netherlands - economic growth. URL https://www.cbs.nl/en-gb/figures/detail/84106ENG.

Chan, K.F., Gray, P., van Campen, B., 2008. A new approach to characterizing and forecasting electricity price volatility. Int. J. Forecast. (ISSN: 01692070) 24, 728–743. http://dx.doi.org/10.1016/j.ijforecast.2008.08.002.

Chanatásig-Niza, E., Ciarreta, A., Zarraga, A., 2022. A volatility spillover analysis with realized semi(co)variances in Australian electricity markets. Energy Econ. (ISSN: 0140-9883) 111, 106076. http://dx.doi.org/10.1016/j.eneco.2022.106076, URL https://www.sciencedirect.com/science/article/pii/S0140988322002407.

Ciarreta, A., Muniain, P., Zarraga, A., 2017. Modeling and forecasting realized volatility in German-Austrian continuous intraday electricity prices. J. Forecast. (ISSN: 02776693) 36, 680–690. http://dx.doi.org/10.1002/for.2463.

Council of the European Union European Parliament, 2009. Directive 2009/28/EC. Off. J. Eur. Union 16–62, OJ L 140.

EPEX Spot, 2023. Market data services - EPEX Spot. URL https://www.epexspot.com/en/marketdataservices.

Frömmel, M., Han, X., Kratochvil, S., 2014. Modeling the daily electricity price volatility with realized measures. Energy Econ. (ISSN: 0140-9883) 44, 492–502. http://dx.doi.org/10.1016/j.eneco.2014.03.001, URL https://www.sciencedirect.com/science/article/pii/S0140988314000425.

Grinsztajn, L., Oyallon, E., Varoquaux, G., 2022. Why do tree-based models still outperform deep learning on tabular data?. http://dx.doi.org/10.48550/ARXIV.2207.08815, URL https://arxiv.org/abs/2207.08815.

Hagemann, S., Weber, C., 2013. An empirical analysis of liquidity and its determinants in the German intraday market for electricity. SSRN Electron. J. (ISSN: 1556-5068) http://dx.doi.org/10.2139/ssrn.2349565.

Haugom, E., 2011. Some stylized facts about high-frequency nord pool forward electricity prices. J. Energy Markets (ISSN: 17563607) 4, 21–49. http://dx.doi.org/10.21314/JEM.2011.052.

Haugom, E., Ullrich, C.J., 2012. Forecasting spot price volatility using the short-term forward curve. Energy Econ. (ISSN: 01409883) 34, 1826–1833. http://dx.doi.org/10.1016/j.eneco.2012.07.017.

Hu, X., Jaraitė, J., Kažukauskas, A., 2021. The effects of wind power on electricity markets: A case study of the Swedish intraday market. Energy Econ. (ISSN: 01409883) 96, 105159. http://dx.doi.org/10.1016/j.eneco.2021.105159.

Huang, X., Tauchen, G.E., 2005. The relative contribution of jumps to total price variance. J. Financ. Econom..

Jain, M., AlSkaif, T., Dev, S., 2022. Are deep learning models more effective against traditional models for load demand forecasting? In: 2022 International Conference on Smart Energy Systems and Technologies. SEST, pp. 1–6. http://dx.doi.org/10.1109/SEST53650.2022.9898424.

Karakatsani, N.V., Bunn, D.W., 2010. Fundamental and behavioural drivers of electricity price volatility. Stud. Nonlinear Dyn. Econom. (ISSN: 1558-3708) 14, http://dx.doi.org/10.2202/1558-3708.1657.

Karanfil, F., Li, Y., 2017. The role of continuous intraday electricity markets: The integration of large-share wind power generation in Denmark. Energy J. (ISSN: 01956574) 38, http://dx.doi.org/10.5547/01956574.38.2.fkar.

Kiesel, R., Paraschiv, F., 2017. Econometric analysis of 15-minute intraday electricity prices. Energy Econ. (ISSN: 01409883) 64, 77–90. http://dx.doi.org/10.1016/j.eneco.2017.03.002.

Kollewe, J., 2022. EDF cuts output at nuclear power plants as French rivers get too warm. The Guardian Europe URL https://shorturl.at/uDL15.

Le, H.L., Ilea, V., Bovo, C., 2019. Integrated European intra-day electricity market: Rules, modeling and analysis. Appl. Energy 238, 258–273.

Lee, S.S., Mykland, P.A., 2007. Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. Rev. Financ. Stud. (ISSN: 0893-9454) 21 (6), 2535–2563. http://dx.doi.org/10.1093/rfs/hhm056, arXiv:https://academic.oup.com/rfs/article-pdf/21/6/2535/24453758/hhm056.pdf.

