Profit-Oriented Electricity Price Forecast Method for Demand Responses

Sinan Cai, Masahiro Mae, Ryuji Matsuhashi
Department of Electrical Engineering & Information Systems
The University of Tokyo
Tokyo, Japan
cai@enesys.t.u-tokyo.ac.jp
ORCID: 0000-0002-7536-0620

***Abstract*—This paper proposes a novel electricity price forecast method for demand responses (DRs) who participate in the electricity market. The conventional forecast methods, especially machine learning-based methods, tend to model the featured training data by minimizing the mean-square-error (MSE). In such methods, the loss function is defined to be the MSE. However, due to the characteristics of the DRs, a forecast result with a good MSE value is not necessarily more beneficial for the DRs in a deregulated market environment. Alternatively, the proposed method modifies the conventional loss function to inclusively consider the gradient of the forecast data. The loss function is designed in such a way as to improve the Extremum Timing Accuracy (ETA), a profit-oriented criterion of the forecast results so that the economic benefits of the DRs can be further boosted. Forecast results on the clearing price of the day-ahead energy wholesale market in Japan, the JEPX spot market, by neural network with the proposed profit-oriented loss function are tested. Performance comparison with the conventional neural network forecast is also reported. It is proved that the proposed forecast method increases the market profit of the DR successfully.**

***Index Terms*—Demand Responses, Electricity Market, Price Forecast, Machine Learning, Neural Network**

I. INTRODUCTION

In a competitive market environment nowadays, the generation side and the consumer side submit bids respectively for selling and purchasing electricity to the trading pool and the electricity price is decided when the market is cleared by the market operator [1]. With the progressing of electricity market deregulation and the increasing needs for external resources to address the supply and demand imbalances caused by renewable generation, the demand-side resources such as DRs are welcomed to participate in the market.

Under such circumstances, the electricity prices might rise or drop excessively depending on the actions of individual participants [2]. Meanwhile, the price can only be known when the market is cleared. Moreover, the existing markets in the world are mostly in a day-ahead or hour-ahead scheme to ensure real-time system operation due to the non-storable nature of electricity. As a result, the prediction task might be extremely complicated.

This work is conducted by the Social Cooperation Research Departments of Power System Innovation Realization with Fuji Electric Co., Ltd. in Collaborative Research Organization for Comprehensive Energy Sciences (CROCES) at the University of Tokyo.

However, the market price is crucial to the market participants. Decision-making processes such as bidding strategy optimization and self-operation scheduling relies highly on the price information, thus many techniques have been adopted to challenge the task. Among all, time-series models capable of prediction based on the historic values without extra information input is a cost-effective approach [3]. The application of the auto-regressive integrated moving average (ARIMA) model in electricity price forecasting can be found in [4]. To further consider the the seasonal trend of market price, the seasonal auto-regressive integrated moving average (SARIMA) model is implemented in [5], [6]. Additionally, signal analysis techniques such as the wavelet transform are also applied to process the historical data for stochastic time-series forecast [7].

The rapid development of machine learning has also provided new and attracting options. Machine learning based forecast model can detect and extract correlations from enormous data autonomously. Forecast method based sorely on neural network can be found in [8]–[10], while combination with stochastic time-series analysis is reported in [11], [12]. [13]–[15] utilized deep learning networks featuring more sophisticated convolution neural network (CNN) and long short-term memory (LSTM) structure. Besides, a random forest regression model is developed to forecast New York electricity price in [16].

So far, most of the existing price forecast techniques only focus on the numerical precision of every forecast point. 27 methods for electricity prices forecasting are compared in terms of symmetric mean absolute percentage error (sMAPE) in [17]. However, it is proved in [18] that numerical precision of the forecast result is not an appropriate criterion for the DRs in the aspect of economic benefits. The DRs are demand-side-owned equipment and participating in the market is usually not their priority. Hence the DRs are very sensitive economically. The main contribution of this paper is proposing a novel forecast method from an economic perspective. Such a forecast method has not been seen in the existing literature.

