



Towards the Prediction of Electricity Prices at the Intraday Market Using Shallow and Deep-Learning Methods

Christoph Scholz^(✉), Malte Lehna, Katharina Brauns, and André Baier

Fraunhofer IEE, 34119 Kassel, Germany
{christoph.scholz,malte.lehna,katharina.brauns,
andre.baier}@iee.fraunhofer.de

Abstract. The percentage of renewable energies (RE) within power generation in Germany has increased significantly since 2010 from 16.6% to 42.9% in 2019 which led to a larger variability in the electricity prices. In particular, generation from wind and photovoltaics induces high volatility, is difficult to forecast and challenging to plan. To counter this variability, the continuous intraday market at the *EPEX SPOT* offers the possibility to trade energy in a short-term perspective, and enables the adjustment of earlier trading errors. In this context, appropriate price forecasts are important to improve the trading decisions on the energy market. Therefore, we present and analyse in this paper a novel approach for the prediction of the energy price for the continuous intraday market at the *EPEX SPOT*. To model the continuous intraday price, we introduce a semi-continuous framework based on a rolling window approach. For the prediction task we utilise shallow learning techniques and present a *LSTM*-based deep learning architecture. All approaches are compared against two baseline methods which are simply current intraday prices at different aggregation levels. We show that our novel approaches significantly outperform the considered baseline models. In addition to the general results, we further present an extension in form of a multi-step ahead forecast.

1 Introduction

Since the liberalization of the European electricity markets in the 1990s, the primary trading location for electricity commodities are the national day-ahead spot markets. However, with the increased electricity production through renewable energies, market prices have become more volatile. As result a rising demand for continuous trading on intraday spot markets emerged in recent years.¹ On the intraday spot markets, participants are able to account for rapid changes

¹ http://static.epexspot.com/document/38579/Epex_TradingBrochure_180129_Web.pdf.

C. Scholz and M. Lehna—Both authors contributed equally to this work.

in demand and supply and integrate these changes in the pricing on a short-term notice. However, in current research the continuous intraday market is still under-represented, especially in the electricity price forecast (EPF). Even though various research has been conducted on different day-ahead markets, as reviewed by Gurtler and Paulsen [6], to the best of our knowledge only a small amount of papers [1, 4, 9, 13, 14, 18] have been published for the price forecast on the intraday market. Furthermore, all these papers only focus on forecasting the final price on the continuous intraday market. Consequently, this paper is intended to contribute new insights for EPF into the intraday market. We propose a novel (semi-continuous) forecast framework in order to capture a more consecutive representation of the intraday market. In addition, we incorporate previous research proposals by combining several external factors in our forecast study. Finally, different machine learning methods are presented and analyzed regarding their predictive power.

1.1 Research Environment

As research environment, we chose the intraday spot market of Germany from the year 2018, as we had enough data available to investigate the continuous nature of the spot price and realize the respective forecast. On the market, the electricity for the same day delivery is traded in two different products, which are the one-hour and quarterly-hour products, and are based on their delivery length. For each of the one-hour (or quarterly-hour) products of the day, the market participants trade in different magnitudes of electricity feed-in through an continuous orderbook system. However, in the continuous intraday market the fundamental challenge of the EPF is the short time horizon in which the products are traded. Between the opening of the market and the termination by the delivery, only a small time interval is available for trading. In terms of hourly intervals, the market opening occurs on the previous day at 15:00 while the trade for the quarterly intervals begins at 16:00. For both, the products are traded continuously up until 30 min before the delivery on the whole market and up to 5 min within the control areas of the distributors. In addition, the majority of transactions take place in the last hours before delivery. In Fig. 1 this effect is visualized, where one can clearly see the increase in trading volume over time. Moreover, note the two drops in the graph which correspond to the described 30 min and 5 min mark. Due the fact that the trading volume gets stronger towards the end, we focus in our scenario on the last four hours till delivery time.

1.2 Contribution

Framework: Given the specific structure of the intraday market, we introduce a semi-continuous framework to enable the forecasting of the continuous intraday prices. However, instead of only focusing on one value of each product (as done in previous papers like [9, 14, 18] for the German market), we propose to forecast the electricity price development at regular intervals. This allows to forecast

multiple observations per product and give a more precise description of the price development. Our intention behind this short-term perspective is the ability to forecast the price movement, while the product is still traded on the market. In comparison to other EPF literature this approach is unique, due to the fact that until now researchers only predicted one value per product.

