

Novel Multimodal Data for Enhanced Electricity Spot Price Forecasting Using A CNN-LSTM Ensemble Learning Model for the Japan Electric Power eXchange (JEPX) Spot Market

Ziyang Wang, Masahiro Mae, Ryuji Matsuhashi
Dept. of Electrical Engineering and Information Systems
The University of Tokyo
Tokyo, Japan

wang-ziyang@ieee.org, mmae@ieee.org, matu@enesys.t.u-tokyo.ac.jp

Abstract—Electricity price forecasting (EPF) is critical in energy markets, particularly with the rise of renewable energy sources (RES) in Japan, which can cause day-ahead spot prices to drop to nearly zero JPY/kWh, impacting retailer profitability. This study demonstrates that the Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) model outperforms the LSTM model in both prediction accuracy and computational efficiency in Japan's electricity spot market. A novel ensemble learning strategy enhances both the robustness and accuracy of the EPF model is proposed. Novel multimodal explanatory variables, including electricity spot price, system price, discretionary cost, sell bid amount, buy bid amount, total contracted volume, actual power generation, actual solar power generation, meteorological forecasts and calendar forecasts, alongside the rolling features of spot prices, are utilized and verified. Furthermore, a "policy-versus-policy" approach addresses the zero-inflated regression issue of the zero price prediction is proposed. Our model, with a comprehensive feature integration, achieves an RMSE of 5.66 JPY/kWh and an R^2 of 0.729 during the test range from 2022.01.01 to 2022.12.31. The paper also introduces a novel method for estimating confidence intervals using ensemble learning.

Index Terms—electricity price forecasting (EPF), renewable energy source (RES), CNN-LSTM, ensemble learning, zero electricity spot price, confidence interval

I. INTRODUCTION

In recent years, the growing integration of renewable energy sources (RES), especially wind and solar, has led to a novel occurrence in the electricity markets of the US and EU: negative spot prices [1], [2]. In contrast, Japan's experience tells a different story. The substantial addition of solar and wind power to the Japanese grid has resulted in wholesale spot market prices nearing zero, recorded lowest at 0.01 JPY/kWh. These occurrences will be referred to as zero prices throughout this paper. The electricity spot prices in Kyushu region, Japan, is depicted in Fig. 1(a). Fig. 1(b) offers a zooming-in observation of the intermittent zero prices. Fig. 1(a) shows that the significant occurrence of zero prices started around 2020, primarily due to the swift adoption of RES. In this shifting energy environment, day-ahead electricity price forecasting

(EPF) in spot markets has become critically important [3]. Various statistical methods, including the Autoregressive Moving Average (ARMA) [4], [5] and the Autoregressive Integrated Moving Average (ARIMA) [6]–[8], have been extensively utilized in EPF research. Although these models offer a solid foundation, their linear nature can present limitations. The growing penetration of RES, coupled with demand variability, leads to non-linear patterns and abrupt price fluctuations that traditional statistical models might not accurately capture.

Additionally, as depicted in Fig. 1(b), the significant periods of zero prices in the target variable introduce a zero-inflated regression challenge in machine learning after data normalization. Nonetheless, most machine learning algorithms, such as Random Forest (RF), Support Vector Regression (SVR), and neural networks [9], usually fail to predict consistent zero values. Conventional approaches to address zero-inflated regression typically require a dual-model setup: one classification model to detect zero values and a separate regression model for non-zero values. This approach effectively doubles the training time and associated costs.

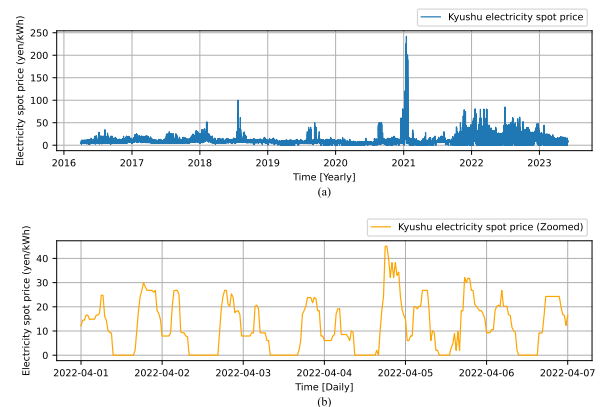


Figure 1. Kyushu electricity spot price [JPY/kWh] (a) and zooming-in zero-inflated prices (b) [10].