The rest of this paper is organized as follows: Section II introduces the profit-oriented forecast criterion for the DRs. The novel loss function design is discussed in Section III. Forecast simulation with the real historical price of the day-ahead energy market in Japan is demonstrated to validate the performance of the proposed forecast method in Section IV, followed by the conclusion in Section V.

II. Profit-Oriented Forecast Criterion

The mean-square-error (MSE) is commonly used to evaluate the forecast result by summing the square of the forecast errors of all the forecast points. The smaller the MSE, the closer the result is to the actual value and the better the result is considered. Likewise, the mean-absolute-error (MAE), sMAPE and the Huber Loss are also regular approaches for regression tasks [17], [19]. All of these methods focus on the numerical precision of every single forecast point. The concept is very straightforward and intuitive since a good forecast result should be as close to the actual values as possible. These error-based evaluations are positive values whose scale might vary with different datasets. To get a general impression of how good a forecast it is without knowing the scale of the actual datasets, the MSE is sometimes converted linearly to the R-squared coefficient which is a unit value in most cases. The forecast result is considered good when the R-squared coefficient is close to 1.

For market participants like a generation company or an electricity retailer, a precise forecast on every time step is desired. The day-ahead generation plan or electricity purchase/sell plan covers the whole day and needs to be scheduled according to the forecast of the energy whole sell market price. However, the situation is different for the DRs. It is first reported in [20] that an EV aggregator gains 88.3% of the theoretical maximum revenue from the frequency regulation market during the parking time with a 'poor' prediction whose R-squared value is only 0.2. Typical DRs are household equipment such as water heat pumps, electric vehicles (EVs), or air conditioners with their own purposes and hence are unlikely to participate in the market for a long time span. Meanwhile, the capacity of an individual DR is very small even when aggregated compared to other traditional resources in the system. In a single-price auction market like most of the actual energy spot markets in the world, the DR will only act as a price taker and cannot affect the market price directly as big generation or consumption resources do.

Given the above inherent characteristics of the DRs, the conventional evaluation such as the MSE might not be appropriate since it is sorely based on the sum of the error of all the forecast points. The most critical information for DRs from a profit perspective is the timing of the peak and dip in the market price. The price peaks and dips are the most profitable moments, thus the DRs should arrange their schedule and participate in the market at that moment to utilize it. Under such circumstances, the accuracy of the timing of the price peak and dip should be taken into consideration when evaluating a forecast result. The Extremum Timing Accuracy (ETA) is proposed in [18] to evaluate the forecast results as follows:

$$ETA = \frac{\sum_{i \in L\max^F} T(i) - \sum_{i \in L\min^F} T(i)}{\sum_{i \in L\max^A} T(i) - \sum_{i \in L\min^A} T(i)} \quad (1)$$

T is the actual price in the time series. $L\max$ and $L\min$ are the indexes of the local maximum and local minimum point, while F and A stand for the forecast result and the actual price respectively. Like the R-squared coefficient, ETA is a unit value with 1 indicating most profitable and 0 indicating unprofitable.

An example of the application of the ETA is given here. Two forecast results are shown in Fig. 1. At first sight, Prediction 2 seems to be the better prediction with an R-squared value of 0.87 apparently, while the R-squared value of Prediction 1 is only 0.19. However, the timing of the highest and the lowest price in Prediction 2 is completely wrong. For a simple battery storage system that intends to purchase energy at a low price and sell at a high price, only 0.62 profit can be obtained. On the contrary, Prediction 1 is correct about the timing of the highest and the lowest price and full profit can be achieved. The ETA of Prediction 1 and Prediction 2 is 1 and 0.3 respectively, consistent with the profit calculation. The details of the example are laid out in Table I. Conclusively, a forecast result with a high ETA value can be regarded as more profitable for DR operation scheduling.