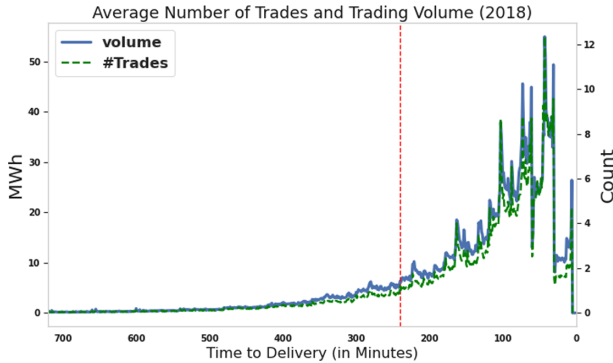


Fig. 1. Average number of both trades and traded volume for time t . The x -axis denotes the remaining time t (in minutes) till the delivery of the product. The left y -axis represent the average number of traded volume at time t , the right y -axis the average number of trades at time t . The red dotted line at minute 240 is the starting time of our forecast scenario.

Forecasting Approaches: One major consequence following from the structure of the research environment is that the implementation of time series approaches such as ARIMA and VAR models proves to be difficult. Due to their reliance on previous observations, it would be necessary to determine the auto-regressive order structure and provide sufficient amount of data to estimate the models for each individual product. As a result, in this paper we disregard the usage of basic time series model and instead analyze three different machine learning techniques for the EPF task. As model candidates, we decided to implement two shallow learning models as well as one deep learning model for the comparison analysis. In terms of the shallow learning models, the *Random Forests* [2] as a simple model and the *XG-Boost* [3] as a more advanced model were selected. However, it is important to note that while both approaches can be applied on a regression setting, they are not designed to capture time related dependencies between the samples i.e. auto-correlation or auto-regressive structures. Nevertheless, due to their flexibility and their robust behavior in terms of noise variables, we employ the two approaches. On the other hand, we decided to employ a *LSTM*-based deep neural network architecture for the prediction of the spot price. This network is well known to detect and use time dependencies for the prediction task. All models are further discussed in Sect. 4.2.

Outline: In general, our contribution in this paper can be summarized as follows:

1. We present the first work forecasting the electricity price development of individual products on the continuous intraday market.
2. We utilize three different machine learning approaches, two shallow learning and one *LSTM*-based deep learning approach and compare the predictive power with state-of-the-art baseline models.
3. We show that our presented novel approaches significantly outperform baseline models by at least 10%.
4. We perform and discuss the performance of a multi-step ahead forecast.

2 Related Work

Within the literature, intraday spot price receives more and more attention in recent years. Kiesel and Paraschiv [10] examine the bidding behaviour of participants on the German intraday market based on 15 min products and analyze the different influencing factors. In their paper, they especially emphasise the importance of wind and photovoltaic forecast errors as one major impact for the trading behaviour. Similar results were published by Ziel [21], who analyzed the effect of wind and solar forecasting errors. Further analyses on the German intraday market were published [8, 16], however with a different research focus. While Pape et al. [16] focused their research on the fundamentals of the electricity production and their influence on the intraday spot price, Kath [8] analyzed the effect of cross-border trade of electricity. Next to the German intraday market, other research was conducted on the Iberian electricity market [1, 4, 13] as well as the Turkish intraday market [15]. On the Iberian spot market MIBEL, the first advances in terms of intraday EPF have been made. While Monteiro et al. [13] use a MLP neural network to forecast the intraday prices, Andrade et al. [1] conduct a probabilistic forecasting approach to further analyze the influence of external data. Further research was published by Oksuz and Ugurlu [15], who examined the performance of different neural networks in the EPF in comparison to regression and LASSO techniques on the Turkish intraday market. In terms of the German, the intraday trade was primarily covered by three papers [9, 14, 18]. Kath and Ziel [9] predicted both the intraday continuous and the intraday call-auction prices of German quarterly hour deliveries based on an elastic net regression. Furthermore, they proposed a trading strategy based on their forecast results to realize possible profits. A different approach was proposed by Uniejewski et al. [18] which analyzed the German intraday spot price in form of the closure price of the *EPEX SPOT* ID3 index. For their forecast they employed a multivariate elastic net regression model to further conduct a variable selection and proposed a simple trading strategy to realize possible gains. With a very similar research setting Narajewski and Ziel [14] also analyzed the *EPEX SPOT* ID3 index through a multivariate elastic net regression model. Their difference to Uniejewski et al. [18] is that they included both the one hour and quarterly-hour in their analysis and focused on the variable selection. However, all previous papers disregard the continuous structure of the intraday spot

price market as discussed in Sect. 1.1 and 1.2. While the forecast of the MIBEL intraday market covered at least multiple trading intervals [1, 13], the German intraday forecast only consider one observation per product [9, 14, 18]. While this approach might be interesting for a long term perspective (e.g. on the day-ahead market), it does not necessarily reflect the needs on the intraday market. Therefore, this paper introduces a new perspective on the intraday market with the aim to establish a more short-term forecasting horizon.

