Analysis of Photovoltaic Power Generation Forecast Accuracy and Battery Capacity in a Factory

Yui Koya[a], Xinjie Li[b], Yoshiharu Amano[c]

[a] Department of Applied Mechanics and Aerospace Engineering, Waseda University, Tokyo, Japan, fse_y.koya@ruri.waseda.jp
[b] Department of Applied Mechanics and Aerospace Engineering, Waseda University, Tokyo, Japan, vincent_li@fuji.waseda.jp
[c] Department of Applied Mechanics and Aerospace Engineering, Waseda University, Tokyo, Japan, yoshiha@waseda.jp

Abstract:
The adoption of variable renewable energy (VRE), which is uncertain and time-varying, is accelerating, causing the energy mix to change over time. Furthermore, the electricity and heat demand generated by factory equipment operation also changes over time. Matching energy supply and demand is essential for stabilizing equipment operation and decarbonizing the plant. In particular, utilizing battery storage as a balancing function is effective from the perspective of local energy production and consumption, as required by the GHG Protocol. In this study, we solve optimal factory equipment operation plans based on such as day-ahead photovoltaic (PV) power generation forecasts, which are based on both historical actual weather data, and forecast weather. Then we analyze the relationship between forecast uncertainty and battery (BT) capacity.
Keywords:
Battery; Factory energy management system; Optimization; Photovoltaic power; Temporal fusion transformer.

1. Introduction

1.1. Research backgrounds

The stable supply of CO2-free energy sources is globally demanded due to the impacts of global warming and other factors. Japan has declared carbon neutrality by 2050 [1]. The Seventh Strategic Energy Plan [2] sets a goal to make renewable energy (RE) sources such as solar and wind power the largest power source in the future, and to have RE account for approximately 40-50 % of the domestic power generation mix by 2040. The industrial sector remains a major source of CO2 emissions, and energy management has become a key theme in decarbonization. Optimization combining the introduction of RE, the utilization of batteries (BT), and demand control is drawing attention as a means to achieve both electricity cost reduction and CO2 reduction simultaneously. Energy management within factories is being standardized with a functional framework including monitoring, analysis, optimization, and instruction, according to IEC 63376 [3]. This standard assumes the integrated management of photovoltaic power (PV), BT, demand facilities, heat sources, and other components as a whole factory, requiring overall optimization rather than partial optimization for each facility. RE sources like PV depend on weather and solar radiation conditions and exhibit significant temporal variability. On the other hand, the factory's energy demand also varies greatly over time due to production schedules, equipment operating patterns, heat demand, and air conditioning loads. Therefore, aligning these fluctuations is key to overall optimization. Factories possess certain demand-side adjustment capabilities, such as product stock shelves and production schedule adjustments. However, these adjustment capabilities have limited adjustable quantities. In contrast, BT offer high flexibility in charging and discharging and are positioned as the most crucial adjustment capability.

1.2. Research objectives

Therefore, in this study, we solve the capacity design problem for a BT in a hypothetical factory and analyze the relationship between the accuracy of PV power generation forecasting and CO2 emissions when creating an equipment operation plan for the following day. Specifically, using predicted PV power generation, we calculate an operational plan for in-factory equipment that minimizes CO2 emissions from equipment

operation. At that time, we will vary the BT capacity and comparatively analyze the prediction reliability, BT capacity, stored energy, and CO2 emissions.

2. Target system

2.1. Factory configuration

The Fig. 1 shows the system targeted in this study. This study focused on an actual factory located in Japan. This factory has multiple buildings on its premises, along with shared on-site PV panels and BT. The energy management target of this study is one of the buildings within a factory (Building no.3) that has multiple buildings. In other words, the equipment operation plan to be solved with CO2 minimization as the objective function is only for Building no. 3. The electricity flowing into the factory is assumed to consist only of the power grid and on-site PV, with no procurement of RE from external sources. This factory is a manufacturing facility for the assembly of sensors and precision equipment, operating 24 hours a day on a two-shift system. It is characterized by high air-conditioning loads and a constant demand for electricity and heat even during the night. Building no. 3, which is one of the buildings within the factory, houses production equipment, air-conditioning systems, lighting, and electrical outlets. It is a two-story building, with the first floor primarily used for production equipment and the second floor mainly for office space. We are currently collecting hourly electricity data for lighting and outlets by floor, production electricity, Heating, Ventilation, and Air Conditioning (HVAC) electricity by building, and other electricity usage. This research utilizes these measured data.