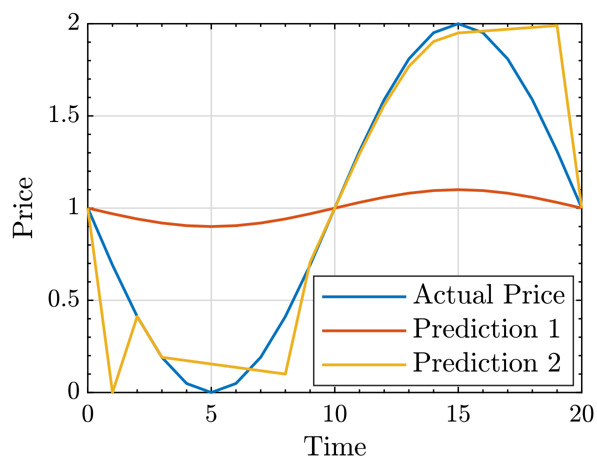

Fig. 1. Example of price forecast result comparison.

TABLE I
Operation Schedule and Profit: An Example

	Charge Time	Discharge Time	R^2	ETA	Profit
Actual	5	15	1	1	2
Prediction 1	5	15	0.19	1	2
Prediction 2	1	19	0.87	0.3	0.62

III. The Profit-Oriented Loss Function

All machine learning techniques require a loss function as a learning target. Typically, the loss function is not only used to evaluate the learning result, but the steepest gradient descent of the loss function also guides the direction of the learning. In deep neural networks, the gradient of the loss function with respect to the learnable weights of the network is calculated via backpropagation to update the network [21].

According to the discussion in Section II, a profit-oriented forecast can be interpreted as producing a forecast result with a high ETA value. The loss function of the profit-oriented forecast should be implemented in a way that maximizes the ETA value. One reason that MSE is frequently chosen as the loss function is that its gradient can be calculated easily. Unfortunately, the ETA defined in (1) is non-differentiable since the definition includes the process of searching for the local extremums. One solution is to find an analytic continuation of the ETA. Denote $Y(t)$ and $T(t)$ to be the forecast and actual value, the analytic continuation of the ETA is expressed as:

$$ETA = \frac{\sum_t T(t)[2\sigma(k\nabla Y^- \mathrm{ReLu}(-\nabla Y^- \nabla Y^+)) - 1]}{\sum_t T(t)[2\sigma(k\nabla T^- \mathrm{ReLu}(-\nabla T^- \nabla T^+)) - 1]} \quad (2)$$

where

$$\nabla Y^- = Y(t) - Y(t-1) \quad (3)$$

$$\nabla Y^+ = Y(t+1) - Y(t) \quad (4)$$

$$\nabla T^- = T(t) - T(t-1) \quad (5)$$

$$\nabla T^+ = T(t+1) - T(t) \quad (6)$$

σ is the sigmoid function and $ReLu$ is the ReLu function. k is a scaling coefficient. Equation (2) then becomes differentiable.

However, the gradient of (2) is mainly zero. In fact, this phenomenon is reasonable. The basic idea of ETA is to evaluate the accuracy of the foretasted local maximum and minimum timing, and the change in the non-local-extremum point values will not affect the ETA since they are not involved in the calculation of the ETA. To sum up, the ETA is a much looser criterion than the MSE by definition. Therefore, even though it is a good criterion for forecast result evaluation, it cannot provide too much information and guidance for the network training process.