3 Data Used for the Continuous Intraday Price Forecast

In order to compare different forecasting approaches, we chose the German hourly electricity intraday spot price as endogenous variable. As time horizon, we analyzed the first half of 2018 and thus constructed a data set of the German intraday price from 02.01.2018 till the 30.06.2018. For the model comparison and their predictive power, we decided to use the June of 2018 as test interval and the previous five months as training interval.² For each day a total of 24 hourly products were traded which results in a total of 3576 products in the training sample and 720 products within the forecasting sample.³ To analyze the predictive power in short-term perspective, we chose as dependent variable the volume weighted 15 min averaged trading results of the one hour German continuous intraday market. Through the averaging process, further described in Sect. 4.1, a total of 40 observations per product were created, thus giving a more detailed process of the price development. In order, to predict the dependent variable, multiple features were included in the forecasting process. These variables are the following:

1. Past transaction results which consist of the previous price as well as trading volume and number of trades.
2. Prediction error of the wind forecasts between the day-ahead forecast and the current intraday wind forecast.
3. *EPEX SPOT M7* orderbook data as well as grid frequency data.
4. Categorical variables describe the weekday, hour of the product and the time till completion for each step.

Due to the fact that some variables have not been frequently used in literature and specific transformations were partly necessary, we shortly discuss the variables in detail.

² Note that it is planned for future work to consider further training and testing periods to investigate the quality of the models.

³ Note that two days were excluded from the observation (01.01.2018 & 25.03.2018). The first was excluded due to missing training data of 2017, while the second was generally missing data at the respective day.

3.1 Transaction Data

The transaction data (for 2018) of the continuous intraday market were bought from the *EPEX SPOT*⁴. The data set contains price and volume information about all executed trades. In addition it also contains information about the execution time (at a one minute resolution) and information about the corresponding product. For the price variable, several aggregation steps were necessary. First, the price was centered through a subtraction of the product's day-ahead price and, due to extreme outliers in 2018, scaled by the standard deviation and a constant c .⁵ Second, the price was averaged through a volume weighted mean with two different resolutions. For the endogenous variable an interval of 15 min was used for the averaging. In terms of the exogenous variable, the previous price information were aggregated in two parts. On the one hand, the previous 15 min mean prices were used as look-back variable with a time horizon of 4 h. However, additional research showed that a more detailed description of the 15 min prior to the prediction offered further insights for the forecast. Thus, we included a one minute volume weighted mean for the last 15 min in order to capture the short-term trend. Next to the price variable, we further included the trade volume and count (i.e. number of trades) variable. Due to the fact that we wanted to model a mid-range development with the variables, we chose to aggregate them similar to the first spot price variable. Hence, the volume and count variable were aggregated 15 min mean values recorded for the last four hours.

3.2 EPEX SPOT M7 Orderbook Data

The *EPEX SPOT* M7 orderbook data contains all historic and anonymous orders submitted to the continuous intraday market of the *EPEX SPOT*. This allows a more precise description of the current market situation, as the data also consist of additional information such as available trading volume and the current bid-ask spread for each point in time. In general the information contained in the orderbook data can be divided in ex-ante and ex-post information, see Martin et al. in [12]. Ex-ante information like price and volume are available at the current time of order creation. Ex-post information like the execution price depends on market developments and is not yet available at the time of order creation.⁶ In our experiments we combined the information to calculate the current buy and sell prices for 10MWh and 50 MWh which we also included in our models. In Fig. 2 we plotted as an example the current price development as well as the sell and buy price for 10MWh. The difference between the buy and sell price is also known as the bid-ask spread.

⁴ <https://www.eex-group.com/eexg/companies/epex-spot>.

⁵ While many researchers further transform the spot price, e.g. Uniejewski and Weron [19], we were not able to detect improvements in our estimation. Instead, we opted for the simple scaling through a constant c so that 99.7% of the data was within the interval $[-1, 1]$.

⁶ The detailed variables in the orderbook data are displayed in the Appendix A based on Martin et al. [12].

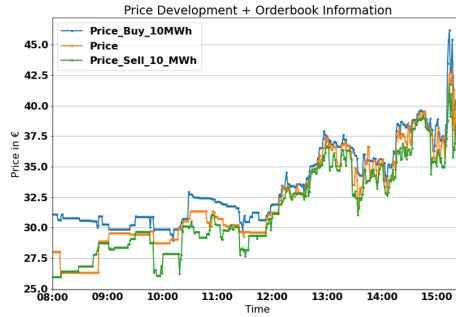


Fig. 2. Price development of the 2018-01-18 16:00:00 product with the corresponding buy and sell prices for 10 MWh. The x -axis represents the time t of the day, the y -axis the corresponding price at time t .

