Phase 4 : DEVELOPMENT
PART 2

Name: Emi Hedrif X

Register Number: 312621243012

College Name: Thangavelu engineering college

Project 4: Electricity Prices Prediction

1. Introduction

Day-ahead electricity market (DAM) prices have historically been driven by fundamental drivers or fundamentals reflecting an intrinsic demand and supply of electricity. Fundamentals are associated with the intrinsic value of a good, commodity, or security. Examples include goods' production costs, economic variables, etc. In a world of perfectly rational investors, only fundamentals are considered to drive prices as investors always optimally maximise their utility.

In recent years, regulatory changes, such as [1], have swept through electricity markets provoking disruptive propagotary shocks. These shocks, which culminated from a growing need to generate cleaner and safer energy [2,3], have indirectly boosted the impacts of technical price drivers or technicals on the DAM and moved prices further from their intrinsic values. The term "technical" comes from the discipline of technical analysis (TA) [4]. Technicals are factors which move prices away from their intrinsic values. They result because, in practice, investors are not perfectly rational and often make seemingly sub-optimal decisions. Behavioural economics focuses on explaining how investors' biases, emotions, and other psychological factors influence decisions and, by extension, prices. As a result of the impacts of technicals, the identification of new model features, capable of capturing the residual impacts of technicals, has become necessary to forecast DAM prices.

Describing the propagation of regulatory shocks across day-ahead and intraday markets, the incentivisation of renewable electricity generation spurred investment in renewable technologies, increasing efficiencies, reducing costs, and accelerating uptake. An accelerating uptake of renewables, however, increased electricity demand and supply variability and, by extension, volume and price forecast uncertainty. Given the need to avoid energy imbalances, a greater uncertainty surrounding delivery volumes impacted traders' planning decisions, amplifying day-ahead and intraday trading. Amplified intraday trading, in turn, increased the market's liquidity and decreased its average bid-ask spread, provoking a feedback loop which spurred yet more intraday trading. Although DAM volumes continued to overwhelmingly exceed intraday volumes throughout, greater intraday liquidity further impacted traders' planning decisions; boosting their willingness to distribute orders across day-ahead and intraday markets.

Because traders' decisions are not always perfectly rational, a greater willingness to distribute orders across day-ahead and intraday markets accentuated the impacts of technicals, moving prices away from their intrinsic values. Exemplifying this point, note that all trades on electricity markets are placed because of an underlying stochastic energy need. Whereas before the boom in renewable generation, traders almost exclusively utilised the DAM to meet this need in the short-term, resulting in planning decisions rooted principally in fundamentals, today, intraday opportunities offer traders many more ways to maximise their respective reward to risk ratio, resulting in more heterogeneous short-term planning decisions and the proliferation of technicals. Day-ahead and intraday substitution can thus be understood to have heightened the impacts of technicals on DAM prices.

Introducing existing modelling approaches for forecasting DAM prices, historically researchers have relied upon features primarily reflecting the fundamentals of demand and supply. On the supply side, features, such as gas and coal prices [5], which capture the cost of generating electricity, have been used. On the demand-side, with weather conditions and circadian rhythm dictating consumers' daily routines, meteorological factors [5] as well as seasonal dummies [6], ranging from the hour of the day to the month of the year, have been used. While the above-mentioned features, along with a multitude of others like grid-load and available generation [5,7], explain a significant portion of the DAM's variance, they nonetheless fail to capture the impacts of technical price drivers.

Lagged price features, which consist of current and past prices, inadvertently capture the impacts of technicals [8]. However, as they are an amalgamation of all historically available information, of which technicals are just a portion, it is difficult for models to identify technical signals from lagged prices. Hence, although researchers have used lagged prices to forecast the DAM [9,10], the identification of new technical features is nevertheless required to capture the residual impacts of technicals.

To capture the impacts of technicals, an approach to forecasting future price movements using historic market data is needed. Technical analysis (TA) [4], widely used by practitioners to identify investment opportunities across financial markets, offers such an approach. Rooted in the theories of behavioural finance [11], TA, as an analysis methodology, focuses on analysing statistical trends in historical market data to forecast future price movements. TA was established around three assumptions: prices move in trends; history repeats itself; studying price fluctuations allows the prediction of future shifts in demand and supply [4]. In short, TA assumes that technicals form reoccurring patterns in market data which careful analysis can identify and predict. One of the ways in which TA identifies emerging price patterns is using technical indicators (TIs). TIs transform market data using various formulas. By transforming the data, TIs can facilitate the identification of complex price patterns; signalling when securities are overbought or oversold, when prices deviate from a central trend, or when they are near support and resistance lines (levels that prices bounce off of and break through). Moreover, in conjunction with price charts, TIs can provide leading indications of future price movements, offering additional explanatory power to models forecasting financial time series. Numerous

studies, such as [12–14], have extensively demonstrated the explanatory power of TIs across stock markets. Given that technicals result from universally exhibited behavioural biases among investors, such studies motivate our examination of TIs with the DAM.

Summarising the principle contribution of our paper, to the best of our knowledge, we are the first to evaluate the explanatory power of TIs in DAM price forecasting. We demonstrate that the simple inclusion of TI features can significantly reduce linear, ensemble, and deep model regression forecast errors, on average reducing the root mean squared error (RMSE) by 3.28% and the mean absolute error (MAE) by 3.32%. Describing the structure of this paper, we introduce our examined TIs, modelling approaches, and case study in Sections 2.1–2.3 respectively. Subsequently, in Section 3 we present and evaluate our results. Finally, in Section 4 we discuss possible extensions to our research.

2. Materials and Methods

2.1. Technical Indicators

While a plethora of TIs exists, none can guarantee the addition of explanatory power. This is because the success of individual TIs is domain and period-dependent. For instance, price-based trend-following indicators may be useful in times of high autocorrelation; however, they cease being indicative of future price moves when autocorrelation vanishes and trends become spurious [15]. We are mindful of this while selecting a list of TIs to examine, and as a result, choose TIs satisfying the following two criteria.

1. TI calculation only requires close prices.
2. TI inclusion is likely to improve predictive performance by highlighting oscillations or trends in DAM prices.

Criterion 1 necessitates from the properties of the DAM, neither *high*, *low*, nor *open* prices are available. Other financial markets record security's *open*/*close*/*high*/*low* prices, which are the initial/last/highest/lowest recorded trading prices over a given period, respectively [16]. Together, these metrics can be used to convey primary information about a security's price movements over a specific period. Unlike other financial markets, however, the DAM only records a single clearing price for each hourly contract. We consider this price to be the *close* price.

Criterion 2 stems from our desire to avoid examining overly intricate TIs which have been optimised for rare use-cases and evaluate TIs which are well suited for the DAM. DAM prices typically move between horizontal support and resistance lines, behaving comparably to ranging markets whose prices move within a band. On occasion, however, DAM prices breakout from this price band, and establish a trend; behaving comparably to trending markets. Overall, such varying behaviour spurs us to evaluate both oscillators, which vary around a central line or within set levels, and trend-following TIs, which measure the direction of a trend. The former TIs are well suited for ranging markets, while the latter TIs are optimised for trending markets.

