# A Novel Criterion of Electricity Price Forecast for Demand-side Responses Participating in the Electricity Market

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 criterion for evaluating electricity price forecasting results for demand-side responses (DRs) who participate in the electricity market. Generally, the DR needs to predict the market price and arrange its bidding and operation schedule according to the forecast result. The mean-square-error (MSE) or the R-squared coefficient is used for evaluating the forecast result conventionally. However, it is shown in this paper that a forecast result with a good MSE or R-squared value is not necessarily more beneficial for the DRs. Instead, the proposed novel criterion can reflect the influence of different forecast results on the DR's market revenue and help the DR identify which forecast result is better regarding economic benefits. The proposed criterion emphasizes the accuracy of predicting the timing of the price peaks and dips, which is more crucial information to the DRs than the total numerical forecast precision over time. Results comparing different forecast methods on the clearing price of the day-ahead energy wholesale market in Japan, the JEPX spot market, are reported. The better forecast method for the DR is identified by the proposed criterion successfully.**

***Index Terms*—Demand-side Responses, Electricity Market, Price Forecast, ARIMA models, Deep Learning**

## I. Introduction

The electricity price has been an important research topic in power systems long before the deregulation of the power market. Models to calculate the price on an hour-by-hour, bus-by-bus level were developed based on future regional demand estimation and regional supply optimal dispatching [1]. The models aim at system cost minimization and are very detailed to include all the crucial components in the system since all the information is available. With the assumption of a low uncertainty environment and no extra exercisable market power, one can easily obtain the electricity price through the model.

The deregulated market environment has replaced the traditional monopolistic scheme in the electricity industry in many countries. Except for the direct bilateral contracts, the electricity price is mostly determined by pool trading, where the generation side and consumer side submit bids respectively for selling and purchasing electricity [2]. The electricity prices might rise or drop excessively depending on the accepted bids in the pool when the market operator clears the market. More importantly, not only the traditional generators and power companies but also the demand-side resources such as DRs can participate in the market.

As the competitive framework has been introduced into the electricity industry and market, a price model with detailed components of all the resources' characteristics and operation behaviors becomes impossible since no entity could have all the information on every individual market participant. The price can only be known when the market is cleared. Moreover, due to the non-storable nature of electrical energy, the markets are usually in a day-ahead or hour-ahead scheme to ensure that during real-time operation balance between demand and supply can be maintained.

The market price is essential to the market participants' decision-making process such as bidding strategy optimization and self-operation scheduling. As a result, short-term price forecasting has emerged as a heated research topic in the past two decades. The price curve in actual electricity markets exhibits a far more complex structure than the load curve such as nonconstant mean and variance, high level of volatility, and unexpected spikes [3]. The prediction task can be extremely complicated and many techniques have been adopted to challenge the task.

The price forecasting methodologies can be generally classified into three categories: the game theory models, the simulation models, and the time-series models [4]. For a normal market participant, since the detailed information of other participants in the market is usually unclear, time-series models focusing on the past behavior of the dependent variables are a more cost-effective approach. The auto-regressive integrated moving average (ARIMA) model is a common stochastic time-series analysis method and predicts future values based on the historic values without extra information input [5]. The application of the ARIMA model in electricity

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.

979-8-3503-8174-0/24/$31.00 ©2024 IEEE

price forecasting can be found in [6], while the seasonal autoregressive integrated moving average (SARIMA) model is used to consider the seasonal trend of Indian electricity market price in [7]. Likewise, the utilization of Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model is reported in [8]. The stochastic time-series forecast can be further improved by adopting the wavelet transform to pre-process the historical price data [9].

The rapid development of machine learning has also inspired the forecast models based on deep learning neural networks. These models are sorely data-based and can detect and extract correlations from historical time-series data autonomously. Common back propagation neural network is used in [10]–[12], while the application of the more sophisticated convolution neural network (CNN) and long short-term memory (LSTM) structure can be found in [13]–[15]. Moreover, the combination of stochastic time-series analysis and neural network models is reported in [16], [17].

However, different forecast techniques might perform differently in different markets and under different situations, especially when most of the real-world electricity market is still evolving with new mechanisms being set up to address current system problems like increasing and intermittent renewable source generation. With various existing forecast techniques to choose from, a market participant would require a criterion to help them select the most suitable forecast techniques. So far, most of the existing studies only focus on the numerical precision of every forecast point, which does not necessarily stand for a good forecast. For example, in the clinical field, the overall trajectory of a patient's blood glucose is more important than the numerical value [18]. The capacity of the DRs is usually small and their operation behaviors are different from the conventional resources in the system. The business model of DRs usually relies on aggregation or change in lifestyle, making them very sensitive to economic profits. For this reason, a different criterion of forecast evaluation is needed. The core of this paper is not to compare the existing forecast methods directly but to provide a way to distinguish the performance of given forecast results from an economic perspective.