3.3 Day Ahead and Short-Term Forecasting of Wind Energy

In order to balance their own balancing group, managers often have to compensate for errors in the day ahead forecast (wind and photovoltaics) on the intraday market. Therefore, in our experiments we use the difference between the actual and the day-ahead wind-power production forecast of Germany as feature.⁷ For this purpose, we include the day-ahead and short-term forecast generated by the Fraunhofer IEE into our forecasting framework. The respective forecasts have a resolution of 15 min and are generated every 15 min, and forecast the wind-power production in Germany. For more details about the generation of the day ahead and short-term wind-power production forecasts we refer to Wessel et al. [20]. In our experiments we included all forecast errors (i.e. the difference between the actual and the day-ahead forecast) of the past four hours in our models.

3.4 Grid Frequency

In general, the European interconnected grid requires a grid frequency 50 Hz. Usually, only small deviations from the grid frequency occur, so that only minimal countermeasures by the grid operators are necessary to balance the frequency. In order to intervene, balancing energy is often used to compensate for potential imbalances. Since the use of balancing power is usually associated with high costs, it is possible that changes in the frequency are factored in the current electricity price. Therefore, in our forecast approach we also include the current network frequency, where data is provided by the French Transmission System Operator (TSO) RTE.⁸

⁷ At the time of writing, we had no adequate photovoltaic feed-in forecast available.

⁸ https://clients.rte-france.com/lang/an/visiteurs/vie/vie_frequence.jsp.

4 Price Forecasting Methodology

4.1 A New Framework for Short-Term Energy Price Prediction

In order to forecast the intraday electricity price in a short-term perspective it is necessary to transform the continuous variables to a semi-continuous forecast framework. Consequently, we developed the following forecast framework for all one hour products. In our setting we use historical data to predict the next 15 min volume-weighted average prices as illustrated in Fig. 3.⁹ In our regular forecast process, for each product the “forecast-time” was limited to an interval between 4 h until 45 min before delivery time (i.e. the termination) of the product. By reason of their low trading volume (as seen in Fig. 1), observations prior to the 4 h limit were discarded and only used as features. Further, due to the fact that in the last 30 min trading is only allowed within the control areas of the distributors, we excluded this time period as well. Thus, we performed the last forecast 45 min before deliver time in order to ensure that there are no overlaps to the last half hour. As a second step, we aggregate the continuous intraday electricity price with a rolling window. The window generates in 5 min steps volume weighted averages of the next four 15 min blocks. The decision to aggregate the spot price through the volume weighted mean is twofold. First, we ensure that enough observations are summarized, given that especially the early observations are relatively sparse. Moreover, the volume averaging further establishes more stable observations that are less affected by outliers i.e. high prices with a small trading volume. Thus, with 5 min steps in the 3.15h interval we receive a total of 40 observations per product.

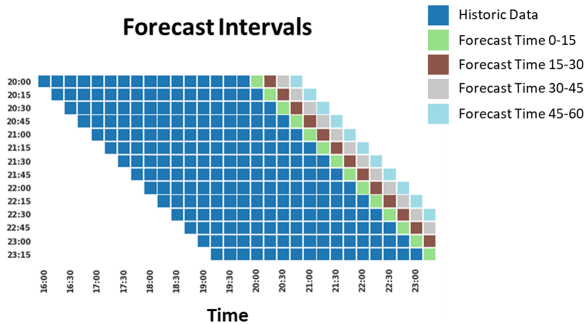


Fig. 3. Forecasting interval of a 00:00 product. The x -axis denotes the time, the y -axis the model run of the forecast (i.e. when the forecast is performed). The first forecast is done four hours before termination (at 20:00), while the last forecast is done 45 min before termination (at 23:15). The dark blue squares symbolize the time horizon of the data points that are used to perform the price forecast for the time symbolized by the remaining squares.

⁹ In addition, in Sect. 5.2 we extend the prediction interval to a length of four 15 min steps which we further display in Fig. 3 as well.

The third step of the research framework is the inclusion of the features. As already stated in the Sect. 3, the features were included in the analysis through three different groups. As a first group, we consider the price development of the product in the last 15 min with a one minute resolution, to capture the short-term information. This inclusion is based on the idea that last minutes changes in the trading behaviour would translate to the next prediction interval. However, we further wanted to describe a mid-range trend of the spot price. Thus, the second group of variables was included in the framework which consisted of a 4h look-back in 15 min steps. Next to the volume weighted price, this group of variables consisted of the count, volume and orderbook data as well as wind error variable. Lastly, as third group of variables, we included three categorical variables that denoted the hour and weekday of the product as well as the remaining time until the product would end. These variables were included to capture structural dependencies that were valid across all products. After the realization of the described framework, the products were later divided into a training and testing data set, as described in the Sect. 3. Under consideration that the rolling window results in 40 observations per product, we have a total of 28800 observations in the test sample and 143040 observations in the training data set.

4.2 Machine Learning Models

In the following, we briefly summarize the characteristics of the machine learning methods that we applied to the continuous intraday price forecast.