Together, criteria 1 and 2 allow us to follow the suggestions of [13] by focusing our research on the simplest and most widely used TIs, which are likely to introduce more explanatory power than noise. Below, we introduce our list of chosen TIs and provide formulas for their calculation. Note that throughout the paper the following notations are used:

t time,
p_t price at time t,
n lag-factor for $\forall n \in \mathbb{Z}$,
s span for $\forall s \in \mathbb{Z}$,

where n and s are hyperparameters tuned using grid-search during the modelling of DAM prices.

2.1.1. Simple Moving Average (SMA)

The SMA [4], a type of moving average, is the most fundamental TI. It is often used as a building block in the calculation of other compound indicators such as Bollinger Bands (BBANDs). The SMA captures trends by smoothing a price series using a lag-factor, n. A single SMA curve, either alone or in conjunction with the price series, can be used to forecast future price movements and generate trading ideas [17]. Specifically, the SMA can be used to identify support and resistance lines. Furthermore, SMA crossovers can be used to identify emerging price trends or consolidations [18] (when the price crosses the SMA from above/below, it can signal that prices have peaked/troughed. A downward/upward trend may follow). The SMA is calculated according to Equation (1).

$$\text{SMA}(p_t, n) = \frac{p_t + \ldots p_{t-(n-1)}}{n} = \frac{1}{n}\sum_{i=0}^{n-1} p_{t-i}. \tag{1}$$

2.1.2. Exponential Moving Average (EMA)

The EMA [4] is a special type of moving average that exponentially averages historic prices. Through its weighting, the EMA can place greater significance on more recent price trends. This weighting distinguishes the EMA from the SMA and allows the EMA to more rapidly reflect immediate price movements. During periods of high volatility, placing more weight on more recent price moves can be an advantage. The EMA is calculated according to Equation (2).

$$\text{EMA}(p_t, s) = \frac{p_t + \alpha \times p_{t-1} + \ldots \alpha^t \times p_0}{1 + \alpha + \ldots \alpha^t}, \tag{2}$$

where $\alpha = \frac{s-1}{s+1}$ is the weighting term. α can be tailored to give more or less importance to the recent past.

2.1.3. Moving Average Convergence Divergence (MACD)

The MACD [4] is a trend-following momentum indicator comprising of three time series: the MACD 'Series', 'Signal', and 'Histogram'. These series can be used in tandem to formulate trading rules, for instance, extending the double crossover trading strategy [19]. As we are not aiming to construct trading strategies, nor wish to introduce too many features at once, we split the MACD into its components, treating each as an individual indicator. We describe these indicators below.

- **'Series'**: Calculated from two EMAs, the 'Series' [17] gives insight into price convergence, divergence, and crossover. The 'Series' reflects the difference between a fast (e.g., $s = 12$) and a slow (e.g., $s = 26$) EMA, capturing the second derivative of a price series. Using Equation (2), the 'Series' is calculated according to Equation (3).

$$\text{Series}(p_t, s_1, s_2) = \text{EMA}(p_t, s_1) - \text{EMA}(p_t, s_2), \tag{3}$$

  where $s_2 > s_1$.
- **'Signal'**: The 'Signal' [17] is an EMA of the 'Series'. It provides a lagging indication of crossovers between fast and slow EMAs. Using Equations (2) and (3), the 'Signal' is calculated according to Equation (4).

$$\text{Signal}(p_t, s_1, s_2, s) = \text{EMA}(\text{Series}(p_t, s_1, s_2), s). \tag{4}$$

- **'Histogram'**: The 'Histogram' [17] is the difference between the 'Series' and 'Signal'. Mathematically, it can be interpreted as the fourth derivative of a price series, anticipating changes in the 'Series'. Using Equations (3) and (4), the 'Histogram' is calculated according to Equation (5).

$$\text{Histogram}(p_t, s_1, s_2, s) = \text{Series}(p_t, s_1, s_2) - \text{Signal}(p_t, s_1, s_2, s).$$

2.1.4. Moving Standard Deviation (MSD)

The MSD [20], measuring the rolling n day volatility of prices, is considered helpful in predicting the size of future price moves. The indicator anticipates periods of low volatility following periods of high volatility and vice versa. Using Equation (1), the MSD is calculated according to Equation (6).

$$\text{MSD}(p_t, n) = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}(p_{t-i} - \text{SMA}(p_t, n))^2}. \tag{6}$$

2.1.5. Bollinger Bands (BBANDs)

BBANDs [4] consist of two bands: the BBAND$^+$, calculated according to Equation (7), and the BBAND$^-$, calculated according to Equation (8).

$$\text{BBAND}^+(p_t, n) = \text{SMA}(p_t, n) + 2 \times \text{MSD}(p_t, n), \tag{7}$$

$$\text{BBAND}^-(p_t, n) = \text{SMA}(p_t, n) - 2 \times \text{MSD}(p_t, n). \tag{8}$$

These bands, each two MSDs away from the SMA, indicate when a security is overbought (price above the BBAND$^+$) or oversold (price below the BBAND$^-$). BBANDs can be used to facilitate the prediction of future increases/decreases in volatility, and to identify technical signals such as the W-Bottom, explained in AppendixA, [21] . As we want to avoid introducing too many features, we explore two compound TIs derived from BBANDs: the %B and the Bandwidth.

- **%B**: The %B [22] scales the price series by the BBAND width. When the underlying security price equals the SMA, the %B equals 0.5. When the price is equal to the BBAND$^-$/BBAND$^+$, the %B equals 0/1 respectively. Similarly to BBANDs, the %B can be used to identify when prices are overbought or oversold, to predict future volatility and to generate trading ideas.
Using Equations (7) and (8), the %B is calculated according to Equation (9).

$$\%B(p_t, n) = \frac{p_t - \text{BBAND}^-(p_t, n)}{\text{BBAND}^+(p_t, n) - \text{BBAND}^-(p_t, n)}. \tag{9}$$

- **Bandwidth**: The Bandwidth [22] measures the BBANDs divergence and is used to anticipate changing volatility and price breakouts. Using Equations (1), (7) and (8), the Bandwidth is calculated according to Equation (10).

$$\text{Bandwidth}(p_t, n) = \frac{\text{BBAND}^+(p_t, n) - \text{BBAND}^-(p_t, n)}{\text{SMA}(p_t, n)}. \tag{10}$$

2.1.6. Momentum (MOM)

The MOM [4] is a trend-following leading indicator. Elaborating, the MOM provides insight into price trends, acting as a signal to buy/sell when crossing above/below the zero line. Unlike the SMA, the MOM can peak or trough before the price, providing a forward-looking ('leading') trend prediction. As a forward-looking indicator, when the MOM peaks or troughs and begins to diverge from the main price trend, it can signal bearish or bullish divergence. The MOM is calculated according to Equation (11).

$$\text{MOM}(p_t, n) = p_t - p_{t-n}. \tag{11}$$

2.1.7. Rate of Change (ROC)

The ROC [4,17] is an oscillator, comparable to the MOM indicator, that expresses change as a percentage instead of an absolute value. As a standardised measure of change, the ROC can be

used to identify overbought or oversold extremes that previously foreshadowed a trend reversal. Note that when above zero, the ROC indicates an overall uptrend, and when below zero, it indicates a downtrend. When prices are moving within a fixed corridor/range, the ROC remains near zero, confirming price consolidation. In these instances, the ROC provides little insight about future price movements. The ROC is calculated according to Equation (12).