To tackle the aforementioned challenges, this paper aims to improve EPF in the Kyushu region of Japan by employing multimodal data as novel features, utilizing LSTM and CNN-LSTM prediction models. Additionally, a new ensemble learning approach is applied to enhance the prediction accuracy of LSTM and CNN-LSTM models, addressing their inherent uncertainties. The performance of LSTM and CNN-LSTM in EPF is evaluated based on prediction accuracy and computation time. To address the zero-inflated problem in the Japan Electric Power eXchange (JEPX) spot market, a novel "policy-versus-policy" strategy is introduced, which forecasts zero prices and halves the computation time compared to the traditional dual-model method. Furthermore, a natural logarithm transformation is used to improve the Skewness and Kurtosis of the spot prices, thereby enhancing prediction accuracy. This paper also introduces a new method for extracting meteorological forecast data using Google Maps. Finally, a novel method for estimating the confidence interval of EPF is demonstrated.

II. METHODOLOGY

A. LSTM and CNN-LSTM forecasting models

An LSTM model and a CNN-LSTM model were developed and utilized for EPF, with a comparative analysis conducted using the Python TensorFlow Keras library. The architectures of these models are illustrated in Fig. 2, and the hyperparameters were chosen empirically to achieve optimal performance.

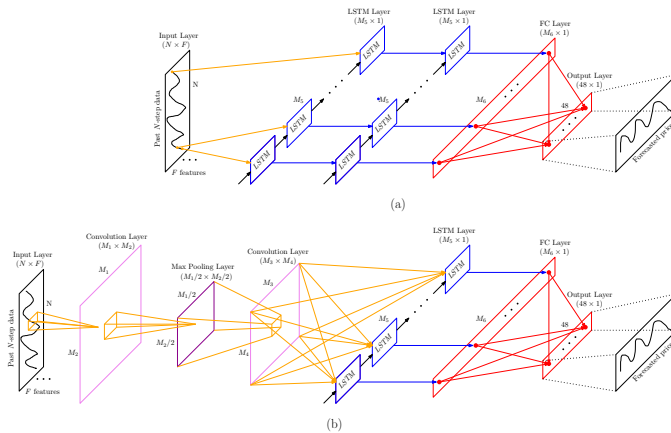


Figure 2. Schematic of the architectures of the (a) LSTM and (b) CNN-LSTM models [10].

B. Ensemble learning strategy

Due to the inherent variability of neural network models, which arises from their sensitivity to initial conditions and the stochastic nature of their training processes, training the same neural network multiple times and averaging the resulting predictions can help mitigate individual model errors. This approach leads to improved prediction performance, as different models will not produce identical errors on the test set [11], [12]. Based on this principle, an ensemble learning approach was adopted. The CNN-LSTM and LSTM models

were trained multiple times, and the individual predictions were then combined using an averaging method to form the final ensemble prediction. This process is depicted in (1), where N represents the total number of predictions and k denotes the index of each individual prediction. The pseudo-code for the ensemble learning procedure is detailed in **Algorithm 1**.

$$\hat{y}_{\text{ensemble}} = \frac{1}{N} \sum_{k=1}^N \hat{y}_k \quad (1)$$

Algorithm 1 Ensemble Learning Procedure [10]

- 1: Perform log transformation of the electricity spot price using (9).
 - 2: Normalize the training and test data.
 - 3: **for** $i = 1$ to 30 **do**
 - 4: Train the model to generate prediction $\hat{y}_i$.
 - 5: **end for**
 - 6: Restore the predicted values to their original scale (reverse data normalization).
 - 7: Apply the exponential transformation to the predicted values using (10) (reverse of the log transformation).
 - 8: Calculate the ensemble prediction $\hat{y}_{\text{ensemble}}$ using (1).
 - 9: Calculate the zero price for $\hat{y}_i$ and $\hat{y}_{\text{ensemble}}$ using (2).
-