The more practical solution is to alter the existing MSE loss function without inserting ETA directly. As the overall trend is more decisive than actual point values in ETA, the proposed profit-oriented loss is given as:

$$L = \sum_t (Y(t) - T(t))^2 + \alpha \sum_t [\sigma(\beta \nabla Y^-) - \sigma(\beta \nabla T^-)]^2 \quad (7)$$

α and β are two tuning parameters. The former part of (7) is the conventional MSE loss function. The latter part is to improve the ETA. The larger α is, the more significant the ETA part in the overall loss function will be. ∇Y^- and the ∇T^- can be regarded as the derivatives of the forecast and the actual price. The sigmoid function and the tuning parameter β convert the derivatives of Y close to 1 when $Y(t) - Y(t-1) > 0$ and close to 0 when $Y(t) - Y(t-1) < 0$. Through such a design, there will be losses when the forecast Y and the actual price T are moving in the opposite direction. If the moving direction of Y and T is consistent, the timing of the local extremums is expected to be the same, and the ETA would increase in turn.

It is notable that by inserting the ETA part into the loss function, the proposed loss function is trying to trade off some degradation in MSE for the improvement of ETA. The trade-off degree is controlled by the tuning parameter β. If β is too big, the gradient of L will become mostly zeros, providing no information for training. If β is chosen too small, only the center part of the sigmoid function will be activated, which is nearly linear. The ETA part then becomes equivalent to the MSE of the derivative of $Y(t)$, and the learning might be directed to the minimization of MSE again, resulting in no improvement of ETA.

IV. Simulation Results

A 3-layer neural network is used to forecast the spot price of the energy wholesale market in Japan: the JEPX spot market. The JEPX spot market is a day-ahead single-price auction market with a 30-minute time step. The market participants need to submit the bid by 10:00 one day before [22]. The structure of the neural network is illustrated in Fig. 2. The proposed loss function in (7) is implemented in the regression layer. The historical data from April 1st, 2022 to May 31st, 2023 is used as the training set. The focus of this paper is to examine the performance of the proposed loss function. The detailed design of the network structure is outside the topic of this paper. The training and testing of the neural network is completed in MATLAB 2023b.

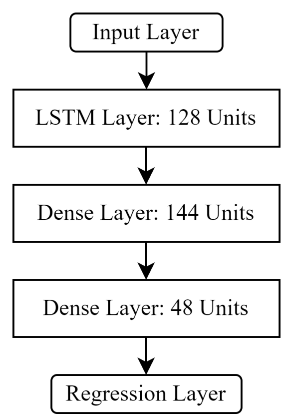

Fig. 2. Structure of the Neural Network

The R-squared and ETA values with the proposed profit-oriented loss function are shown in Fig. 3. The R-squared and ETA values of the training result with the conventional MSE loss function are 0.84 and 0.61 respectively, which are denoted by the horizontal black lines for comparison. The ETA is generally improved compared with the conventional MSE loss function, but the extent of improvement depends on the tuning parameters.

With the increase of α, the influence of the ETA part in the loss function gradually expands and the ETA values rise. However, the over-domination of the ETA part will also result in a severe drop in the R-squared value. On the other hand, $\beta = 10$ provides the best performance compared to other values examined. As pointed out in Section III, the learning

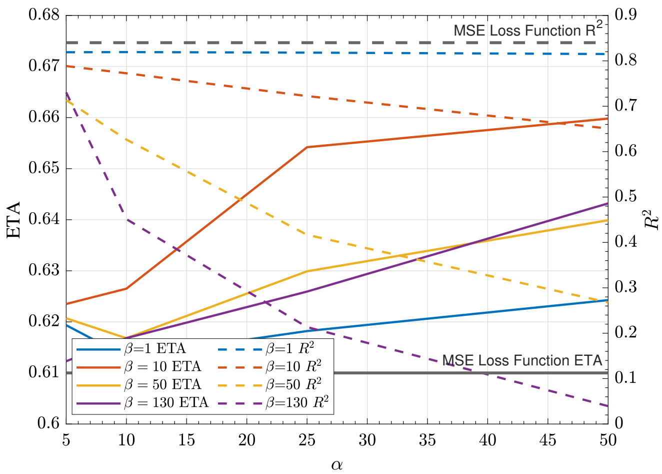

Fig. 3. The Training Result with Proposed Loss Function

is directed to optimize the MSE when $\beta = 1$. The ETA value hardly improved and the R-squared value did not decline with increasing α. When β is over 50, the sparsity of the loss function gradient results in insufficient guidance for learning. The improvement in ETA is limited compared to the $\beta = 10$ case.