$$\text{ROC}(p_t, n) = \frac{p_t - p_{t-n}}{p_{t-n}}. \tag{12}$$

2.1.8. Coppock Curve (COPP)

The COPP [23] is a smoothed momentum oscillator. Although the COPP was originally developed to capture long-term price trends occurring in American equities, since its inception it has been used to identify both long and short-term trends in numerous markets, e.g., [17]. Using Equations (2) and (12), the COPP is calculated according to Equation (13).

$$\text{COPP}(p_t, s, n_1, n_2) = \text{EMA}(\text{ROC}(p_t, n_1) + \text{ROC}(p_t, n_2), s). \tag{13}$$

2.1.9. True Strength Index (TSI)

Providing trend insights, as well as indications of when a security is overbought or oversold, the TSI [24] is a smoothed momentum oscillator. Technical analysts often look for trend lines in the TSI to identify support and resistance price bands. Using Equation (2), the TSI is calculated according to Equation (14).

$$\text{TSI}(p_t, s_1, s_2) = \frac{\text{EMA}(\text{EMA}(p_t - p_{t-1}, s_1), s_2)}{\text{EMA}(\text{EMA}(|p_t - p_{t-1}|, s_1), s_2)}. \tag{14}$$

2.2. Models

Significant performance gains, using multiple models, must be observed if we are to robustly demonstrate the added explanatory power TIs offer. With this in mind, we identify a list of high-performing models to evaluate TIs. After consulting [5,6], we settle on using machine learning (ML) models. ML models can be classified into three primary groups: linear, ensemble and deep. Below, we briefly introduce our chosen ML models from each group.

2.2.1. Linear Models

- **Linear Regression (LR)**: The most fundamental of linear models, LR [25] fits a straight line through a series of points by minimizing the sum of squared errors between its targets and predictions. LRs are sensitive to outliers and correlated features. Nevertheless, as they are one of the primary ML models used for DAM forecasting, we include them in our examination.
- **Huber Regression (HR)**: Extending LRs, HR [26] is a linear model robust to response variable outliers. Unlike LR, HR optimises both an absolute and squared loss function, reducing the impact of outliers. To switch between loss functions, HR uses an epsilon hyperparameter. Despite implementing an enhanced optimisation procedure, HR remains sensitive to explanatory variable outliers and correlations.

2.2.2. Ensemble Models

- **Random Forest (RF)**: RF [27] fits several decision trees on random samples of the data and averages them to obtain a final result. Individual trees are fit by recursively splitting the data in such a way that maximises the information gain. The hyperparameters used by RF are the number of fitted estimators, the maximum number of features, and the minimum sample leaf size.
- **AdaBoost (AB)**: AB [28] is an adaptive boosting algorithm used to sequentially train an ensemble of weak learners. The algorithm begins by fitting a weak learner, and continues by training copies

of this learner, placing a greater instance weight on incorrectly predicted values. The algorithm proceeds until the final model, a weighted sum of all trained weak learners, becomes a strong learner. We use the algorithm to train an ensemble of decision trees.

- **Gradient Boosting (GB)**: Another boosting algorithm, GB [29] focuses on sequentially improving model predictions by fitting copies of learners to residuals. Residual predictions are repeatedly added to model predictions until the sum of residuals stops decreasing. Similarly to AB, we choose to apply the GB algorithm to train an ensemble of decision trees.

2.2.3. Deep Models

Deep models, or artificial neural networks, consist of an input, output, and hidden layer. Of these three, the hidden layer is the most varied, with architectures differing in depth and layer make-up. Below, we describe some of the most fundamental layers and modules used in deep models.

- **Fully Connected Layer (FCL)**: Fully connected neurons, comprising of linear regression with an added non-linearity [30], are stacked to build an FCL. FCLs can be used to approximate any continuous function [31], explaining why, with increasing computational power, they are frequently used in state-of-the-art DAM predictors.
- **Convolutional Layer (CONV)**: Locally connected neural networks, or convolutional neural networks (CNNs) [32], are used for feature mapping/extraction. The primary module used by these networks, CONV, works by sliding equally sized filters with trainable parameters across input data producing 2D activation maps. While FCLs tune the parameters of every neuron, CONVs implement parameter sharing to remain computationally feasible. For a single CONV filter of size (NxM), *NxMxdepth* parameters are trained. The inclusion of the depth parameter is a consequence of CONV's fully connected architecture across the final depth dimension. Overall, CONVs have proved adept at identifying features in images [33] and time series [34], making them potentially very powerful modules for technical analysis.
- **Residual Module**: A residual module, ResNet [35], adds the inputs from one module to the outputs of another module. It thus creates a direct identity mapping in a network between module inputs and outputs, combating both the vanishing gradient problem and the degradation problem, which otherwise impede the training of deep networks.

We explore the performance of TIs with each of the above-described layers and modules. Specifically, we use neural networks with: two FCLs (**2NN**); a single CONV (**CNN**); two CONVs (**2CNN**); two CONVs and a single FCL (**2CNN_NN**); seven CONV residual modules (**ResNet**).

2.3. Case Study

We focus on demonstrating the explanatory power of TIs by boosting forecasting accuracies of Belgian DAM prices. Introducing the Belgian DAM, like to other European countries, short-term energy contracts are traded on three markets in Belgium; the day-ahead, the intraday, and the imbalance markets. Of these, the DAM, offering hourly denominated contracts for next day delivery, is the largest, continually recording the greatest volume of trades. Several characteristics distinguish the DAM from the other energy markets. Most prominently the clearing procedure is different. DAM clearing requires trades to be submitted before noon, after which point hourly prices (in €/MWh) are computed using a matching engine. The matching engine sets prices where supply and demand curves intersect. More information about the Belgian DAM can be found at [36]. Below, we introduce our data, data processing, TI calculation method, as well as our training and evaluation procedures.

2.3.1. Data

A dataset of historic DAM prices, spanning four and a half years, is gathered from [36]. Three and a half years of prices, from 1 January 2014 to 29 June 2017, are used for hyperparameter tuning and model training. A single year of prices, from 30 June 2017 to

30 June 2018, is used for testing. Figure 1 visualises the dataset and highlights the train /test split. The summary statistics (in €/MWh) of the train/test data are: mean, 41.12/44.96; standard deviation, 20.16/19.39; maximum, 696.02/330.77; minimum, −5.00/−41.31; lower quartile, 30.12/33.50; median, 39.05/41.90; upper quartile, 48.99/52.91.
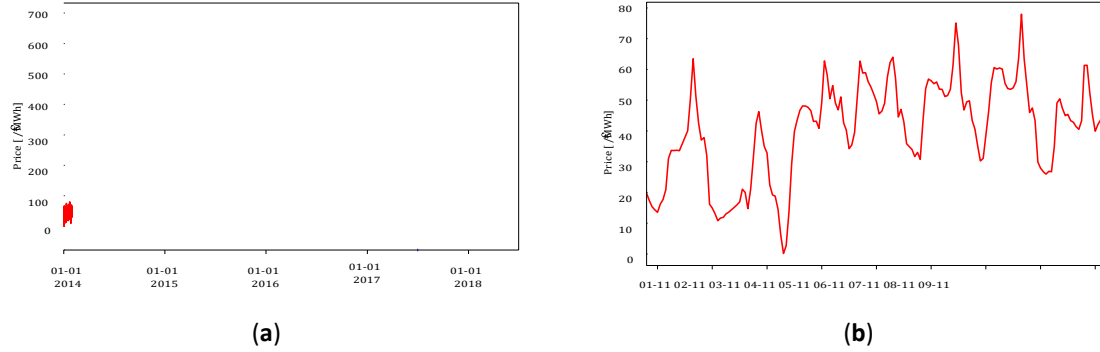

(a)   (b)

Figure 1. Belgian DAM prices. (**a**) Historic DAM prices, covering the training and test spans: 1 January 2014 to 30 June 2018. The dashed blue line represents the train/test split. (**b**) Single week of DAM prices from 1 November 2014 to 9 November 2014.