C. "Policy-versus-policy" zero prices forecasting strategy

This study presents an innovative approach to address the zero-inflated regression issue in EPF within the JEPX spot market. The solution starts with an analysis of broader trends in global electricity spot markets. According to Seel et al. [1], an overabundance of RES can result in negative electricity spot prices in the US and EU. From this, we can deduce that negative pricing is a natural outcome of high RES penetration. In Japan, however, regulations stipulate that electricity spot prices cannot fall below 0.01 JPY/kWh, preventing negative prices. Assuming similar factors that lead to negative prices in the US and EU are present in Japan, it is logical to assume that the explanatory variables in these contexts would show similar patterns. When these Japanese variables are input into a machine learning regression model, the model would predict negative prices since it is not limited by Japan's minimum pricing regulation. Therefore, by using this logic, zero prices can be forecasted by converting any model negative outputs to zeros, as described in **Algorithm 1**. This method operates as a "policy-versus-policy" forecasting strategy, mirroring actual market conditions. Equation (2) demonstrates the procedure for calculating zero prices.

$$\hat{y}_i = \max(0, \hat{y}_i) \quad (2)$$

D. Performance evaluation

Performance metrics such as root mean squared error (RMSE) and the coefficient of determination (R^2) were used for evaluation. The formulas for calculating RMSE and R^2 are provided in (3)-(4), respectively.

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}} \quad (3)$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (4)$$

III. DATA PREPARATION

This study leverages multimodal data to improve EPF, incorporating actual total power generation, actual solar power generation, rolling features of spot prices (minimum, maximum, mean, and standard deviation), system spot price, discretionary cost, sell bid amount, buy bid amount, total contracted volume, meteorological forecast data, and calendar forecast data. The comprehensive data architecture and associated time frame are depicted in Fig. 3. The input data are divided into three green blocks based on their temporal delay. To ensure consistency with the JEPX spot price data, all inputs were linearly interpolated to a 30-min time resolution. A 7-day-long moving window was applied to the input data before being input into the CNN-LSTM and LSTM models. In the JEPX spot market, all transactions must be finalized by the 10:00 JST bidding deadline. To ensure a thorough and precise prediction process, sufficient time is allocated for the execution and refinement of neural network calculations. Considering the computational demands and the complexities involved, a 5-hour buffer before the deadline is established. The forecasting time is set to 05:00 JST, covering the entire following day from 00:00 JST to 23:30 JST, encompassing a total of 48 time frames.

In this study, photovoltaic (PV) power is the primary RES in the Kyushu region due to its significantly higher installed capacity compared to wind power. Consequently, features related to wind power are not included in this investigation.

A. Feature engineering

Feature engineering was performed by calculating rolling statistics (minimum, maximum, mean, and standard deviation) of the electricity spot price using a rolling window. This technique is vital for identifying temporal patterns and trends, which are key for extracting time series features. The rolling window length was set to 24 hours, or $w = 48$. The equations for these calculations are detailed in (5)-(8).

$$\text{Min}_t = \min(y_{t-w+1}, y_{t-w+2}, \dots, y_t) \quad (5)$$

$$\text{Max}_t = \max(y_{t-w+1}, y_{t-w+2}, \dots, y_t) \quad (6)$$

$$\text{Mean}_t = \frac{1}{w} \sum_{i=t-w+1}^t y_i \quad (7)$$

$$\text{Std}_t = \sqrt{\frac{1}{w-1} \sum_{i=t-w+1}^t (y_i - \text{Mean}_t)^2} \quad (8)$$

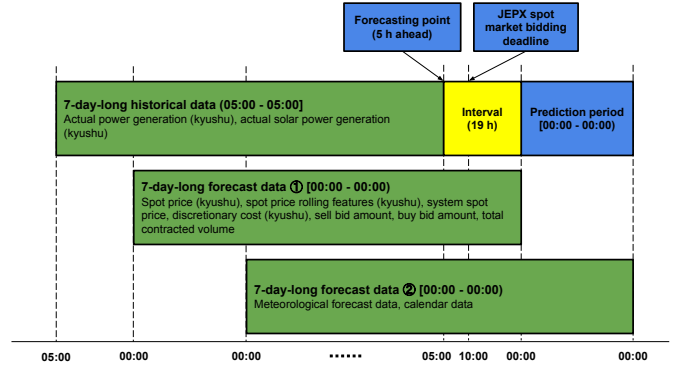


Figure 3. Illustration of the data architecture with a 30-min time interval, highlighting the time delays among different data [10].