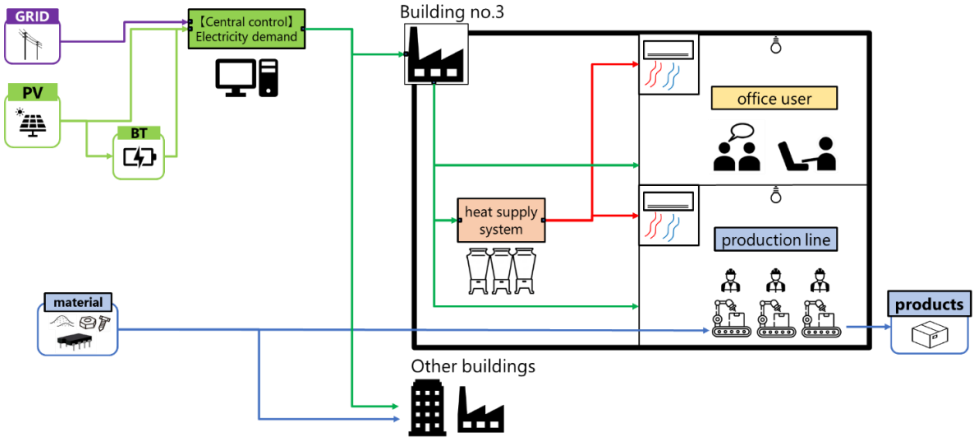

Fig. 1 The factory configuration

2.2. PV panel, battery (BT), and heat sources

The total shared PV panel capacity for the entire factory is 2,048 kWh. Similarly, the total shared BT capacity for the entire factory is 2,400 kWh, with a charge/discharge rate of 400 kW. However, since the energy management target of this study is only Building no.3, the condition of using shared PV and BT (shared among multiple buildings) solely for Building no.3 is not realistic. Therefore, the PV available for Building no.3 will be determined by applying the ratio of Building no.3's electricity consumption to the total measured electricity consumption of the entire factory. For heat sources, Building no.3 has three heat pumps (HP) of the same performance, each with a rated capacity of 540 kW. This means that Building no.3's heat demand is met by these HPs converting electricity into heat. However, there are no thermal storage facilities such as water or ice storage tanks. In this study, using the currently measured electricity data described in the previous section, we will formulate an operation plan for the three HPs (the heat sources) and schedule the electricity supply and demand, including the BT. We will also calculate the amount of CO2 emitted when operating according to that plan. Since there is no measured data for the PV power generation to be input as an exogenous variable, we will use open-source test data to make predictions using AI. Details will be described in Chapter 3.

2.3. Demand of Building no.3

The evaluation period is a single day during the summer, and the time slice is one hour; in other words, we calculate an operational plan from 0:00a.m. to 11:00p.m. Fig. 2 shows the power consumption data for Building no.3, categorized by item, measured during the evaluation period. Electricity demand is categorized into four areas: production, air conditioning, lighting/outlets, and others, all of which are covered by the power grid and PV. However, PV power specifically can be supplied directly to the demand side or stored in BT to be discharged at the required time.

2.4. Configuration of Energy Management System (EMS)

2.4.1. Evaluation period

The evaluation period for the optimization calculation is one day, July 18, 2024. This means the heat pump is operating in cooling mode. This day is a weekday, and production is taking place during the daytime. Optimizing the medium- to long-term schedule is outside the scope of this discussion.

2.4.2. Data used

This section explains the parameters that are input into the optimization calculation as exogenous variables. The exogenous variables to be input are the CO2 emission factor of the power grid, outside air temperature, electricity and heat demand and on-site PV power generation. The CO2 emission factor is an index indicating the amount of CO2 emitted to generate electricity, which was created from the power supply and demand performance data of electric power companies [4]. For outside air temperature, data from the Japan Meteorological Agency [5] was used. Electricity demand is the value explained in the previous section, and heat demand was calculated by multiplying the power consumption of air conditioning by a Coefficient of Performance (COP) equivalent of 3. These are shown in Fig. 3 through Fig. 5.