The rest of this paper is organized as follows: the potential problem of the conventional forecast result evaluation criterion is identified in Section II. Section III introduces the proposed novel criterion. The simulation with the real historical price of the day-ahead energy market in Japan is demonstrated to validate the proposed criterion in Section IV, followed by the conclusion in Section V.

## II. Deficiency of Conventional Forecast Result Evaluation

The mean-absolute-error (MAE) and the mean-square-error (MSE) are commonly used criteria for forecast result evaluation, both evaluate the forecast result by summing the forecast errors of all the forecast points. The smaller the sum of the errors, the better the result is. Likewise, the R-squared coefficient converts the sum of forecast errors to a coefficient below 1. The forecast result is considered good if the R-squared coefficient is close to 1.

The methods are very straightforward and intuitive, as a good forecast result is expected to be as close as possible to the actual values. Generally, this is the case for market price forecasts, whether in stock markets or electricity markets. The market participants have the motivation to keep track of the prices at every time step so that better strategies can be taken. For example, a generation company or an electricity retailer can schedule their day-ahead generation plan or electricity purchase/sell plan according to the forecast of the energy whole sell market price. The plan covers the whole day, therefore, a precise forecast on every time step is desired.

However, the case is different for the DRs, mainly for two reasons. First, the DRs are unlikely to participate in the market during a long time span. Typical DRs are household equipment such as water heat pumps, electric vehicles (EVs), or air conditioners. Those equipment have their own purposes and cannot always respond to the system's demand. Even for battery storage systems, the charging and discharging time only lasts for a few hours. Second, the capacity of a DR is generally very small compared to other resources in the market even when aggregated. In most cases, the DR will act as a price taker in the market and does not affect the market price directly like big generation or consumption resources.

A detailed example is given in Fig. 1. The blue line denotes the actual clearing price of a single-price market and the two forecast results are shown in red and yellow respectively. At first glance, Prediction 2 is clearly the better prediction with a R-squared value of 0.87, while the R-squared value of Prediction 1 is only 0.19. To maximize the profit, the DRs are likely to consume energy when the price is low and output energy when the price is high. In this case, assuming the DR is a battery system that charges at the lowest price and discharges at the highest price according to the forecast results, the operation schedule and the profit of the DR are listed in Table I. In spite of a better R-squared value, Prediction 2 wrongly forecast the timing of the highest and the lowest price, leading to 0.62 profit of the DR. On the other hand, even with a worse R-squared value, Prediction 1 is correct about the timing of the highest and the lowest price, hence full profit can be obtained.

TABLE I
DR's Operation Schedule and Profit

| | Charge Time | Discharge Time | Profit | $R^2$ |
|---|---|---|---|---|
| **Actual** | 5 | 15 | 2 | 1 |
| **Prediction 1** | 5 | 15 | 2 | 0.19 |
| **Prediction 2** | 1 | 19 | 0.62 | 0.87 |

To sum up, the critical market price information for DRs is to find out the timing of the peak and dip in the price, so that they can arrange their schedule and participate in the market when the peak or dip comes. This phenomenon is first pointed out in [19], where aggregated EVs participate in the frequency regulation market during their parking time.

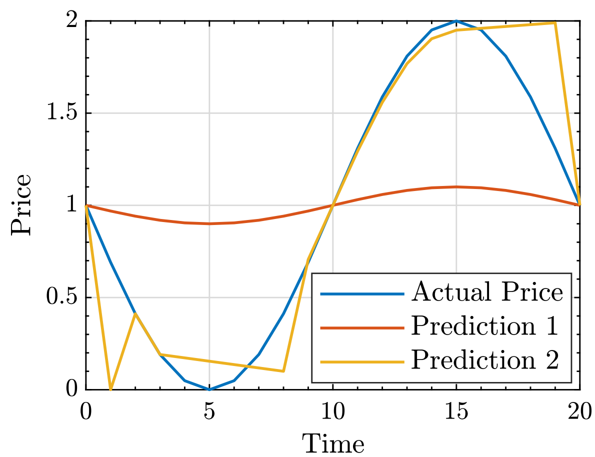


Fig. 1. Example of price forecast result comparison.