Based on the above analysis, $\alpha = 25$ and $\beta = 10$ are used to forecast the JEPX spot price from June 1st, 2023 to November 15th, 2023. A fraction of the forecast results are plotted in Fig. 4. Compared with the forecast using the conventional MSE loss function represented by the red line, the forecast result using the proposed method is less sharp at the price peaks and dips. However, the DR cares more about the timing of the peaks and dips rather than their exact value. The deterioration of the MSE in these parts is traded for the improvement in ETA. The ETA of the conventional MSE forecast is only 0.57, while the ETA of the proposed method is 0.61.

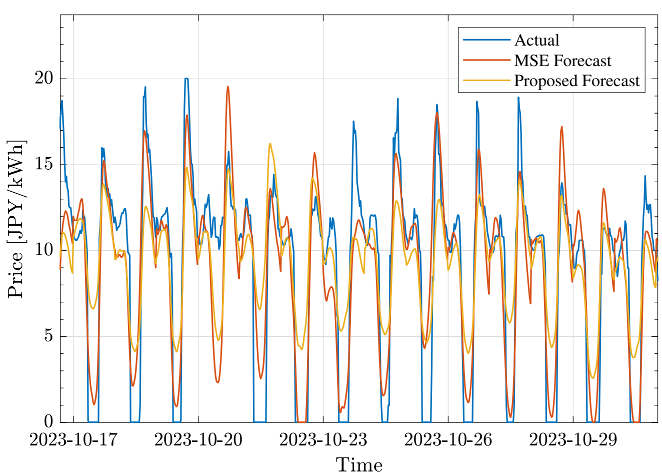

Fig. 4. Forecast Result

The forecast results are provided for a simple demand-side battery storage system that intends to purchase charging power from the energy spot market and sell the power back to the market by discharging power to gain profit. The daily operation scheduling of the DR is described as an optimization problem in (8):

$$\underset{State(t)}{\text{maximize}} \quad P_{max} \sum_t Y(t) State(t) \tag{8a}$$

$$\text{subject to} \quad SOC(t) = SOC_{initial} - P_{max} \Delta t \sum_t State(t), \tag{8b}$$

$$0 \leq SOC(t) \leq 1, \tag{8c}$$

$$-\sum_t State(t) \leq n \text{ if } State(t) < 0 \tag{8d}$$

The DR aims at maximizing its one-day profit by the objective function (8a) and the given forecast result $Y(t)$. P_{max} is the maximum power output of the DR. $State(t) \in [-1, 1]$ denotes the charging and discharging schedule to be optimized. The state-of-charge (SOC) of the DR is calculated by constraint (8b) and maintained by constraint (8c). To prevent massive battery deterioration, the number of charging cycles is limited by (8d) where n is the number of charging cycles allowed per day. Δt is the length of the market time-step.

In this simulation, the DR's capacity is $10\,\text{MWh}$ with a maximum power output of $10\,\text{MW}$. $n = 2$ charging cycles are allowed per day. The optimization is executed on a daily basis by Mixed Integer Linear Programming. The profit of the DR is calculated as in (9) since the DR acts as a price-taker in the market. $State(t)$ is the optimized charging and discharging schedule in (8) and $T(t)$ is the actual spot price.

$$Profit = P_{max} \Delta t \sum_t T(t) State(t) \tag{9}$$

The DR's profit from June to October is shown in Table II. The Oracle case is a case in which the DR predicts the spot price perfectly, indicating the theoretical maximum profit. The proposed forecast method boosts the total profit by 6% compared with the conventional MSE forecast, reaching 70% of the theoretical maximum value.