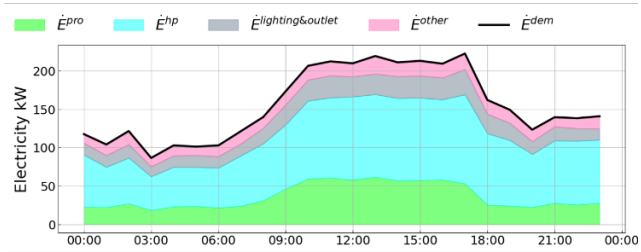

Fig. 2 Breakdown of electricity demand in Building no.3

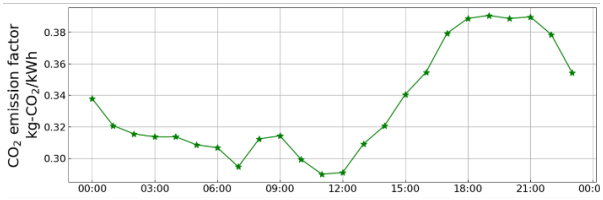

Fig. 3 CO2 emission factor

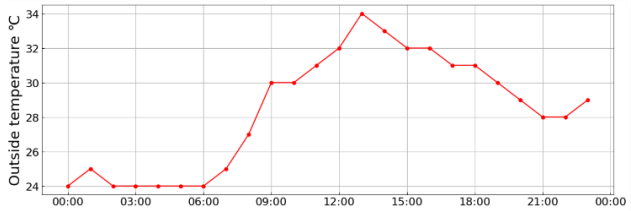

Fig. 4 Outside temperature

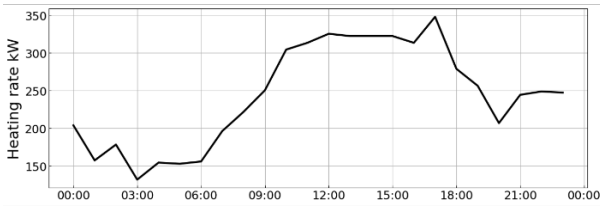

Fig. 5 Heat demand

Regarding PV power generation, we will use values predicted using AI. Since PV power generation significantly impacts facility operation plans and the resulting CO2 emissions, we will analyze the impact of forecast accuracy. This will be discussed in Chapter 3.

3. Photovoltaic power generation prediction using Temporal Fusion Transformer (TFT)

3.1. Photovoltaic power generation prediction within Industrial Energy Management Systems

The solar energy is the most popular and fastest-growing variable renewable energy (VRE) resource to replace traditional power resources. However, due to the uncertainty and time-varying of photovoltaic (PV) power and the ever-changing energy demand of the factory's production, balancing energy supply and demand is essential for stabilizing equipment operation. Therefore, a reliable PV generation prediction is critical. In this study, we solve optimal factory equipment operation plans based on a designed hourly day-ahead PV power generation prediction, which are based on both historical actual weather data and forecasted weather. The results of the prediction will be introduced as a variable to analyze the battery capacity.

3.2. The advantage of the Temporal Fusion Transformer (TFT) in renewable energy forecasting

For the past decades, renewable energy forecasting has been largely based on statistical models, such as the Auto-Regressive Integrated Moving Average (ARIMA) model, which is the most widely used. However, as these statistical approaches reach their limits in modelling non-linear spatiotemporal relationships, conventional machine learning algorithms have been utilized, such as Random Forests and Extreme Gradient Boosting (XGBoost). Unfortunately, although tree-based models provide consistent performance, they lack an

inherent understanding of sequential order and treat time-series forecasting as a regular regression task. In this way, they are difficult to predict complex situations such as PV power prediction, which involves a large number of meteorological features. Recently, with the quick development of deep learning, models such as the Long Short-Term Memory (LSTM) network have attracted more attention. However, conventional deep learning architectures face two core limitations: firstly, they struggle to consistently and coherently process heterogeneous data types; secondly, they function as black-box models, with little way to interpret which features informed a given prediction [6-8].