B. Electricity data

The electricity spot prices (JPY/kWh) at a 30-min resolution, including the Kyushu regional spot price, system spot price, discretionary cost, sell bid amount, buy bid amount, and total contracted volume, from 00:00 JST to 23:30 JST (over a 7-day period), were downloaded from the JEPX [13]. The distribution of electricity spot prices in the Kyushu region, shown in Fig. 4(a), is characterized by notable Skewness and Kurtosis, suggesting a departure from a normal distribution. Neural networks generally prefer input data that are normally distributed or symmetric, as this standardizes the input feature scale, aiding in the model's learning effectiveness. When data exhibits Skewness and Kurtosis, it can lead to prediction bias and hinder the efficiency of the learning process. To address these issues, the electricity spot prices underwent a natural logarithm transformation, as illustrated in Fig. 9. This transformation method is widely used in statistics to normalize data, helping to compress the range of the distribution, improve symmetry, and diminish the effect of extreme values. The success of this transformation in reducing the Skewness and Kurtosis is documented in Table I. After the model generates its predictions, an exponential back-transformation, as described in (10), is utilized to revert the predicted values to their original scale.

$$y = \log_e(y + 1) \quad (9)$$

$$y = e^y - 1 \quad (10)$$

TABLE I. Skewness and Kurtosis of the original electricity spot price and the natural logarithm-transformed electricity spot price.

	Original	Log-transformed
Skewness	8.77	-0.68
Kurtosis	123.00	3.81

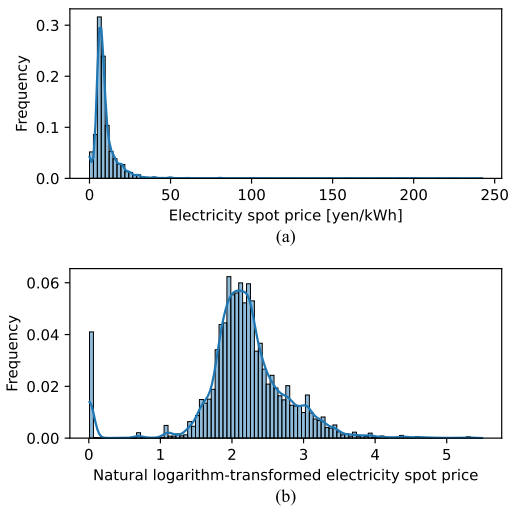


Figure 4. Distribution of Kyushu region's electricity spot prices (a) and the corresponding distribution after a natural logarithm transformation (b) [10].

The actual electricity generation data (MW) at a 30-min resolution in the Kyushu region, including the actual power generation and actual solar power generation over a 7-day period from 05:00 JST to 05:30 JST, were downloaded from the Organization for Cross-regional Coordination of Transmission Operators, Japan (OCCTO) [14]. These data were used as input features and are depicted in the leftmost green block in Fig. 3.

C. Meteorological forecast data

The meteorological forecast data of Kyushu region, including the wind speed, air temperature, relative humidity, cloudiness, precipitation, solar radiation, were downloaded from the Japan Meteorological Business Support Center (JMBSC) [15], as illustrated in the rightmost green block in Fig. 3. The map of the Kyushu region was captured as a rectangular screenshot using the Google Maps Styling Wizard [16]. The process of identifying the land areas on this map involved using the Python OpenCV library [17] to analyze the screenshot from Google Maps. Fig. 5(a) presents the initial screenshot of the Kyushu region map. During this analysis, pixels falling outside the specified range of blue colors were classified as land. Fig. 5(b) then illustrates these identified land areas, marked with red dots on the initial map screenshot. Following the land identification, the values of the meteorological data within the land regions were averaged and used as input features.