TABLE II
DR'S PROFIT (THOUSAND JPY)

	June	July	August	Sepetember	October	Total
Oracle	391.63	504.18	482.18	546.94	497.85	2422.78
MSE Forecast	247.48	278.30	300.00	366.96	416.03	1608.75
Proposed Forecast	263.60	314.74	305.69	396.26	425.58	1705.86

A detailed operation schedule on Fig. 5. The blue bars are the bids for purchasing or selling energy submitted to the market. The proposed forecast method is more precise about the timing of the peak on 10:00 and the dip on 11:30 despite that the peak and the dip are less obvious in terms of numerical value. The MSE forecast predicts the same peak and dip earlier than they actually are, leading to undesirable profit loss.

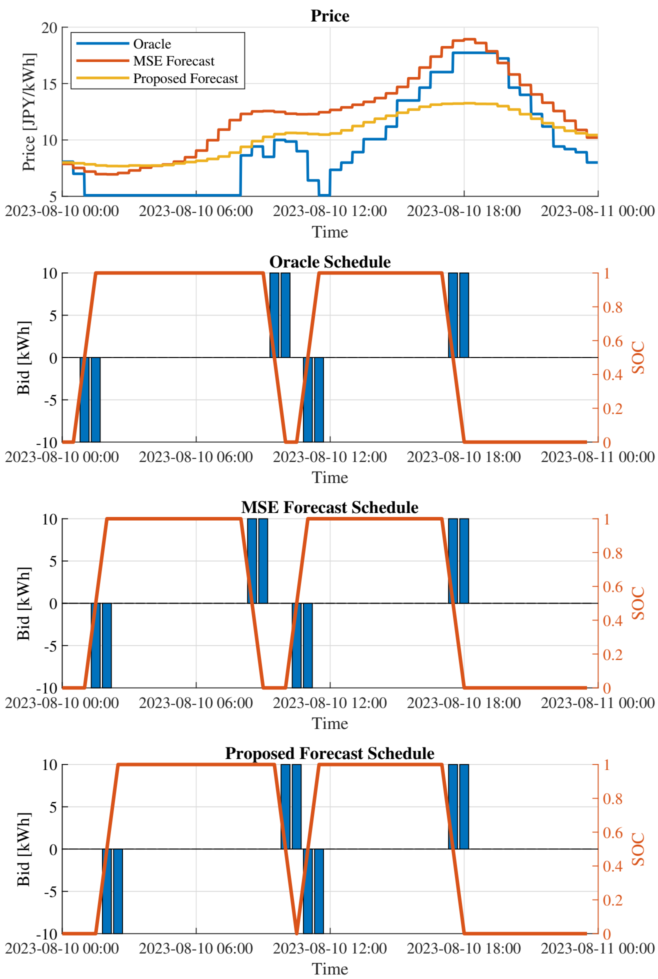

Fig. 5. Operation Schedule on August 10th, 2023

V. Conclusions

This paper proposes a novel loss function for electricity price forecasts. The proposed method is specifically designed to consider the unique characteristics of the DRs. Unlike the conventional loss functions for regression, the proposed loss function involves the gradient of the forecast data and endeavors to improve the profit-oriented criterion ETA by trading off some performance in the aspect of MSE. A neural network equipped with the proposed loss function is capable of producing forecast results with higher ETA, and the DR's profit is successfully boosted compared with the conventional forecast based on MSE.

The proposed loss function can not only be used in neural networks but also in other machine learning forecasts such as random forests regression. The performance analysis on deep learning networks with the proposed loss function might also be a potential future research direction.

References

[1] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn, *Spot pricing of electricity*. Springer Science & Business Media, 2013.