The EV aggregators gain 88.3% of the theoretical maximum revenue with a 'poor' prediction whose R-squared value is only 0.2. The conventional way of forecast result evaluation sorely based on the sum of the error of all the forecast points is not appropriate due to the inherent characteristics of DRs. When evaluating forecast results for DRs, the accuracy of the timing of the price peak and dip should be taken into consideration as well.

## III. The Novel Criterion for Demand-side Response

The most intuitive and accurate way to evaluate a forecast result from a profit perspective is to directly compare the actual profit from the market as in [19]: the optimal operation schedule and market bidding strategy was first determined to maximize the EV aggregators' profit and then the actual profit was calculated by market simulation with the actual market price. Different forecasts resulted in different operation strategies, which in turn led to different profit losses.

However, such an approach has a few drawbacks:

- The operation strategy of the DRs must be determined and modeled in advance. Meanwhile, the capacity and output power also need to be fixed in order to perform operation optimization and market simulation for profit calculation.
- The result of direct profit comparison is only valid for the specific case.

Therefore, a general way to evaluate the forecast results can be more helpful and informative for the DRs, especially when the business model is not decided. The proposed novel criterion, Extremum Timing Accuracy ($ETA$), is defined in (1):

$$ETA = \frac{\sum_{i \in LMax^F} Y(i) - \sum_{i \in LMin^F} Y(i)}{\sum_{i \in LMax^A} Y(i) - \sum_{i \in LMin^A} Y(i)} \qquad (1)$$

$Y$ is the actual price in the time series. $Lmax$ and $Lmin$ 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. The basic idea of $ETA$ is to evaluate whether the foretasted local maximum and minimum timing is correct in the actual price, the larger the $ETA$ is, the better the forecast result. Note that the numerator part of the $ETA$ alone can be used for evaluation as well. The denominator part is used to scale the $ETA$ down to a value not bigger than 1, and it is the same for all forecast results. When the timing of the local maximum and the local minimum of the forecast result is exactly the same as those of the actual price, the $ETA$ reaches 1, indicating a perfect forecast in terms of the $ETA$.

A calculation example using the price and forecast results shown in Fig. 1 is given here. For the denominator part, the actual price's local maximum and local minimum index are 15 and 5, respectively. Accordingly, the denominator part of the $ETA$ is $Y(15) - Y(5) = 2 - 0 = 0$. For Prediction 1, the local maximum and the local minimum index are identical to the actual price, thus the numerator part equals 2 as well and the $ETA$ is 1. On the other hand, Prediction 2 has two local maximums(at 2 and 19) and two local minimums(at 1 and 8). The numerator part can be calculated as $\sum_{i=2,19} Y(i) - \sum_{i=1,8} Y(i)$. Despite a R-squared value of 0.87, the $ETA$ drops to 0.3, indicating that the accuracy of the extremum timings is low and the forecast result is considered poor for the DRs. The above calculation process is summarized in Table II.

TABLE II
ETA Calculation

| | $Lmax$ | $Y(i)$ | $Lmin$ | $Y(i)$ | $ETA$ |
|---|---|---|---|---|---|
| **Prediction 1** | 15 | 2 | 5 | 0 | $\frac{2-0}{2-0} = 1$ |
| **Prediction 2** | 2 | 0.4 | 1 | 0.7 | $\frac{(0.4+1.3)-(0.7+0.4)}{2-0} = 0.3$ |
| | 19 | 1.3 | 8 | 0.4 | |

## IV. Market Simulation

The example given in Section III is an exaggerated situation to demonstrate the difference between the proposed criterion $ETA$ and the conventional R-squared value. In this section, three given forecast results are to be evaluated. A DR optimizes its operation and bidding schedule based on the given forecast results respectively. The market simulation is performed to compare the profits and validate the evaluation in MATLAB.

### A. DR Model

The DR is a 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 schedule optimization problem is described in (2):

$$\begin{aligned}
\underset{State(t)}{\text{maximize}} \quad & P_{max}\sum_t Price(t)State(t) && \text{(2a)}\\
\text{subject to} \quad & SOC(t) = SOC_{initial} - P_{max}\Delta t \sum_t State(t), && \text{(2b)}\\
& 0 \le SOC(t) \le 1, && \text{(2c)}\\
& -\sum_t State(t) \le n \text{ if } State(t) < 0 && \text{(2d)}
\end{aligned}$$

The objective function (2a) is to maximize the DR's one-day profit. $Price(t)$ is the given forecast market price and $P_{max}$ is the maximum power output of the DR. $State(t) \in [-1, 1]$ is the charging and discharging schedule to be optimized. The state-of-charge (SOC) of the DR is maintained by constraint (2b) and (2c). The number of charging cycles is limited by (2d) to prevent massive battery deterioration, and $n$ is the number of charging cycles allowed per day. $\Delta t$ is the length of the market time-step.