D. Calendar forecast data

The calendar forecast data, by integrating the cyclic data and culturally significant Japanese holidays, aimed to enhance the EPF accuracy by accounting for the unique electricity consumption behaviors associated with these special occasions, covering a 7-day period from 00:00 JST to 23:30 JST, were

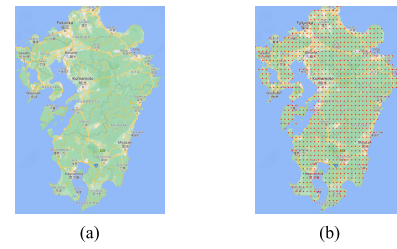


Figure 5. Kyushu area map (a) and the identified land areas (b). Map data: ©2024 Google, TMap Mobility [10].

used as features, as illustrated in the rightmost green block in Fig. 3.

1) *Cyclic data:* In this research, we utilize sine and cosine transformations to capture cyclical patterns with periods of 1 day, 1 week, 3 months, and 1 year. These transformations are well-regarded in fundamental time series analysis as an effective approach for representing periodic trends in data [18]–[20]. The input features include sine and cosine waves corresponding to the periods of 1 day, 1 week, 3 months (seasonal variations), and 1 year. An example of the sine and cosine waves with a period of 1 year is illustrated in Fig. 6.

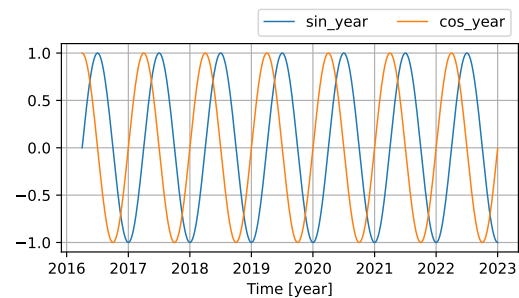


Figure 6. Sine and cosine waves with a period of 1 year [10].

2) *Holiday data:* Holidays and non-holidays were represented by numerical labels, with holidays assigned a value of 1 and non-holidays assigned a value of 0.

E. Prediction approach

A One-time prediction schematic, shown in Fig. 7(b), is used to validate the proposed features as it requires much less training time compared with the day-by-day prediction schematic shown in Fig. 7(a). In this scenario, the training data spanned from April 1, 2016 — the date marking the full liberalization of the electricity retail market — to December 31, 2021. The testing data encompass the period from January 1, 2022 to December 31, 2022. All the features proposed in Section III has been validated for prediction accuracy enhancement.

After validating the proposed features using the test data which spans the entire 2022 year, the day-by-day prediction approach was conducted using all proposed features to explore the actual prediction accuracy in real-application scenarios, as

illustrated in Fig. 7(a), which shares the same test range in the one-time prediction schematic. The training of the models was conducted using two NVIDIA Quadro RTX 8000 GPUs using the Python keras package in Windows OS.

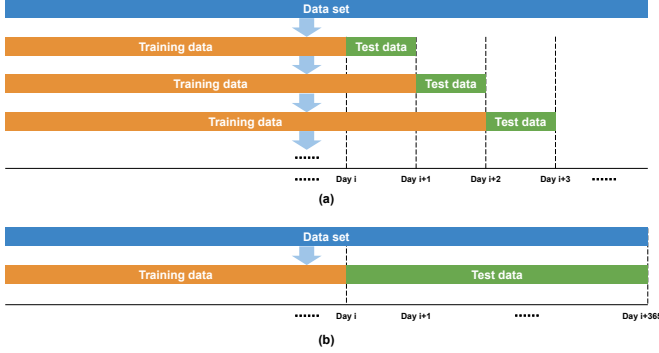


Figure 7. Day-by-day prediction schematic (a) and one-time prediction schematic (b).

IV. RESULTS AND DISCUSSION

A. Ensemble learning results

An ensemble learning technique was applied to each feature set to generate ensemble predictions. Table II presents the prediction accuracy and the computation time of the proposed LSTM and CNN-LSTM models. According to Table II, the CNN-LSTM only has half the computation time of the LSTM for same training times of ensemble learning, while still boasts higher prediction accuracy over the LSTM using less computation time.