Random Forests (RF). The Random Forests algorithm, introduced by Breiman [2], is an ensemble method that uses a set of (weak) decision trees to build a strong regression model. In the tree building process each tree is trained on bootstrapped samples of the training data. The split of each node is selected on a random subset of d input features.¹⁰ The final prediction is the average of all trees individual predictions.

XG-Boost (XGB). The *Extreme Gradient Boosting* algorithm was first introduced by Chen and Guestrin [3] and is a parallel tree boosting that is designed to be “efficient, flexible and portable”.¹¹ In general, the algorithm is based on the *Gradient Boosting* algorithm which is similar to *RF* an ensemble method that combines multiple weak learners to a stronger model. In comparison, the *XG-Boost*-algorithm further improves the framework on the one hand from an algorithmic perspective, by enhancing the regularization, weighted quantile sketching and sparsity-aware splitting. On the other hand, the system design was improved through parallelization, distributed tree learning and out-of-core computation. Thus, through the combined improvements of both algorithmic and system design, the *XGB* is the next evolutionary step of the *Gradient Boosting*.

¹⁰ For regression tasks typically $m/6$ features are selected at random.

¹¹ <https://xgboost.readthedocs.io/en/latest/>.

Long Short-Term Memory (LSTM). The *Long Short-Term Memory* was introduced by Hochreiter and Schmidhuber [7] in 1997. A *LSTM* is a special type of *Recurrent Neural Network (RNN)* [17] that can capture long-term dependencies. In contrast to a *RNN*, in practice, a *LSTM* is able to deal with the vanishing and exploding gradient problem. A *LSTM* consists of the four components: memory cell, input gate, forget gate and output gate that interact with each other. Each component is represented by a neural network, where the input gate controls the degree to which level a new information is stored in the memory cell. The forget gate controls the degree to which level an information is kept in the memory cell, and the output gate controls the degree to which level an information is used within the activation function.

5 Evaluation

5.1 Experimental Setting

Given the research data and the respective framework, the three machine learning models *RF*, *XGB* and *LSTM* were applied and optimized for their hyper parameters. Due to the fact that all three models inherit some randomness in their estimation, we aggregate multiple predictions, in order to receive more stable results. The aggregation is implemented by re-estimating the optimized model ten times, forecast based on the test data set and thereafter take the median value of the ten predictions. While the general experimental setting is enforced on all three models, there are still some differences in the structure and the hyper parameter optimization. Accordingly, we shortly summarize the implementation of all three models and further present the baseline as well.

RF and XGB: In terms of the *Random Forests* and the *XG-Boost*, the data was implemented in a regression setting. Because both approaches are in general not able to distinguish time structured data, no further arrangement was necessary. For the hyper parameter optimization of both models, we decided to implement a randomized *Cross-Validation* to ensure that all training data was included in the selection process. Thereafter, the models were repeatedly re-estimated and the median forecast was computed.

LSTM: In the architecture of the *LSTM* model (see Fig. 6 in the Appendix), the features were integrated in three different input gates. For both the 4h mid-range data as well as the 15 min short-term data, the respective features were integrated through separate *LSTM* layers followed by an individual dense layer. In terms of the categorical variable, we decided against the frequently used one-hot encoding and instead chose to embed the variables, based on the results of Guo and Berkahn [5]. Thus, each variable was first inserted into an embedding layer and then combined with the other two through one dense layer. As a next step, the output of the embedded layers was combined with the two *LSTM* dense layer outputs. Thereafter, the concatenated data is fed through the final two dense layers that summarize the data. For all dense layers, we chose the *Leaky ReLU* activation function with the exception of the last dense layer, were only a linear

activation function was implemented. For the hyper parameter optimization, we decided to use the *Hyperband* approach [11], due to its ability to select the best hyper parameters while managing the resources effectively. Given the results of the *Hyperband* optimization, the forecast median was conducted as mentioned before in the experimental setting.

Baseline Models: To examine the forecast quality of our approaches two different and (in practice) commonly used baselines are proposed to guarantee an absolute necessity and minimum performance threshold. The first baseline is the previous observation i.e. the volume weighted mean of the last 15 min which we denote as BL_{15} . As second baseline, we include the last one minute price of the product before the prediction interval, denoted as BL_1 .¹² The reason for the decision of two performance thresholds is based on the idea that the models have to compete against both current impulses (BL_1) as well as a more robust price development (BL_{15}) of the product.