2.3.2. Data Processing

Data processing is conducted to maximise the performances of the models. Firstly, using Min–Max scaling, both features and response variables are scaled. Min–Max scaling bounds each value $a \in A$ between 0 and 1 according to the formula: $a^* = (a − A_{min})/(A_{max} − A_{min})$, where A_{min} and A_{max} are the minimum and maximum values of a set A, and a^* is the scaled value of a. Secondly, to capture significant serial correlation, an autocorrelation plot (ACF) is examined. ACFs plot lags on the x-axis and autocorrelation coefficients on the y-axes [37]. The plots visually facilitate the selection of an optimal look-back period. Capturing all significant terms in our ACF plot, we select a 6-day look-back period with an additional averaged 8-week look-back. Formally, to forecast day $d + 1$ prices, we use six prices from days d to $d − 5$ along with the average of $d − 6, d − 13, \ldots, d − 55$ prices.

2.3.3. TI Calculation

The TI formulas, presented in Section 2.1, require sequential inputs, i.e., continuous linear time. The DAM, however, is not sequential because it releases 24 prices simultaneously upon clearing. Consequently, we treat the dataset as an assortment of prices from 24 separate markets when calculating DAM TIs. Formally, we use h hour prices to calculate h hour TIs. For instance, the 12 h SMA ($n = 3$) at 9 November 2014 is calculated by taking the average of three 12 h prices: the 7, 8 and 9 of November. Finally, in order to identify tailored TIs for every model specified in Section 2.2, we optimise TI hyperparameters using grid-search separately for each model.

2.3.4. Model Training and Prediction

To be able to evaluate TI performance, we train each model introduced in Section 2.2 twice: once as a benchmark model accepting only lagged prices, and once as a TI model accepting both lagged prices and a lagged TI. Depending on the TI model, one of two approaches is used to concatenate inputs. Length-wise concatenation is used with linear and ensemble models, while channel-wise concatenation is used with deep models. Figure 2 visualises these approaches. To select optimal hyperparameters for each model, we divide the training set in two using a 3:1 ratio to get training and validation sets. Grid-search, which maximises benchmark model validation performance, is subsequently conducted using these sets. A complete specification of the resulting model hyperparameters can be found in Appendix B.
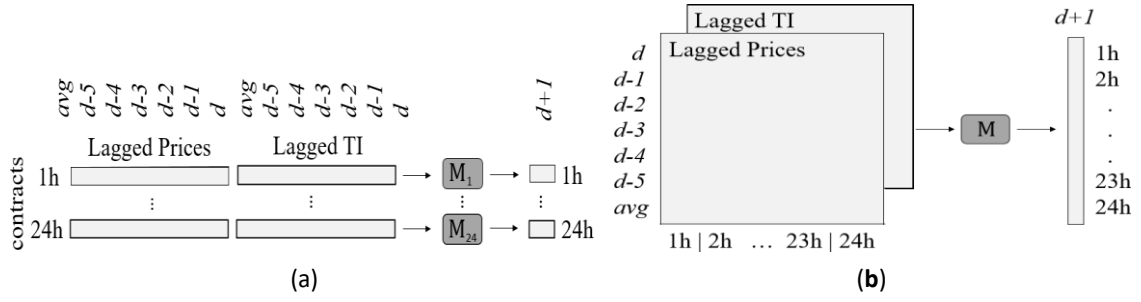
Figure 2. Diagrams displaying TI model inputs—lagged prices concatenated with a lagged TI—and outputs. Per Section 2.3.2, d to $d-5$ represent the 6-day look-back period, while *avg* stands for the averaged 8-week look-back. **M** is a black box representation of ML models. (**a**) Linear and ensemble models; (**b**) Deep models.

Forecasting the DAM necessitates predicting 24 h prices. Because linear and ensemble models do not support multi-output prediction, but deep models do, to forecast the DAM we use different training methods for linear, ensemble, and deep models. These methods are visualised in Figure 2 for TI models. To forecast all 24 prices with linear and ensemble models, we split the data into 24 separate sets and train 24 single-output regressions. For instance, to predict the next day's 12 h price, a benchmark model is trained using a set of lagged 12 h prices. On the other hand, to forecast DAM prices with deep models, we train a multi-output regression. Thus, to predict next days 01–24 h prices, a benchmark deep model is trained using the entire training set of lagged prices.

2.3.5. Evaluation

To evaluate the explanatory power of TIs, benchmark and TI model performances are compared. In order to assess performance, various metrics measuring the discrepancies and similarities between a model's targets and predictions are computed. Firstly, the root mean squared error (RMSE) and the mean absolute error (MAE) are calculated for benchmark and TI models. By analysing the percentage change in these discrepancy metrics, %RMSE and %MAE, overall accuracy improvements are determinable. The %RMSE and %MAE are calculated according to Equation (15).

$$\%ME = 100 \times \frac{ME^{Bench} - ME^{TI}}{ME^{Bench}}, \quad \Sigma \tag{15}$$

where ME^{Bench} is the mean benchmark model error and ME^{TI} is the mean TI model error. A positive %RMSE and %MAE indicate that the inclusion of TI features, overall, reduces forecast errors.

Secondly, the Pearson correlation coefficient (PCC) is calculated for benchmark and TI models. The PCC is a similarity metric, whose square is equal to the coefficient of determination. By analysing the percentage change in this metric, %PCC, overall improvements in a model's goodness of fit are ascertainable. The %PCC is calculated according to Equation (16).

$$\%PCC = 100 \times \frac{PCC^{TI} - PCC^{Bench}}{PCC^{Bench}}. \quad \Sigma \tag{16}$$

A positive %PCC indicates that the inclusion of TI features, overall, increases a model's goodness of fit.