TABLE II. Prediction accuracy and computation time comparison.

Model	Computation time	Ensemble times	R^2	RMSE
LSTM	59 min	15	0.5782	7.062
LSTM	120 min	30	0.5821	7.028
CNN-LSTM	60 min	30	0.5825	7.027
CNN-LSTM	40 min	20	0.5824	7.031

The results of the predictions using all features in the day-by-day prediction method are illustrated in Fig. 8. It is evident from Fig. 8 that each training iteration generates a distinct prediction on the test set, highlighting the inherent uncertainty in the neural network training process. The ensemble learning method we proposed effectively minimizes the variability present in individual model predictions. While some predictions may significantly deviate from the actual electricity spot prices, the ensemble approach, as detailed in (1), averages these forecasts to yield a more accurate final prediction. The model's capability to predict zero prices validates the "policy-versus-policy" zero prices forecasting strategy. The performance metrics of the ensemble learning for the day-by-day prediction method using all features are presented in Table III. This day-by-day prediction approach significantly enhances prediction accuracy compared to the one-time prediction method.

TABLE III. Ensemble learning metrics for the day-by-day prediction approach using all proposed features.

Metrics	
RMSE	R^2
5.66	0.729

B. Implications for EPF's confidence interval

In Fig. 8, it is noteworthy that when the ensemble learning prediction accuracy is high, the 30 individual predictions tend to highly overlap each other and also demonstrate high prediction accuracy, as shown for the period 2022.05.09-2022.05.14 in Fig. 8(b). Conversely, when the ensemble learning fails to predict the actual value accurately, the 30 individual predictions tend to diverge from each other without overlapping. The standard deviation of the 30 individual predictions at the same time point is calculated and used to compute the Person correlation efficient with the ensemble learning prediction error at the same time point. The correlation efficient value is 0.462, indicating a moderate correlation. By utilizing the above-mentioned phenomenon, a novel confidence interval estimation method for EPF can be developed in our future work.

V. CONCLUSIONS AND FUTURE WORK

This study has proposed an innovative EPF framework that utilizes multimodal data augmented by an ensemble learning technique. The CNN-LSTM model is superior over the LSTM model in both prediction accuracy and computation time. The logarithm transform of the pre-processing for the electricity spot price data has been shown crucial for EPF in the Japan electricity spot market. A "policy-versus-policy" strategy has been proposed to solve the zero-inflated regression problem, which halves the computation time compared with traditional two-stage method. By using the day-by-day prediction approach, the ensemble learning method achieved achieved a RMSE and (R^2) of 5.66 JPY/kWh and 0.728 over the test range of the full year 2022. An implication for a novel estimation method of the EPF's confidence interval using the ensemble learning approach has been proposed.

REFERENCES

- [1] J. Seel, D. Millstein, A. Mills, M. Bolinger, and R. Wiser, "Plentiful electricity turns wholesale prices negative," *Advances in Applied Energy*, vol. 4, p. 100073, Nov. 2021.
- [2] A. Shiri, M. Afshar, A. Rahimi-Kian, and B. Maham, "Electricity price forecasting using support vector machines by considering oil and natural gas price impacts," in *2015 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*. IEEE Explore, Aug. 2015, pp. 1-5.
- [3] K. Özen and D. Yıldırım, "Application of bagging in day-ahead electricity price forecasting and factor augmentation," *Energy Econ.*, vol. 103, p. 105573, Nov. 2021.
- [4] K. He, Y. Xu, Y. Zou, and L. Tang, "Electricity price forecasts using a curvelet denoising based approach," *Physica A: Statistical Mechanics and its Applications*, vol. 425, pp. 1-9, May 2015.
- [5] Z. Yang, L. Ce, and L. Lian, "Electricity price forecasting by a hybrid model, combining wavelet transform, ARMA and kernel-based extreme learning machine methods," *Appl. Energy*, vol. 190, pp. 291-305, Mar. 2017.