5.2 Results and Discussion

Single Step Prediction: Given the aggregated median forecast results, we are now able to analyze and discuss the prediction quality of the three models. For this purpose we primarily evaluate the forecast values based on the *RMSE* error metric. As one can see in Table 1, all three models outperform the BL_{15} baseline by more than 14% as well as the BL_1 baseline by more than 10%. In terms of the overall performance, the *LSTM* shows the best results. However, the results between the *XG-Boost* and the *LSTM* models are relative close which is surprising considering that the *XGB* does not incorporate any time series relationship. Furthermore, when consulting the standard deviation in Table 1 one can see that both shallow learning models are more stable in their prediction. Next, we analyze the performance of the forecast models within the specific time intervals i.e. we evaluate the forecast quality with respect (to the remaining) time till the product’s delivery time. The result are visualized in Fig. 4. In this

Table 1. RMSE Results of the respective models and their percentage change in comparison to BL_{15} and BL_1 . In the last column we denote the standard deviation of the *RMSE* based on the 10 forecast runs prior to aggregation.

Model	<i>RMSE</i>	$\Delta\% BL_{15}$	$\Delta\% BL_1$	Standard deviation
BL_{15}	2.3214	0	-3.78	
BL_1	2.2370	3.64	0	
<i>Random Forests</i>	1.9957	14.03	10.78	0.00145
<i>XG-Boost</i>	1.9461	16.17	13.00	0.00766
<i>LSTM</i>	1.9422	16.34	13.18	0.01517

¹² As example, for the prediction interval 20:00-20:15 the price of 19.59 is taken as BL_1 baseline and the volume weighted mean of 19:45-20:00 as BL_{15} .

context, some interesting results are visible. First of all, we are able to see that a large proportion of the $RMSE$ errors are induced through the forecast results at the end of the product. While the baseline models BL_{15} and BL_1 jump up to 5.099 and 4.980 in the last forecast step, the machine learning models achieve significant lower values with 3.942 (RF), 3.999 (XGB) and 4.394 ($LSTM$). The most probable explanation for the high rise of the $RMSE$ might be the increase in trading volume as seen in Fig. 1. Furthermore, one can see that all models are frequently able to beat the baseline, with especially good performances at minute 210, 95 and in case of the $LSTM$ also minute 190. It can therefore be assumed that the models will in many cases increase the performance of trading at the intraday market. Lastly, the direct comparison between the $LSTM$ and the XGB in Fig. 4 reveals that the $LSTM$ is generally performing better in the interval between minute 250 and 100, while XGB is showing better results in the last observations. The implication arising from the different performances hints that in the beginning a time series relation might drive the intraday spot price. The advantage of the $LSTM$ is later lost, especially in the last prediction step, thus one might increase the overall performance by combining the two models.

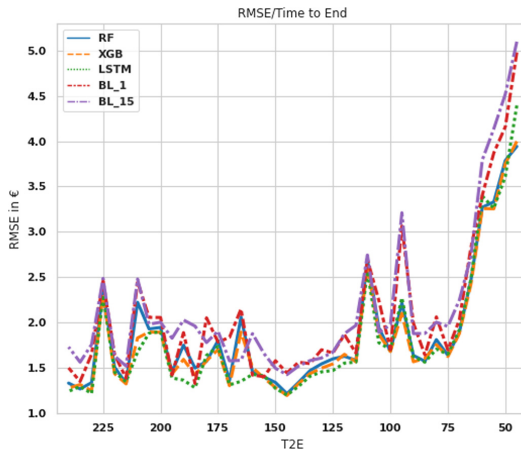


Fig. 4. Display of the $RMSE$ values for different points of time till the end of the product. The x -axis is denoted in minutes till delivery, while the y -axis shows the $RMSE$ in Euro.

The Multi-Step Forecast Extension: With the previous forecast success, we further want to present one possible extension of the forecasting framework. In the prior analysis, we were only interested in a one-step ahead prediction which is most useful for short-term trading. However, for (automated) trading it might be

advantageous to increase the forecast horizon in order to model the price trend of the product. Thus, we extended the forecast horizon by three additional 15 min steps. Note that our total forecast interval still ends 30 min before the delivery. This results in fewer forecast intervals in the last steps.

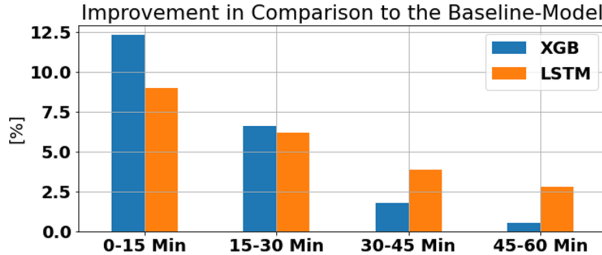


Fig. 5. Bar plot of $RMSE$ values for different forecast horizons.

Implementation: For the analysis we applied the two best performing models (XGB and $LSTM$).¹³ Therefore, in terms of the $XG-Boost$ we used a Multi-Output-Regressor structure which basically calculates for each of the four prediction steps an individual model. For the $LSTM$ we were able to implement a different approach, because the model is able to predict multiple steps ahead. Hence, a repetition vector prior to the dense layers was embedded and the last two dense layers were enhanced by a time distributed layer.^{14,15} Under consideration that the multi-step prediction is only seen as extension, we abstained from a re-optimization of the hyper parameters. Instead we only re-estimate the two models with our median aggregation approach.