### B. Simulation Result

The target market is the energy wholesale market in Japan: the JEPX spot market. The JEPX spot market is a day-ahead single-price market with a 30-minute time step. The market participants need to submit the bid by 10:00 one day before [20].

Two of the forecast results to be evaluated are obtained by a SARIMA and a Machine Learning model. The SARIMA model is derived by the Econometrics Toolbox in MATLAB. The Machine Learning model is developed by the Japan Weather Association, utilizing historical price data, weather data, and calendar data [21]. The price data for June 2023 is used in this paper. The actual and two forecast prices are depicted in Fig. 2. The other forecast is to simply use the day-ahead price as the forecast, namely the Yesterday model.

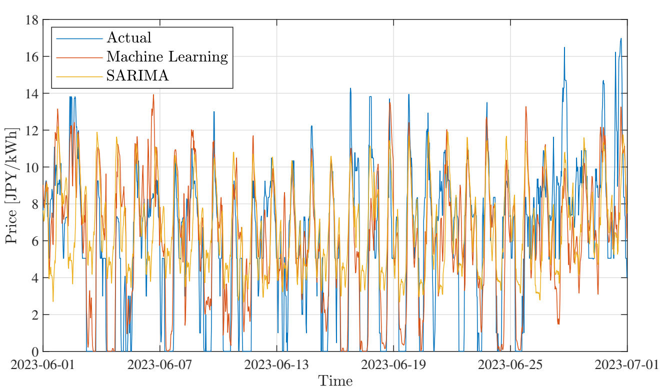


Fig. 2. Price Data

The DR acts as a price-taker in the market. Therefore the profit can be calculated as in (3). $State(t)$ is the optimized charging and discharging schedule in (2). The DR's capacity is 10 MWh with a maximum power output of 10MW. 2 charging cycles are allowed per day. The optimization is executed on a daily basis by Mixed Integer Linear Programming, and the profit is calculated according to the optimization result.

$$Profit = P_{max}\Delta t \sum_t Y(t)State(t) \tag{3}$$

The simulation result is shown in Fig. III. The Oracle case is a benchmark case standing for the perfect forecast and the profit of the DR reaches the theoretical upper limit. The Machine Learning forecast has the highest R-squared value of 0.5 and can be considered the best forecast result among the three results conventionally. However, the $ETA$ of the Machine Learning forecast result is only 0.54, close to that of the Yesterday forecast. The R-squared value of the SARIMA forecast is 0.34, but the $ETA$ reaches 0.69, indicating the SARIMA forecast should be the most profitable forecast result. The outcome of the actual profit calculation agrees with the $ETA$ evaluation. The profit of the SARIMA forecast is 356.1 thousand JPY, reaching 81% of the theoretical limit. In contrast, despite the high R-squared value, the Machine Learning forecast obtained around 260 thousand ¥. On the other hand, the Yesterday forecast's R-squared value is extremely poor, but the actual profit is close to the Machine Learning forecast since their $ETA$ values are close too.

TABLE III
SIMULATION RESULT

| **Forecast** | $R^2$ | $ETA$ | **Profit (Thousand JPY)** | **Profit Ratio** |
|---|---|---|---|---|
| **Oracle** | 1 | 1 | 391.63 | 100% |
| **Machine Learning** | 0.55 | 0.54 | 265.36 | 67.76% |
| **SARIMA** | 0.34 | 0.69 | 318.11 | 81.23% |
| **Yesterday** | 0.08 | 0.53 | 249.07 | 63.60% |

The detailed operation schedule on June 10th, 2023 is shown in Fig. 3. The positive blue bar indicates selling energy by discharging, while the negative one indicates purchasing energy to charge. The battery SOC is illustrated by the red line. The Machine Learning forecast seems to be closer to the actual price, especially at around 6:00 to 15:00. However, since the price is dipping during this period, the DR is expected to purchase energy. As long as the trend of the price dipping trajectory is reflected by the forecast, the DR will be optimized to purchase energy no matter how much the predicted price is. In the meantime, the Machine Learning forecast introduces unnecessary dynamics at 8:00, 10:30, and 22:30 regardless of the high R-squared value, hence the operation bidding schedule strays away from the Oracle case. Contrariwise, the SARIMA forecast reflects the tendency of the price dipping at noon and consequently, the operation schedule is close to the Oracle case and more profitable.