Thirdly, interquartile and tail squared residuals are computed for benchmark and TI models. Interquartile squared residuals, $\{(e_t)^2\}^{N_{IQR}}_{t=1}$, are the forecast errors of interquartile targets, $\{\{y_t\}^{N_{IQR}}_{t=1} \mid Q1 \leq y_t \leq Q3\}^T_{t=1}$, where T is the test size, and $Q1$ and $Q3$ are the lower and upper quartiles respectively. Similarly, tail squared residuals, $\{(e_t)^2\}^{N_{Tail}}_{t=1}$, are the forecast errors of tail targets, $\{\{y_t\}^{N_{Tail}}_{t=1} \mid y_t < Q1 \cup y_t > Q3\}^T_{t=1}$. Four custom measures of forecasting accuracy, %W, %L, $\bar{W}$,

and $\bar{L}$ are calculated from both the interquartile and tail errors. %W and %L measure the percentage of squared TI residuals smaller/larger than squared benchmark residuals. $\bar{W}$ and $\bar{L}$ measure the mean squared error difference between TI and benchmark model forecasts. By analysing these four metrics, improvements in interquartile and tail accuracies are determinable. Formally, with either the interquartile or tail sets, %W, %L, $\bar{W}$ and $\bar{L}$ are calculated according to Equations (17) and (18).

$$\%W = 100 \times \frac{n_w}{N}, \quad \bar{W} = \frac{1}{n_w} \times \sum_{t=1}^{n_w} d_t, \quad d \quad d^{n_w} d \quad 0^N, \quad \forall_t \in \{\{d_t\}_{t=1} \mid t < \}_{t=1} \tag{17}$$

$$\%L = 100 \times \frac{n_l}{N}, \quad \bar{L} = \frac{1}{n_l} \times \sum_{t=1}^{n_l} d_t, \quad d \quad d^{n_l} \quad d\text{``}0^N, \quad \forall_t \in \{\{d_t\}_{t=1} \mid t \quad \}_{t=1} \tag{18}$$

where $d_t = (e^{TI}_t)^2 - (e^{Bench}_t)^2$, and N is either N_{IQR} or N_{Tail}. A %W higher than 50% and a negative $\bar{W} + \bar{L}$ are indicative of more accurate TI model forecasts.

Finally, to determine the statistical significance of any performance improvements, one tailed Diebold–Mariano (DM) tests [38] with a mean squared error loss function are conducted. The DM test statistically evaluates differences between two models' forecasts by comparing their residuals, in our case $\{e^{Bench}_t\}^T_{t=1}$ and $\{e^{TI}_t\}^T_{t=1}$ The test converts a loss-differential, $l_t = g(e^{Bench})_t - g(e^{TI})$ where $g(.)$ is the loss function, into an asymptotically normal DM statistic. Using the DM, we can statistically evaluate whether TI forecast accuracies are equal to or worse than benchmark forecast accuracies (H$_0$: E(l_t) $\leq$ 0) or weather they are better (H$_1$: E(l_t) > 0). We reject H$_0$ when DM values are above 2.33, indicating statistically significant performance improvement at a 1% significance level.

3. Results

The repeated selection of a handful of TI hyperparameters—%B ($n = 58$), EMA ($s = 6$), ROC ($n = 49$), etc.—by the grid-search optimisation, points to the prevalence of real and identifiable behavioural biases across the Belgian DAM. Our results, presented in Tables 1 and 2 as well as Figures 3–5, demonstrate that TI features can help in the identification of these biases by adding explanatory power and significantly boosting the regression accuracies of linear, ensemble, and deep models. On average, per the results in Table 1, the *best TIs* reduce forecast RMSEs by 3.28%, MAEs by 3.32%, and increase PCCs by 1.96%. These empirical metric improvements are statistically significant at a 1% level, yielding DM values above 2.33, in 9/10 cases.

Summarising TI performances across the ML models: the ROC increased nine models' accuracies; EMA, 8; SMA, 7; MOM, 6; 'Series', 4; 'Signal', 4; 'Histogram', 4; MSD, 4; %B, 4; COPP, 4; TSI, 3; and Bandwidth, 1. Broadly, the Bandwidth was found to be the worst-performing TI, while the ROC and the EMA were found to be the best. In Sections 3.1–3.3, we further breakdown the performance impacts of TIs; determining optimal indicators for each of the three ML model groups.
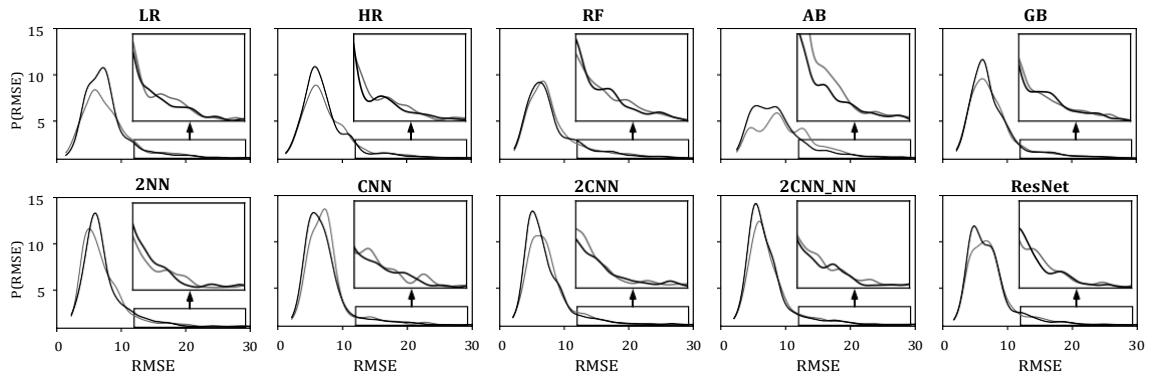
Figure 3. Scaled univariate kernel density estimates of daily model forecast RMSEs. **Black** curves present the distribution of forecast errors with the *best TI* features, while **grey** curves present benchmark model errors without TIs.
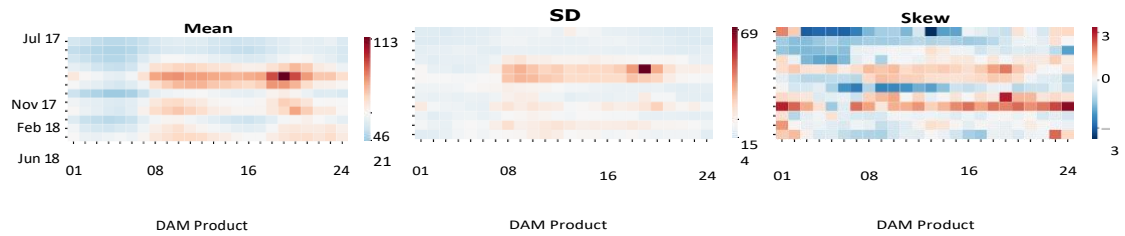

Figure 4. Heatmaps showing monthly means (**left**), standard deviations (**centre**), and skewness (**right**) of the test set prices (€/MWh). These metrics are calculated for each hour denominated DAM contract, to highlight both the periods and products exhibiting higher than average price, volatility, and skewness. Note that the colour bars are centred at each metrics mean value to facilitate our analysis.
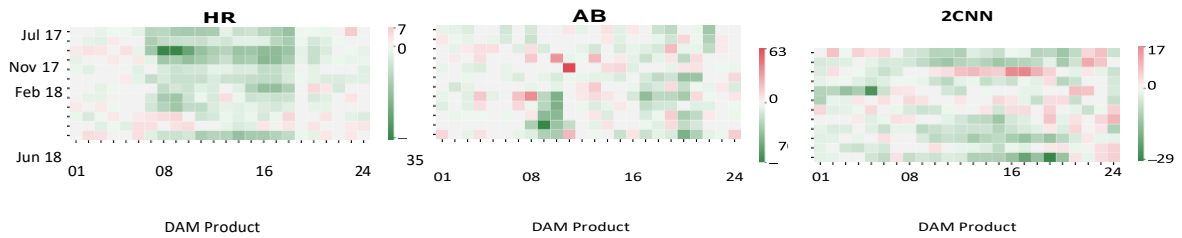

Figure 5. Heatmaps showing the %RMSE for each month of the test set and each hour denominated DAM contract, resulting from including EMA ($s = 2$), MOM ($n = 58$), and ROC ($n = 49$) features with HR, AB, and 2CNN models respectively. Negative values (**green**) indicate a reduction in forecast error, while positive values (**red**) indicate an increase in forecast error.