Lybek, T., Sarr, A., 2002. Measuring liquidity in financial markets. 02, p. 1. http://dx.doi.org/10.5089/9781451875577.001, IMF Working Papers.

Maciejowska, K., Nitka, W., Weron, T., 2019. Day-ahead vs. intraday—Forecasting the price spread to maximize economic benefits. Energies (ISSN: 1996-1073) 12, 631. http://dx.doi.org/10.3390/en12040631.

Martin, H., Otterson, S., 2018. German intraday electricity market analysis and modeling based on the limit order book. In: 2018 15th International Conference on the European Energy Market. EEM, IEEE, ISBN: 978-1-5386-1488-4, pp. 1–6. http://dx.doi.org/10.1109/EEM.2018.8469829.

Mulder, M., Scholtens, B., 2013. The impact of renewable energy on electricity prices in the Netherlands. Renew. Energy (ISSN: 09601481) 57, 94–100. http://dx.doi.org/10.1016/j.renene.2013.01.025.

Our World in Data, 2023. Share of primary energy from renewable sources. URL https://ourworldindata.org/renewable-energy.

Pape, C., Hagemann, S., Weber, C., 2016. Are fundamentals enough? Explaining price variations in the German day-ahead and intraday power market. Energy Econ. (ISSN: 01409883) 54, 376–387. http://dx.doi.org/10.1016/j.eneco.2015.12.013.

Qu, H., Duan, Q., Niu, M., 2018. Modeling the volatility of realized volatility to improve volatility forecasts in electricity markets. Energy Econ. (ISSN: 0140-9883) 74, 767–776. http://dx.doi.org/10.1016/j.eneco.2018.07.033, URL https://www.sciencedirect.com/science/article/pii/S014098831830286X.

Scharff, R., Amelin, M., 2016a. Trading behaviour on the continuous intraday market Elbas. Energy Policy 88, 544–557.

Scharff, R., Amelin, M., 2016b. Trading behaviour on the continuous intraday market Elbas. Energy Policy (ISSN: 03014215) 88, 544–557. http://dx.doi.org/10.1016/j.enpol.2015.10.045.

scikit-learn developers, 2023a. sklearn.ensemble.RandomForestRegressor. URL https://shorturl.at/hvANU.

scikit-learn developers, 2023b. sklearn.feature_selection.RFE. URL https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.RFE.html.

scikit-learn developers, 2023c. sklearn.inspection.permutation_importance. URL https://scikit-learn.org/stable/modules/generated/sklearn.inspection.permutation_importance.html.

Shinde, P., Kouveliotis-Lysikatos, I., Amelin, M., 2021. Analysing trading trends in continuous intraday electricity markets. In: 2021 56th International Universities Power Engineering Conference. UPEC, IEEE, ISBN: 978-1-6654-4389-0, pp. 1–6. http://dx.doi.org/10.1109/UPEC50034.2021.9548168.

Shiryaevskaya, A., Mazneva, E., Mathis, W., 2022. European energy prices soar on summer crisis, winter gloom. Bloomberg URL https://shorturl.at/klvA7.

Tanaka, K., Matsumoto, K., Keeley, A.R., Managi, S., 2022. The impact of weather changes on the supply and demand of electric power and wholesale prices of electricity in Germany. Sustain. Sci. (ISSN: 1862-4065) 17, 1813–1825. http://dx.doi.org/10.1007/s11625-022-01219-7.

Uyterlinde, M.A., Junginger, M., de Vries, H.J., Faaij, A.P., Turkenburg, W.C., 2007. Implications of technological learning on the prospects for renewable energy technologies in Europe. Energy Policy (ISSN: 03014215) 35, 4072–4087. http://dx.doi.org/10.1016/j.enpol.2007.02.004.

Valitov, N., Maier, A., 2020. Asymmetric information in the German intraday electricity market. Energy Econ. (ISSN: 01409883) 89, 104785. http://dx.doi.org/10.1016/j.eneco.2020.104785.

Visser, L., AlSkaif, T., van Sark, W., 2022. Operational day-ahead solar power forecasting for aggregated PV systems with a varying spatial distribution. Renew. Energy 183, 267–282.

Weber, C., 2010. Adequate intraday market design to enable the integration of wind energy into the European power systems. Energy Policy (ISSN: 03014215) 38, 3155–3163. http://dx.doi.org/10.1016/j.enpol.2009.07.040.