To overcome the constraints mentioned above, this study adopts the Temporal Fusion Transformer (TFT) model, which is an innovative attention-based deep neural network first proposed by Lim [7]. The TFT has a great effect in multi-horizon time-series forecasting, such as PV generation prediction [8]. Because it directly considers the inherent heterogeneity of real-world datasets. At the same time, compared to other deep learning models, TFT has great levels of interpretability. TFT's architecture combines the strengths of multiple neural network designs: it not only uses recurrent layers for local processing to retain critical short-term information but also takes advantage of multi-head self-attention to capture long-term dependencies across the full forecasting horizon [7,8]. And remarkably, the TFT combined Variable Selection Networks (VSNs) to automatically filter out noisy or irrelevant inputs at each time step, directing the model's computational resources to the most impactful features. It also employs Gated Residual Networks (GRNs) to suppress non-essential model components, supporting more stable and reliable convergence during training [7,8]. In addition, by using a quantile loss function, the TFT is able to generate probabilistic forecasting intervals, rather than a deterministic single-point estimate. This feature is critical for predicting PV power in real-world production environments. Limiting the PV power generation of day-ahead futures to a reasonable forecast range can provide EMS with sufficient flexibility to cope with uncertainties in real-world environments.

3.3. Dataset characterization and data processing

3.3.1. Dataset characterization

To investigate the predictive performance of the TFT model in a real-world environment, the PV power dataset in this study is collected at one-hour granular generation for a continuous two-year period from January 1, 2022, to December 31, 2023. Historical PV power generation and actual meteorological data, such as temperature, wind speed, etc., are collected from the PHOTOVOLTAIC GEOGRAPHICAL INFORMATION SYSTEM [11] by the European Commission at one PV plant in Japan supply for the studied factory. However, traditionally, prediction models trained exclusively on perfect historical weather observations suffer from performance degradation when deployed, as they are forced to rely on imperfect day-ahead weather forecasts. To improve this predictive performance, this study incorporates historical weather forecast data as features in Photovoltaic (PV) generation prediction during the training process. By directly utilizing historical weather forecasts during the training phase, the TFT can implicitly learn to identify and adapt to the inherent systematic biases and error patterns of the specific meteorological forecasting system. We aimed that this dataset design could eliminate the feature mismatch between the training and application phases, and improve the accuracy, reliability, and robustness of day-ahead PV power predictions. Historical weather forecasts for the same time period and resolution as the power plant's location were also incorporated into the dataset from the Historical Forecast API by open-meteo.com [12]. To preserve the temporal continuity, cyclic features for time components were also created as features. Specifically, the Euclidean distance between hour 23 (23:00) and hour 0 (00:00) is 23, whereas the actual temporal distance is only 1 hour [17]. Therefore, the projection of temporal components onto a two-dimensional unit circle is necessary. An example is as follows:

$$hour_{sin} = sin\left(\frac{2\pi \cdot hour}{24}\right) \quad (1)$$

Overall, the dataset contains three categories of feature vectors: time-varying unknown reals (PV generation information), time-varying known reals (meteorological information), and time-varying known categorical (temporal information).

3.3.2. Target variable analysis

The target variable to be predicted by the model is PV production, representing the actual solar power generation. Exploratory statistical analysis of the target variable shows a highly skewed, zero-inflated distribution. Taking the hourly power generation data of this PV station throughout 2022 as an example shown in fig. 6, it has a mean generation value of 269.09 kW and a standard deviation of 397.15 kW. The minimum value and the 50th percentile (median) are exactly 0.00 kW, reflecting the absence of solar irradiance during nighttime hours, while the 75th percentile is 487.32 kW. The absolute maximum peak generation output reaches 1,559.34 kW.