Table 1. Summary results, showing the three best performing—lowest **RMSE**—TIs for each model. Raw evaluation metrics—**RMSE**, **MAE**, and **PCC**—measure TI model performances. Percentages—**%RMSE**, **%MAE**, and **%PCC**—measure TI model performance improvements. Bold **DM** values represent statistically significant performance improvements at a 1% significance level. The following hyperparameter (**HP**) keys are used: * = (n_1 = 18, n_2 = 24, s = 18); ** = (s_1 = 2, s_2 = 26, s = 9); *** = (s_1 = 58, s_2 = 116, s = 9); **** = (n_1 = 58, n_2 = 74, s = 54).

		LR	HR	RF	AB	GB	2NN	CNN	2CNN	2CNN_NN	ResNet
Best TI	TI	%B	EMA	EMA	MOM	MOM	EMA	ROC	ROC	ROC	ROC
	HP	n = 58	s = 2	s = 6	n = 58	n = 58	s = 22	n = 49	n = 49	n = 9	n = 27
	RMSE	11.47	11.57	12.21	12.66	11.64	10.96	11.14	11.03	11.10	11.09
	%RMSE	4.49	4.50	1.66	5.42	3.74	4.09	2.73	2.39	2.36	1.38
	MAE	7.87	7.67	8.07	8.58	7.84	7.59	7.65	7.58	7.66	7.56
	%MAE	3.26	5.59	2.10	6.20	2.22	1.56	4.21	2.91	3.41	1.75
	PCC	0.79	0.79	0.76	0.74	0.79	0.81	0.80	0.81	0.81	0.80
	%PCC	2.92	2.75	1.08	3.79	2.80	1.85	1.40	1.19	0.95	0.82
	DM	**10.39**	**14.86**	1.74	**4.18**	3.68	**3.23**	**7.15**	**4.51**	**5.15**	**4.28**
Second-Best TI	TI	EMA	%B	MOM	ROC	%B	SMA	-	EMA	SMA	COPP
	HP	s = 2	n = 58	n = 58	n = 57	n = 54	n = 22	-	s = 18	n = 18	*
	RMSE	11.57	11.59	12.22	12.71	11.78	11.08	-	11.26	11.27	11.16
	%RMSE	3.73	4.35	1.57	5.01	2.66	3.00	-	0.37	0.91	0.77
	MAE	7.70	7.78	8.15	8.68	7.81	7.68	-	7.77	7.78	7.60
	%MAE	5.36	4.17	1.10	5.14	2.60	1.56	-	0.44	1.91	1.16
	PCC	0.79	0.79	0.77	0.73	0.78	0.81	-	0.80	0.80	0.80
	%PCC	2.55	2.89	1.57	3.23	1.71	1.38	-	0.06	0.01	0.40
	DM	**13.67**	**12.30**	1.23	**4.36**	1.61	2.08	-	0.16	1.36	**2.40**
Third-Best TI	TI	'Histogram' **	SMA n = 18	ROC n = 57	EMA s = 6	EMA s = 6	'Histogram' ***	-	-	COPP	-HP ****
	RMSE	11.62	11.66	12.23	12.82	11.79	11.12	-	-	11.28	-
	%RMSE	3.25	3.76	1.50	4.19	2.53	2.66	-	-	0.75	-
	MAE	7.88	7.79	8.18	8.78	7.84	7.68	-	-	7.78	-
	%MAE	3.11	4.02	0.77	4.06	2.20	1.51	-	-	1.89	-
	PCC	0.79	0.79	0.76	0.73	0.78	0.80	-	-	0.80	-
	%PCC	2.22	2.43	0.97	3.01	1.67	1.13	-	-	0.01	-
	DM	**11.92**	**16.02**	0.56	**3.92**	2.43	2.49	-	-	1.32	-

Table 2. Results highlighting differences between TI and benchmark model squared errors for targets inside the interquartile range (IQR) and outside the IQR (Tails). Bold values are used to highlight the highest scores.

			LR	HR	RF	AB	GB	2NN	CNN	2CNN	2CNN_NN	ResNet
		TI	%B	EMA	EMA	MOM	MOM	EMA	ROC	ROC	ROC	ROC
		HP	n = 58	s = 2	s = 6	n = 58	n = 58	s = 22	n = 49	n = 49	n = 9	n = 27
Best TI	IQR	%W	47.18	53.35	52.53	51.29	49.91	48.44	**56.45**	54.20	52.00	51.12
		$\bar{W}$	−24.76	−20.59	−29.67	**−66.34**	−27.34	−42.92	−30.00	−28.25	−32.05	−16.43
		%L	52.82	46.65	47.47	48.67	50.09	51.56	**43.55**	45.80	48.00	48.88
		$\bar{L}$	28.15	21.94	32.19	27.28	30.17	33.27	23.98	20.81	23.60	**13.31**
		$\bar{W} + \bar{L}$	3.40	1.35	2.52	**−39.06**	2.83	−9.65	−6.02	−7.44	−8.45	−3.12
	Tails	%W	55.79	**60.10**	53.17	53.88	53.21	51.65	54.55	54.38	53.37	55.03
		$\bar{W}$	−92.42	−67.83	−80.61	**−154.69**	−108.50	−129.77	−54.81	−54.21	−63.78	−37.31
		%L	44.21	**39.90**	46.83	46.07	46.79	48.35	45.45	45.62	46.63	44.97
		$\bar{L}$	43.91	29.69	67.52	118.70	63.10	91.44	49.99	50.09	55.09	**32.97**
		$\bar{W} + \bar{L}$	**−48.51**	−38.14	−13.08	−35.99	−45.41	−38.33	−4.82	−4.11	−8.69	−4.34

3.1. Best Performing TIs

Analysing the performances of linear models, on average linear models were the most successful at leveraging the added explanatory power TIs provide. All TIs except the Bandwidth, the COPP, and the ROC successfully reduced LR's RMSE, while all TIs except the Bandwidth reduced HR's RMSE. Out of the TIs that successfully improved the performances of linear models, we found that %B (n = 58) and EMA (s = 2) added the most explanatory power. Per the results in Table 1, across LR and HR models, %B (n = 58) on average reduced the RMSE and MAE by 4.42% and 3.72% respectively, and increased the PCC by 2.91%. EMA (s = 2) improved the same metrics on average by 4.13%, 5.48%,

and 2.65% respectively. Analysing %B and EMA linear model performances across different target groups, Table2highlights that both %B (n = 58) and EMA (s = 2) significantly reduced a majority of tail forecast errors, without, however, improving IQR forecasts. Specifically, %B (n = 58) and EMA (s = 2) yielded both negative tail $\bar{W} + \bar{L}$ and %Ws greater than 50%: 55.79% and 60.10%. Neither TI, however, yielded a negative $\bar{W} + \bar{L}$ for targets inside the IQR. Moreover, only EMA (s = 2) successfully reduced a majority (53.35%) of IQR forecast errors.