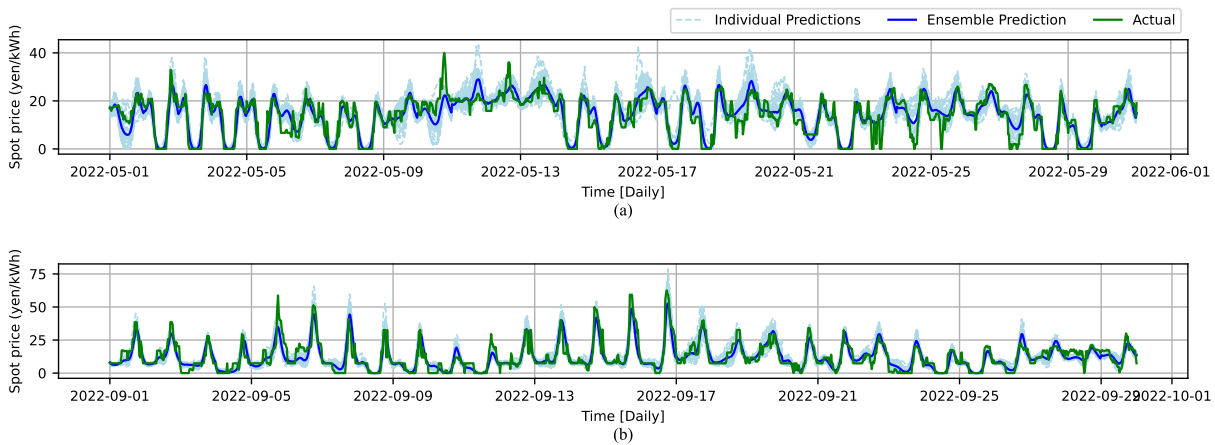


Figure 8. Individual and ensemble predictions for the test range of May and September in 2022, using all the proposed features and day-by-day prediction approach.

- [6] N. Chaâbane, "A hybrid ARFIMA and neural network model for electricity price prediction," *Int. J. Electr. Power Energy Syst.*, vol. 55, pp. 187–194, Feb. 2014.
- [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 Trans. Power Syst.*, vol. 20, no. 2, pp. 1035–1042, May 2005.
- [8] G. P. Girish, "Spot electricity price forecasting in indian electricity market using autoregressive-GARCH models," *Energy Strategy Reviews*, vol. 11–12, pp. 52–57, Jun. 2016.
- [9] Z. Wang, R. Matsuhashi, and H. Onodera, "Intrusive and non-intrusive early warning systems for thermal discomfort by analysis of body surface temperature," *Appl. Energy*, vol. 329, p. 120283, Jan. 2023.
- [10] Z. Wang, M. Mae, T. Yamane, M. Ajisaka, T. Nakata, and R. Matsuhashi, "Enhanced Day-Ahead electricity price forecasting using a convolutional neural Network–Long Short-Term memory ensemble learning approach with multimodal data integration," *Energies*, vol. 17, no. 11, p. 2687, Jun. 2024.
- [11] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, Nov. 2016.
- [12] J. Chen, G.-Q. Zeng, W. Zhou, W. Du, and K.-D. Lu, "Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization," *Energy Convers. Manage.*, vol. 165, pp. 681–695, Jun. 2018.
- [13] J. E. P. Exchange, "Day ahead market," <https://www.jepx.jp/en/electricpower/market-data/spot/>, 2023, accessed: 2023-8-19.
- [14] J. Organization for Cross-regional Coordination of Transmission Operators, "Menu," https://occtonet3.occto.or.jp/public/dfw/RP11/OCCTO/SD/LOGIN_login, 2023, accessed: 2023-8-1.
- [15] J. M. B. S. Center, "Numerical weather prediction model gpv-msm," <http://www.jmbc.or.jp/jp/online/file/f-online10200.html>, 2023, accessed: 2023-7-15.
- [16] "Google maps," <https://www.google.com/maps/@36.2932467,137.3408308,6z?entry=ttu>, accessed: 2024-5-3.
- [17] O. team, "OpenCV library," <https://opencv.org/>, 2023, accessed: 2023-8-19.
- [18] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, May 2015.
- [19] S. J. Taylor and B. Letham, "Forecasting at scale," *Am. Stat.*, vol. 72, no. 1, pp. 37–45, Jan. 2018.
- [20] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, May 2018.