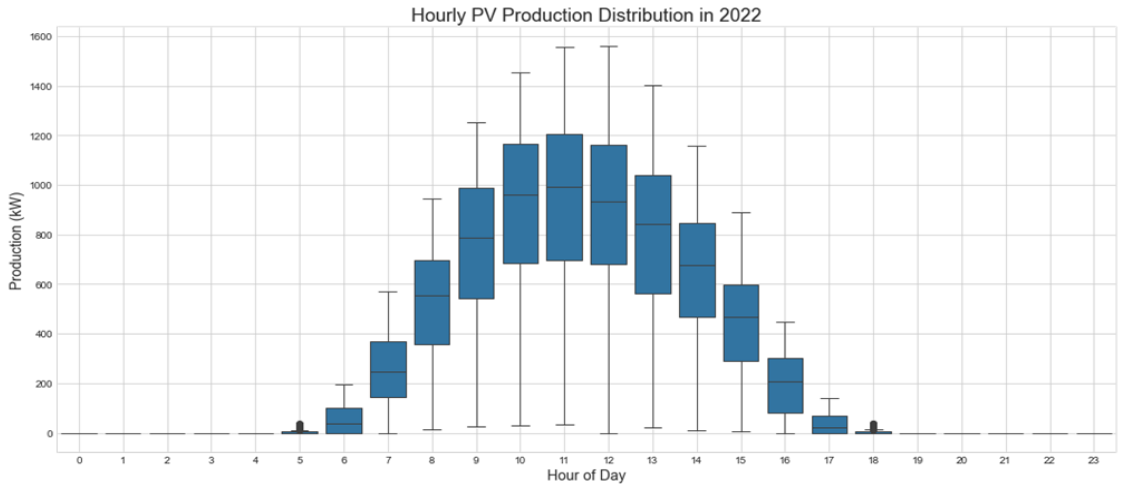

Fig. 6 Historical hourly PV production distribution

3.3.3. Dataset splitting, normalization, and scaling

To ensure the predictive performance and preserve the temporal character, the dataset is split into three distinct subsets by time period: the training set contains the entire year of 2022 data, the validation set includes the first half of 2023 (January to June), and the testing set covers the second half of 2023 (July to December).

To ensure stable model convergence while preventing data leakage, two independent scalers are implemented. The first scaler is fitted on the training set to normalize the input features, including meteorological data and cyclical time variables, into a [0, 1] range. The second scaler independently normalizes the target variable (photovoltaic production), to expose the absolute peak variances directly to the model's loss function.

To design the sliding window sequence generation feed into the TFT model, we apply time series analysis for sequence length selection and map a 168-hour (7 days) historical lookback to a 24-hour (1 day) prediction horizon. A dynamic feature concatenation is employed in each prediction step.

3.4. Temporal Fusion Transformer model design and architecture

The forecasting framework is mainly based on a customized Temporal Fusion Transformer (TFT) architecture, which is designed for day-ahead hourly photovoltaic (PV) power generation prediction with uncertainty quantification. The pipeline is structured as follows: processing a 168-hour (7-day) historical lookback window to predict a 24-hour future horizon by the TFT model. The core TFT network takes advantage of a Variable Selection Network (VSN), Gated Residual Networks (GRN), and multi-head temporal attention [7,8]. The output of the TFT contains specific prediction bounds, 16 %, 50 %, and 84 % quantiles, to achieve a 68 % confidence interval of prediction. An exponentially weighted multi-quantile loss is used during model training to overcome underestimation during daytime peak generation hours. Finally, the TFT prediction results are calibrated by post-processing. The design model architecture is shown in figure 7.

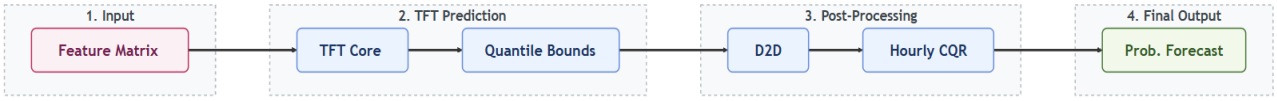

Fig. 7 Day-ahead PV production prediction model architecture

3.4.1. Hyperparameter configuration

The optimal hyperparameter space is defined dynamically via Optuna [13] in Python. The optimal model with follows configuration: 32 hidden state dimension, 8 parallel attention heads, and 2 transformer encoder layers. The dropout rate is 0.1457. The network is trained via the Adam optimizer with weight decay 1.2191^{-5} and learn rate 0.0051.

3.4.2. Input layer and feature embeddings

At each time step t, the input matrix is formed by concatenating the exogenous variables (weather and time) and the historical PV production. Continuous features are projected into a uniform high-dimensional latent space via linear embeddings; and time features are embedded into a temporal vector. Two embedding are fused by additive fusion.