Results: With the above-mentioned model-adaptation, the following results are achieved, as displayed in Fig. 5 and Table 2. Overall, both models are still able to outperform the BL_1 -baseline in the multi-step ahead forecast. Nevertheless one can observe a weakening of the predictive power the larger the forecasting horizon is. Furthermore, in comparison to Table 1 it is interesting to see that both models perform worse in the first step prediction. For the $LSTM$ model, this decline might be induced by the fact that multiple time steps must be optimized. On the other hand, in regard to the second till fourth interval, the $LSTM$ is showing a better prediction performance in comparison to the XGB . Thus, it might be interesting for future work to identify additional mid-term influencing factors to increase the predictability of the $LSTM$.

¹³ Since the RF had similar values to the XGB we skip the analysis of the RF .

¹⁴ https://keras.io/api/layers/recurrent_layers/time_distributed/.

¹⁵ The three input gates are kept in the original structure.

Table 2. Multi-step RMSE Results of the *XGB* and *LSTM* model as well as the BL_1 .

<i>RMSE</i>	00:00–15:00 min	15:00–30:00 min	30:00–45:00 min	45:00–60:00 min
BL_1	2.2370	3.3569	4.0416	4.7050
<i>LSTM</i>	2.0363	3.1496	3.8851	4.5730
<i>XG-Boost</i>	1.9609	3.1356	3.9687	4.6785

5.3 Future Work

In order to further increase the performance of our algorithms, additional influencing factors of the intraday market must be analyzed in depth and included into the machine learning models. The order *EPEX SPOT* book data, for example, offer high potential here. In this context an interesting research question is which additional features can be gained from the orderbook data to further increase the predictive power. Furthermore, factors such as the forecast error of the photovoltaic feed-in and weather forecasts could be integrated into the models. In addition to more input data, an extension of the forecasting periods should be considered as well. While this paper only examined one month of 2018 it could be of interest, to what extent forecast behaviour is changing with different months and seasons. In the same context, analysis should be conducted on the length of the training data, since it can be assumed that the trading behaviour of (automated) traders changes regularly and “older” data does not adequately reflect the current state of the market. Here, additional fundamental analyses are necessary to obtain meaningful statements. Finally, another aspect is the further development of the proposed *LSTM* model. By integrating *CNN-LSTMs*, new (cnn) features may be generated to obtain better results.

6 Conclusion

In this article we presented the first approach to forecast the electricity price development of individual products at the continuous intraday market of the *EPEX SPOT*. We evaluated the predictive power of two shallow learning algorithm (*XG-Boost* and *Random Forests*) as well as one *LSTM*-based deep learning architecture, complemented by comparing the results with two state-of-the-art baseline models. We show that all considered machine learning models perform significantly better than the baseline models. Our new developed *LSTM*-based

model showed the best performance, closely followed by the *XG-Boost*, which had similar results. The remarkable performance of the *XG-Boost* was unexpected, since the *XG-Boost* is not explicitly designed to detect relationships in time series data (in contrast to a *LSTM*). Furthermore we also performed and analyzed a multi-step time series forecast, where we forecast not only the next time step, but the next four. In this context we showed that the quality of the forecast significantly decreases with the longer forecast horizon, however the performance of the baselines was beaten nonetheless.

Acknowledgement. This work was supported as Fraunhofer Cluster of Excellence Integrated Energy Systems *CINES*.

Appendix A Orderbook Data

(See Table 3)

Table 3. Historic ex-ante and ex-post information available in the M7 orderbook data of the of the continuous intraday market, c.f. [12]. The term “delivery date” is used in this paper equivalently to the term product, and means the time of the delivery start of the corresponding product. The “start validity date” is the time the submitted order is valid from, and the ‘end validity date’ the time the order is no longer valid. The flag “active order” symbolizes whether an order is active or deactivated. The variable ‘side’ indicates, if it is a buy or a sell-order. The variables price and volume specify the offered price and volume, in contrast to the ‘execution price’ and ‘execution volume’ that define final the price and volume of the respective trade. In this paper we refer to the ‘execution price’ also with intraday spot price.

Ex-ante	Ex-post
Delivery date	Is executed (yes/no)
Start validity date	End validity date
Active order (Yes/No)	Canceling date
Side (buy/sell)	Execution price
Price	Execution volume
Volume	
ID	

Appendix B LSTM Architecture



Fig. 6. Structure of the *LSTM* model. On the left side, the three embedding layers process the categorical input. In the middle branch, the look-back variables of the last 4h are feed into the *LSTM* layer. On the right side, the last 15 min are also implemented into the network.