Examining the performances of ensemble models, in Table1, TIs recorded statistically significant reductions in forecast error with two out of three ensemble models. EMA (s = 6) was the only *best TI* to not yield a statistically significant performance improvement. It nonetheless reduced RF's RMSE and MAE by 1.66% and 2.10% respectively, and increased its PCC by 1.08%. Overall, MOM (n = 58), which significantly boosted both AB and GB performances, was observed to add the most explanatory power to ensemble models: MOM (n = 58) was the *best TI* for both AB and GB models, and *second-best TI* with RF. Per Table1, across ensemble models MOM (n = 58) on average reduced the RMSE and MAE by 3.58% and 3.17%, respectively, and increased the PCC by 2.72%. Furthermore, Table2highlights that MOM (n = 58) increased the accuracy of a majority of AB and GB tail, and AB IQR forecasts: yielding a negative $\bar{W} + \bar{L}$ and %Ws greater than 50%. MOM (n = 58), however, did not increase the accuracy of a majority of GB IQR forecasts: yielding a positive $\bar{W} + \bar{L}$ (2.83), as well as a %W below 50% (49.91%). Nonetheless, given the results in Tables1and2, we consider MOM (n = 58) to be our best ensemble model TI.

Finally, analysing the performances of deep models, Table1highlights that, when added to deep models containing CONV layers, the ROC overwhelmingly outperformed a majority of TIs. While other indicators, including the %B and the MOM, struggled to add explanatory power to the four CONV models, on average the ROC significantly reduced CONV models' RMSE and MAE by 2.22% and 3.07% respectively, and increased their PCC by 1.09%. Beyond CONV models, Table1 also highlights that the EMA performed best with the fully connected deep model. EMA (s = 22) significantly increased 2NN's predictive power: reducing its RMSE and MAE by 4.09% and 1.56% respectively, and increasing its PCC by 1.85%. Analysing deep models' TI performances across the target groups, Table2highlights that the ROC successfully improved the accuracy of a majority of CONV model IQR and tail forecasts. Unlike with linear and ensemble models, the best performing CONV model TI, the ROC, yielded both tail and IQR: $\bar{W} + \bar{L}$s below zero and %Ws greater than 50%. This was not the case with EMA (s = 22), however. EMA (s = 22) increased 2NN tail forecast accuracies, but not IQR accuracies. Overall, we consider ROC (n = 49) and EMA (s = 22) to be our best performing deep model TIs.

In addition to highlighting which TIs provide the most explanatory power to deep models, Table1 also affirms an inverse relationship between model depth and accuracy improvement. We observe reductions in *best TI* model improvement rates, %RMSE/%MAE/%PCC, with each additional CONV layer: from 2.73/4.21/1.85 (CNN), to 2.39/2.91/1.19 (2CNN), to finally 1.38/1.75/0.82 (ResNet). A gradual reduction in performance improvement occurs because stacking multiple CONV layers allows models to identify larger and more complex features, reducing the need for feature engineering.

3.2. Distribution of Errors

In Table1, we summarised TI model performances, highlighting the added explanatory power offered by TIs. Below, we set out to extend this analysis, and verify that the inclusion of TI features does not induce fat error tails. Fat error tails pose a significant risk to traders that rely on accurate model forecasts to make profitable investment decisions. Because traders are loss-averse, they often prefer models which forecast prices with a lower tail risk, even if this comes at the expense of a marginally higher model RMSE. With this in mind, in Figure3we present the distributions of model errors with and without *best TIs*.

Before analysing the plots of Figure3, let us describe the distribution shift that successful feature engineering should optimally result in. Ideally, the addition of features with added explanatory power

should induce a leftward shift of the error distribution—i.e., towards zero error—and reduce error tail mass. Successful prediction improvement should thus result in a higher central error peak without rousing fatter tails. We observe such error distribution shifts in the plots of Figure 3: *best TIs* (black) induce notably higher central error peaks, often shifted to the left, with thinner or unchanged tail masses. While in many cases tail masses remain unchanged or get slightly smaller, importantly none become fatter. Overall, Figure 3 reflects both Tables 1 and 2 results. Elaborating, except with RF,

we observe significant performance improvements in the error distributions. Further, we observe a greater leftward shift of the central peak with models whose IQR performance was boosted by the inclusion of TI features: notably CONV models. Figure 3 thus supports our findings from Section 3.1, offering evidence that TI features can yield more consistently accurate forecasts, without increasing the frequency of extreme errors. This is a significant finding because it demonstrates that the inclusion of TI features does not increase the tail risk.

3.3. Monthly Performance Improvements

In Section 3.1, we identified EMA ($s = 2$), MOM ($n = 58$), and ROC ($n = 49$) as being among our best performing linear, ensemble, and deep model TIs. Table 2 emphasized that each of these TIs, however, impacted forecast accuracies differently. While EMA ($s = 2$), for instance, significantly improved the accuracy of tail forecasts, it did little to improve the accuracy of IQR forecasts. ROC ($n = 49$), on the other hand, improved both IQR and tail forecast accuracies. Below, in an attempt to identify any prevailing patterns in forecast improvements, we visualise and analyse TI performance across the entire test set. Despite performance being model dependent, we aim to identify a link between forecast improvement and the use-cases of each TI.

Focusing on EMA ($s = 2$), overall we found that this TI increased the accuracy of 74.65% monthly HR forecasts, as visualised in Figure 5. Examining the distribution of EMA forecast improvements in the figure, we observe that a majority of large forecast improvements occurred during peak hours: between 08-20h. Moreover, by overlaying Figures 4 and 5, we discern that these generally correspond

to periods of above-average price and volatility. To potentially understand why the inclusion of EMA ($s = 2$) features results in such improvements, consider the information short-term EMAs convey, as well as their use-case. Short-term EMAs alongside lagged prices are used to identify short-term trend formations, reversals, etc. Practitioners especially use short-term EMAs during periods of increased volatility. This use-case may potentially explain why EMA ($s = 2$) increased peak-hour forecast accuracies the most.

Turning to MOM ($n = 58$), the MOM increased the accuracy of 64.24% of monthly AB forecasts. At a glance, Figure 5 improvements appear to be sporadically distributed; not clusterable by season or hour. Nevertheless, overlaying Figures 4 and 5 suggests that greater %RMSE improvement occurred during periods of average volatility, after prices had peaked or troughed. Outlining the potential reasons for this, recollect from Section 2.1.6 that the MOM, as an unbounded oscillating TI, is used to

forecast emerging trends. Specifically, MOM/price divergence and zero line crosses are thought to indicate bullish or bearish signals. For the MOM to provide additional explanatory power, however, volatility must be near the normal range (IQR) to facilitate the formation of regular and sustained price fluctuations. This is because extreme volatility can generate whipsaws—i.e., a MOM curve that crosses the zero line too frequently, with sporadic regularity. Overall, this potentially explains why MOM ($n = 58$) improved the accuracies of forecasts after prices had peaked or troughed, and after momentum had established longer-term trends.