3.4.3. Variable Selection Network (VSN)

The Variable Selection Network (VSN) dynamically weighs the predictive relevance of each input feature at every time step, filtering out noisy variables [7]. Let the encoded input of the j-th variable at time t be $\xi_t^{(j)}$. The inputs are concatenated into a flattened vector Ξ_t. A Gated Residual Network (GRN) combined with a SoftMax function generates the variable selection weights v_{x^t} based on the inputs and an optional static context vector C_s:

$$v_{x^t} = SoftMax\left(GRN_{weights}(\xi_t^{(j)}, C_s)\right) \tag{2}$$

The final processed feature representation $\tilde{\xi}_t$ is computed as an element-wise weighted sum of the individual transformed features:

$$\tilde{\xi}_t = \sum_{j=1}^{m_x} v_{x^t}^{(j)} GRN_j\left(\xi_t^{(j)}\right) \tag{3}$$

3.4.4. Gated Residual Network (GRN)

The GRN provides adaptive non-linear processing only at necessary, which grants the model the flexibility [7]. GRN is defaulting to an identity mapping to prevent gradient vanishing during training process. For a primary input a and an optional context c:

$$DRN(a, c) = LayerNorm\big(a + GLU(\eta)\big) \tag{4}$$

Where η is intermediate layers, and the Gated Linear Unit (GLU) selectively suppresses irrelevant features pathways using a sigmoid activation (σ) and elementwise Hadamard product ($\odot$) for weights W and biases b as follows:

$$DLU(\gamma) = \sigma\big(W_{1\gamma} + b_1\big) \odot \big(W_{2\gamma} + b_2\big) \tag{5}$$

3.4.5. Multi-Head Temporal Attention Layer

To capture long-range temporal features within the 168-hour historical window, the temporal sequence projects into Query Q, Key K , and Value V matrices. The scaled dot-product attention for each head is:

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{6}$$

where d_k is the scaling factor to prevents the activation functions flatten out and makes gradients close to zero. Multiple heads allow the model to attend to different sub-spaces at the same time, not only capturing both short-term fluctuations but also long-term cycles [7].

3.4.6. Weighted Multi-Quantile Loss

Standard quantile loss penalizes errors based on pinball loss, which equally penalizes regardless of the time of day. However, PV generation presents extreme skewed distribution (zero at night, massive variance at noon). To align the model with PV generation patterns, a dynamically exponential penalty weight w_t is introduced to heavily penalize errors during peak generation hours:

$$w_t = \exp(y_t \cdot \beta) \tag{7}$$

where β is setting to a target hyperparameter and y_t is true penalty weight magnitude. And the model is optimized using the dynamically Weighted Pinball Loss across all target quantiles $q \in Q$:

$$\mathcal{L} = \sum_t \sum_{q \in Q} w_t \cdot Max\big(q(y_t - \hat{y}_t), (q-1)(y_t - \hat{y}_t)\big) \tag{8}$$

3.4.7. Complete prediction and post-processing flow

To prevent systematic errors in deep neural network models when predicting constantly changing real-world data distributions, and to avoid the need for computationally intensive daily model retraining, a series of post-processing steps are introduced into the prediction model results:

Step 1: D2D Post-Processing: a Dynamic Sliding Window deterministic-to-deterministic (D2D) [9] corrector algorithm is executed as the first stage of the post-processing pipeline: the algorithm operates by continuously monitoring a rolling historical window of size W (defined as W = 3 days) at each day d in W and each hour of day h. the unweighted hourly mean error C_h is calculated from the deterministic 50-th percentile prediction:

$$C_h = \frac{1}{W}\sum_{j=1}^{W}\big(y_{d-j,h} - \hat{y}_{d-j,h}^{50\%}\big) \tag{9}$$

This correction vector is sequentially added to the future day's quantile predictions to shifting the intervals to correct for TFT's potential systematic biases without any possibility of forward data leakage.