Appendix C Model Hyper Parameters

In this section we display the hyper parameters for the *RF* in Table 4, the *XGB* in Table 5 and the *LSTM* in Table 6.

Table 4. *RF* hyper parameter

Hyper parameter <i>XGB</i>	
N Estimators	1500
Min Samples Split	5
Min Samples Leaf	2
Max Depth	20

Table 5. *XGB* hyper parameter

Hyper parameter <i>XGB</i>	
Colsample by Tree	0.8176985004103839
Gamma	0.0979914312095726
Learning Rate	0.04356818667316142
Max Depth	9
N Estimators	184
Subsample	0.7554709158757928

Table 6. *LSTM* hyper parameter

Hyper parameter <i>LSTM</i>	
Embedding Layer Output	$\frac{1}{2}$ of Input Shape
<i>LSTM</i> 4h Input	32
<i>LSTM</i> 15min Input	120
<i>LSTM</i> Dropout	0.0001
<i>LSTM</i> Recurrent Dropout	0
<i>LSTM</i> l1,l2 Regularizer	0, 0.000001
<i>LSTM</i> Statefull	False
Dense Layer Embedded	32 units
Dense Layer-4h	40 units
Dense Layer-15min	24 units
Dense Layer Concatenate	50 units
Dense Layer Final	1 units
Leaky ReLU α	0.1
Learning Algorithm	Adam
Learning Rate	0.0005
Batchsize	40

References

1. Andrade, J.R., Filipe, J., Reis, M., Bessa, R.J.: Probabilistic Price Forecasting for Day-Ahead and Intraday Markets: Beyond the Statistical Model (2017)
2. Breiman, L.: Random forests. *Mach. Learn.* **45**, 5–32 (2001)
3. Chen, T., Guestrin, C.: XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd International Conference on Knowledge Discovery and Data Mining, pp. 785–794 (2016)
4. Frade, P., Vieira-Costa, J.V., Osório, G.J., Santana, J.J., Catalão, J.P.: Influence of wind power on intraday electricity spot market: a comparative study based on real data. *Energies* **11**, 2974 (2018)
5. Guo, C., Berkhahn, F.: Entity Embeddings of Categorical Variables. arXiv (2016)
6. Gürtler, M., Paulsen, T.: Forecasting performance of time series models on electricity spot markets: a quasi-meta-analysis. *Int. J. Energy Sect. Manage.* **12**, 103–129 (2018)
7. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997)
8. Kath, C.: Modeling intraday markets under the new advances of the cross-border intraday project (XBID): evidence from the german intraday market. *Energies* **12**, 4339 (2019)
9. Kath, C., Ziel, F.: The value of forecasts: quantifying the economic gains of accurate quarter-hourly electricity price forecasts. *Energy Econ.* **76**, 411–423 (2018)

10. Kiesel, R., Paraschiv, F.: Econometric analysis of 15-minute intraday electricity prices. *Energy Econ.* **64**, 77–90 (2017)
11. Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., Talwalkar, A.: Hyperband: a novel bandit-based approach to hyperparameter optimization. *J. Mach. Learn. Res.* **18**, 6765–6816 (2017)
12. Martin, H.: A Limit Order Book Model for the German Intraday Electricity Market
13. Monteiro, C., Ramirez-Rosado, I.J., Fernandez-Jimenez, L.A., Conde, P.: Short-term price forecasting models based on artificial neural networks for intraday sessions in the iberian electricity market. *Energies* **9**, 721 (2016)
14. Narajewski, M., Ziel, F.: Econometric Modelling and Forecasting of Intraday Electricity Prices. *Journal of Commodity Markets* p. 100107 (2019)
15. Oksuz, I., Ugurlu, U.: Neural network based model comparison for intraday electricity price forecasting. *Energies* **12**, 4557 (2019)
16. Pape, C., Hagemann, S., Weber, C.: Are fundamentals enough? explaining price variations in the german day-ahead and intraday power market. *Energy Econ.* **54**, 376–387 (2016)
17. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986)
18. Uniejewski, B., Marcjasz, G., Weron, R.: Understanding intraday electricity markets: variable selection and very short-term price forecasting using LASSO. *Int. J. Forecast.* **35**, 1533–1547 (2019)
19. Uniejewski, B., Weron, R.: Efficient forecasting of electricity spot prices with expert and lasso models. *Energies* **11**, 2039 (2018)
20. Wessel, A., Dobschinski, J., Lange, B.: Integration of offsite wind speed measurements in shortest-term wind power prediction systems. In: 8th International Workshop on Large-Scale Integration of Wind Power into Power Systems, pp. 14–15 (2009)
21. Ziel, F.: Modeling the impact of wind and solar power forecasting errors on intraday electricity prices. In: 2017 14th International Conference on the European Energy Market (EEM), pp. 1–5 (2017)