## V. CONCLUSIONS

This paper proposes a novel criterion of electricity price forecast, $ETA$, for the DRs. Unlike traditional resources, the DRs generally do not participate in the market for the whole

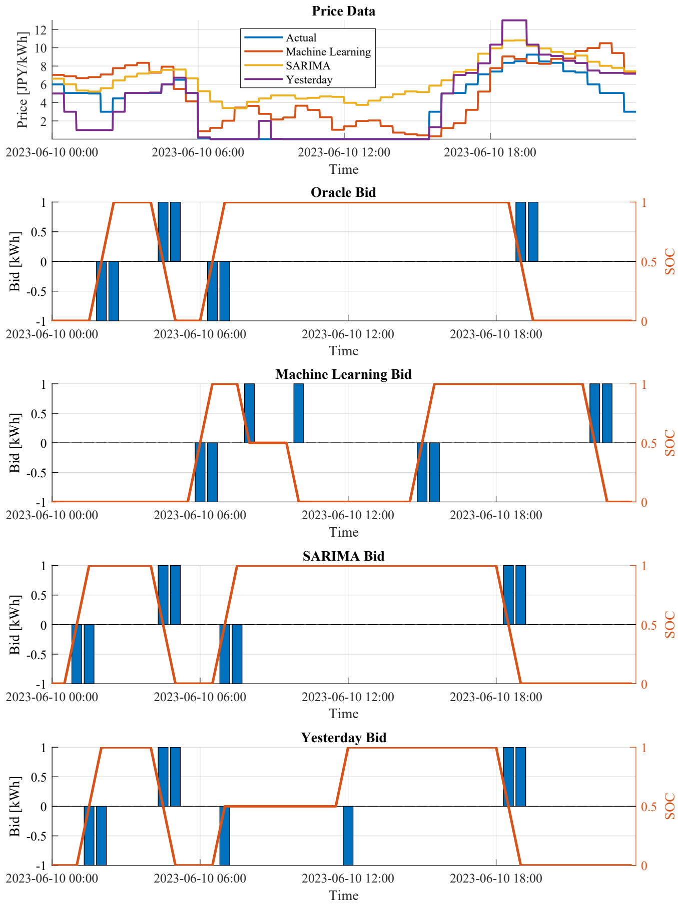


Fig. 3. Detailed Operation Schedule on June 10th, 2023

time, hence a forecast result recognized by conventional R-squared value or MSE is not necessarily beneficial for the DRs. The proposed criterion $ETA$ is designed to emphasize the timing of the price peaks and dips, which is the most critical price information for the DRs. With the help of the proposed criterion, the DRs can easily identify the most profitable forecast under different situations without developing a detailed operation model and performing simulations to calculate and compare the actual profit.

The future research direction includes the combination of $ETA$ with the R-squared value. When the DR is participating in multiple markets simultaneously, the accuracy of the forecast value becomes more important because the DR might need to compare and choose which market to participate. In such a situation, evaluation merging both $ETA$ and the R-squared value seems more reasonable and appropriate. Another potential future research direction is new forecast algorithm and technique development based on $ETA$. For example, the numerator part of $ETA$ can be used as the loss function in machine learning to design a forecast model producing high $ETA$ forecast results.

## References


[1] A. Angelus, "Electricity price forecasting in deregulated markets," *The Electricity Journal*, vol. 14, no. 3, pp. 32–41, 2001.

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

[3] 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.

[4] A. M. González, A. S. Roque, and J. García-González, "Modeling and forecasting electricity prices with input/output hidden markov models," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 13–24, 2005.

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

[6] 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.

[7] 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.

[8] R. C. Garcia, J. Contreras, M. Van Akkeren, and J. B. C. Garcia, "A garch forecasting model to predict day-ahead electricity prices," *IEEE transactions on power systems*, vol. 20, no. 2, pp. 867–874, 2005.

[9] 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.

[10] 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

[11] 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.

[12] 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.

[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] 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.

[17] 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.

[18] I. Fox, L. Ang, M. Jaiswal, R. Pop-Busui, and J. Wiens, "Deep multi-output forecasting: Learning to accurately predict blood glucose trajectories," in *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 2018, pp. 1387–1395.

[19] 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.

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

[21] "Price forecast (jepx forecast)," Japan Weather Association. [Online]. Available: https://www.jwa.or.jp/service/weather-and-data/weather-and-data-03/