Finally considering ROC ($n = 49$), we found that the ROC increased the accuracy of 71.53% monthly 2CNN forecasts. Unlike the EMA with HR, however, performance improvements were not divided broadly between peak and off-peak hours. Instead, as overlaying Figures 4 and 5 highlights, the performance improved more significantly during periods of roughly normal (IQR) volatility as prices peaked or troughed. An understanding of the ROC's use-cases allows us to potentially explain the reasons behind this improvement. Specifically, recall from Section 2.1.7 that the ROC is principally

used for the identification of overbought or oversold extremes that historically foreshadowed a trend reversal. Little information is captured by the ROC during periods of price consolidation and low volatility. Similarly, during periods of extreme volatility, when sporadic price changes ensue, previously informative overbought or oversold signals lose their predictive power. This potentially explains why ROC (n = 49) added more explanatory power during periods of rising/falling prices when volatility was near its IQR.

4. Conclusions and Discussion

This paper evaluated the explanatory power of TIs by examining whether the inclusion of TI features could boost forecasting accuracies of DAM prices. Overall, we have demonstrated that TIs can capture the residual impacts from traders' behavioural biases; resulting in statistically significant reductions in forecast errors with ML models. Our case study has identified four TIs well suited to forecasting Belgian DAM prices: the EMA, the %B, the MOM, and the ROC. More specifically, our results advise using %B (n = 58) and EMA (s = 2) with linear models, MOM (n = 58) with ensemble models, EMA (s = 22) with NNs, and ROC (n = 49) with CNNs. While we find evidence that TI performance is model dependent, the ROC and the EMA succeeded in reducing the RMSE and MAE of 90% and 80% of the ML models respectively.

As the first to use TI features for DAM forecasting, we hope that this paper spurs others to examine the potentials of technical feature engineering. While our research examined the explanatory power of some of the simplest and widely used TIs in finance, further research could examine the predictive power of more complex TIs. Moreover, given the unique qualities of the DAM, future research may succeed in engineering new TIs, which incorporate DAM volumes, order book data, etc. to optimally capture short-term energy traders' behavioural biases. Faced with growing DAM volatility, we believe that the inclusion of tailored TIs alongside fundamental features shall become increasingly important, even if forecasting models grow in depth.

Abbreviations

The following abbreviations are used in this manuscript:

AB	AdaBoost
BBAND(s)	Bollinger Band(s)
CNN	Neural Network with CONV
CONV	Convolutional Layer
COPP	Coppock Curve
DAM	Day-ahead Market
DM	Diebold-Mariano (test statistic)
EMA	Exponential Moving Average
FCL	Fully Connected Layer
GB	Gradient Boosting
HR	Huber Regression
LR	Linear Regression
MACD	Moving Average Convergence Divergence
MAE	Mean Absolute Error
ML	Machine Learning
MOM	Momentum

MSD	Moving Standard Deviation
NN	Neural Network with FCL
PCC	Pearson Correlation Coefficient
RF	Random Forest
RMSE	Root Mean Squared Error
ROC	Rate of Change
SMA	Simple Moving Average
TI(s)	Technical Indicator(s)
TSI	True Strength Index

Appendix A. Technical Analysis

Unlike overbought/oversold signals occurring in isolation, identification of a potential W-Bottom pattern can directly lead to a price target. Specifically, the W-Bottom pattern establishes a resistance band, which is broken by an eventual strong upward price move. Detailing the indicative markers of a W-Bottom pattern, the pattern occurs in a price down-trend when firstly a security becomes oversold. Subsequently, the price moves towards the SMA, establishing a resistance level, before continuing its previous downward-trend, moving towards the $BBAND^-$. This second price decrease is marked by two

identifiers: the price does not breach the $BBAND^-$ (%B > 0), yet it establishes a new low. The W-Pattern is confirmed by a strong price move towards the resistance band, leading to an eventual breakout.
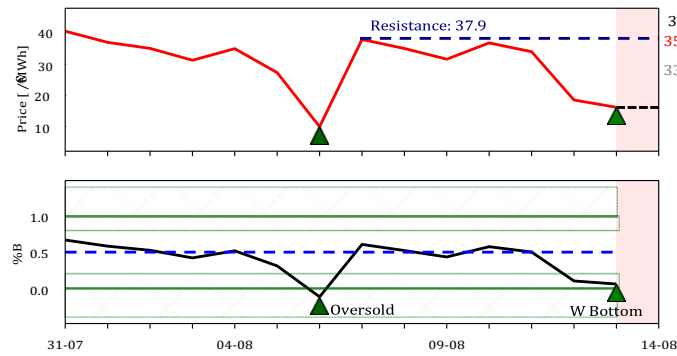

Figure A1. Single hour prediction of the 14/08/17 08h price, generated using an HR model, demonstrating the added explanatory power offered by TIs. The red line in the top graph shows raw prices, while the black line in the bottom graph shows the %B (n = 26). Arrows and horizontal lines mark the identifiers of a W-Bottom pattern that foreshadows a strong price move towards the resistance line. The dashed lines, appearing in the red band, correspond to the benchmark forecast (grey: 33.1 €/MWh), TI forecast (black: 36.1 €/MWh), and price target (red: 35.9 €/MWh).

Examining Figure A1, we can identify all but one of the markers necessary for a potential W-Bottom pattern. Elaborating, we observe oversold prices, followed by a move to the SMA (%B = 0.5), followed by a significant price drop that remains above %B = 0. While we do not see a new low emerge because of the DAM's sideways trend, the failure of the %B to breach 0 is indicative of a trend reversal. Overall, the model is thus able to identify the trend reversal better when %B features are included, predicting a next-day price 4.7% from the resistance band and 0.6% from the target. This prediction proves to be considerably more accurate than the benchmark prediction; the benchmark model forecasts a price 12.7% from the resistance band and 7.8% from the target.

Appendix B. Hyperparameters

Table A1 presents grid-search selected optimal model hyperparameters. Note that we use scikit-learn [39] to run linear and ensemble models, and Keras [40] to run deep models. As a result, unless otherwise specified in Table A1, either scikit's, or Keras' default model hyperparameters are used.

Table A1. Grid-search selected model hyperparameters. Benchmark model performance is optimised to select these. For a full specification and description of the models' hyperparameters, we refer interested readers to [39,40].

	Hyperparameters
LR	-
HR	epsilon:1.35
RF	n_estimators: 100
AB	n_estimators: 100, loss: square, learning rate: 0.1
GB	loss: huber

2NN	$neuron_1$: 500, $neuron_2$: 250, learning rate: 0.001
CNN	kernel: (1, 3), filter: 16, CONV layer: 1, learning rate: 0.001, dropout: 0.25
2CNN	kernel: (2, 3), filter: 13, CONV layer: 2, learning rate: 0.001, dropout: 0.25
2CNN_NN	kernel: (1, 3), filter: 32, CONV layer: 2, neuron: 123, learning rate: 0.001
ResNet	kernel: (3, 3), filter: 23, CONV layer: 7, learning rate: 0.001