[2] J. P. d. S. Catalão, S. J. P. S. Mariano, V. Mendes, and L. Ferreira, "Short-term electricity prices forecasting in a competitive market: A neural network approach," *electric power systems research*, vol. 77, no. 10, pp. 1297–1304, 2007.

[3] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[4] T. Jakaša, I. Andročec, and P. Sprčić, "Electricity price forecasting—arima model approach," in *2011 8th International Conference on the European Energy Market (EEM)*. IEEE, 2011, pp. 222–225.

[5] P. Rajan and K. V. Chandrakala, "Statistical model approach of electricity price forecasting for indian electricity market," in *2021 IEEE Madras Section Conference (MASCON)*. IEEE, 2021, pp. 1–5.

[6] S. Cai and R. Matsuhashi, "Model predictive control for ev aggregators participating in system frequency regulation market," *IEEE Access*, vol. 9, pp. 80 763–80 771, 2021.

[7] A. J. Conejo, M. A. Plazas, R. Espinola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and arima models," *IEEE transactions on power systems*, vol. 20, no. 2, pp. 1035–1042, 2005.

[8] D. Singhal and K. Swarup, "Electricity price forecasting using artificial neural networks," *International Journal of Electrical Power & Energy Systems*, vol. 33, no. 3, pp. 550–555, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061510002231

[9] D. Keles, J. Scelle, F. Paraschiv, and W. Fichtner, "Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks," *Applied energy*, vol. 162, pp. 218–230, 2016.

[10] D. Wang, H. Luo, O. Grunder, Y. Lin, and H. Guo, "Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and bp neural network optimized by firefly algorithm," *Applied Energy*, vol. 190, pp. 390–407, 2017.

[11] N. Amjady and F. Keynia, "Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method," *International Journal of Electrical Power & Energy Systems*, vol. 30, no. 9, pp. 533–546, 2008.

[12] J. C. R Filho, C. M. Affonso, and R. C. Oliveira, "Energy price forecasting in the north brazilian market using nn-arima model and explanatory variables," in *2014 IEEE Symposium on Computational Intelligence for Engineering Solutions (CIES)*. IEEE, 2014, pp. 171–175.

[13] P.-H. Kuo and C.-J. Huang, "An electricity price forecasting model by hybrid structured deep neural networks," *Sustainability*, vol. 10, no. 4, p. 1280, 2018.

[14] K. Wang, M. Yu, D. Niu, Y. Liang, S. Peng, and X. Xu, "Short-term electricity price forecasting based on similarity day screening, two-layer decomposition technique and bi-lstm neural network," *Applied Soft Computing*, vol. 136, p. 110018, 2023.

[15] F. H. Mohammadreza Heidarpanah and M. Fazeli, "Daily electricity price forecasting using artificial intelligence models in the iranian electricity market," *Energy*, vol. 263, p. 126011, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360544222028973

[16] J. Mei, D. He, R. Harley, T. Habetler, and G. Qu, "A random forest method for real-time price forecasting in new york electricity market," in *2014 IEEE PES general meeting— conference & exposition*. IEEE, 2014, pp. 1–5.

[17] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Applied Energy*, vol. 221, pp. 386–405, 2018.

[18] S. Cai, M. Mae, and R. Matsuhashi, "A novel criterion of electricity price forecast for demand-side responses participating in the electricity market (under reviewing)," in *2024 20th International Conference on the European Energy Market (EEM)*. IEEE, 2024.

[19] P. J. Huber, "Robust estimation of a location parameter," in *Breakthroughs in statistics: Methodology and distribution*. Springer, 1992, pp. 492–518.

[20] S. Cai and R. Matsuhashi, "Model predictive control for ev aggregators participating in system frequency regulation market," *IEEE Access*, vol. 9, pp. 80 763–80 771, 2021.

[21] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[22] *Japan Electric Power eXchange Guide*, Japan Electric Power Exchange, 2019. [Online]. Available: https://www.jepx.jp/electricpower/outline/pdf/Guide_2.00.pdf