Step 2: Hour-Conditional Conformalized Quantile Regression (CQR) Calibration: While Weighted Multi-Quantile loss and D2D correction ensure high deterministic accuracy, standard quantile regression methods cannot guarantee enough sample coverage of prediction. In EMS, certainty is critical. If a model outputs a 68 % prediction interval, exactly or higher 68 % of actual future observations is expected within that interval, regardless of the underlying data distribution. This absolute mathematical certainty can be achieved by using Conformalized Quantile Regression (CQR) as the final calibration step. CQR fundamentally combines the locally adaptive properties of standard quantile regression with the distributional independence and finite sample coverage guarantee of conformal prediction theory [10]. This provides strong assurance for the prediction results in battery capacity analysis and optimization analysis. CQR is computed conditionally per hour ($h \in \{0, 23\}$). At each hour, a non-conformity score is defined as:

$$E_{i,h} = \max\left(\hat{y}_{lower,h} - y_{i,h}, y_{i,h} - \hat{y}_{upper,h}\right) \quad (10)$$

Based on target fixed confidence interval (e.g., 68 %) and non-conformity score $E_{i,h}$, a calibration scalar $\hat{q}_h$ is calculated at each hour. Nighttime scalars are strictly clamped to zero to maintain tight bounds, and daytime D2D-corrected bounds are then symmetrically expanded or shrunk as follows:

$$\hat{y}_{lower}^{calibrated} = \max(\hat{y}_{lower}^{D2D} - \hat{q}_h, 0) \quad (11)$$

$$\hat{y}_{upper}^{calibrated} = \max\left(\hat{y}_{upper}^{D2D} + \hat{q}_h, 0\right) \quad (12)$$

3.5. Predictive performance and evaluation metrics

After the model training, the model's predictive performance was evaluated using the test dataset. Three standard statistical error metrics were used to evaluate the model's accuracy at the 50th quantiles: Mean Absolute Error (MAE), which measures the average absolute deviation; Root Mean Square Error (RMSE), which penalizes larger errors by calculating the square root of the mean square deviation; and R-squared (R^2), representing the proportion of actual data variance explained by the predicted values. The results show robust model performance: MAE of 80.1442 kW, RMSE of 166.6757 kW, and R^2 of 0.0651. Considering the physical parameters of the target PV plant (average power generation of 269.09 kW, standard deviation of 397.15 kW, and maximum peak power of 1559.34 kW), the MAE of 80.1442 kW indicates a sufficiently small hourly deviation between the predicted and actual values. Although the huge fluctuation of solar power generation typically reduces the R^2 value, the model remains suitable for day-ahead prediction. Such accuracy provides a robust and reliable reference for downstream optimization algorithms. Furthermore, this study employs probabilistic evaluation metrics to evaluate the model's uncertainty range, whether the 16th to 84th quantiles can cover the actual power generation during peak periods. Probabilistic evaluation metrics contain the Probability of Predicted Interval Coverage (PICP), Average Coverage Error (ACE), and Prediction Interval Normalized Average Width (PINAW). Evaluation based on the entire test set average (6-month interval), PICP and ACE scored 0.8065 and 0.1265 respectively. High PICP and low ACE indicates that the percentage of actual observations falling within the prediction bounds safely exceeds 68% target and avoids potential “overconfidence”, which that narrow intervals lead to critical missed predictions during the PV production. Regarding prediction sharpness, PINAW is 0.1974. This reveals that although the prediction model tends to safely coverage, it does not achieve this through the unconstrained inflation of interval boundaries. It proves that the model balanced high safety margin for PV production and keep the model remain highly informative. A one-week sample prediction result and an average probabilistic evaluation metrics result for the complete test set shows in figure 8(a) and 8(b), respectively. Meanwhile, we selected an extreme weather environment, where the weather information input to the prediction model on that day has the maximum variance for testing. The results are shown in figure 8(c), demonstrating that the model is highly likely to converge the actual power generation to the model's prediction range. Finaly, to meet the ultimate goal of this study, optimize energy management system of Building no.3 rather than entire factory, a spatial mapping is applied to final prediction results to estimate the PV energy supply for Building no.3. The final total entity predictions range (Lower, Median, Upper) are multiplied by a historically 24-hour dynamic ration derived from electricity consumption of Building no.3 to the total measured electricity consumption of the entire factory.

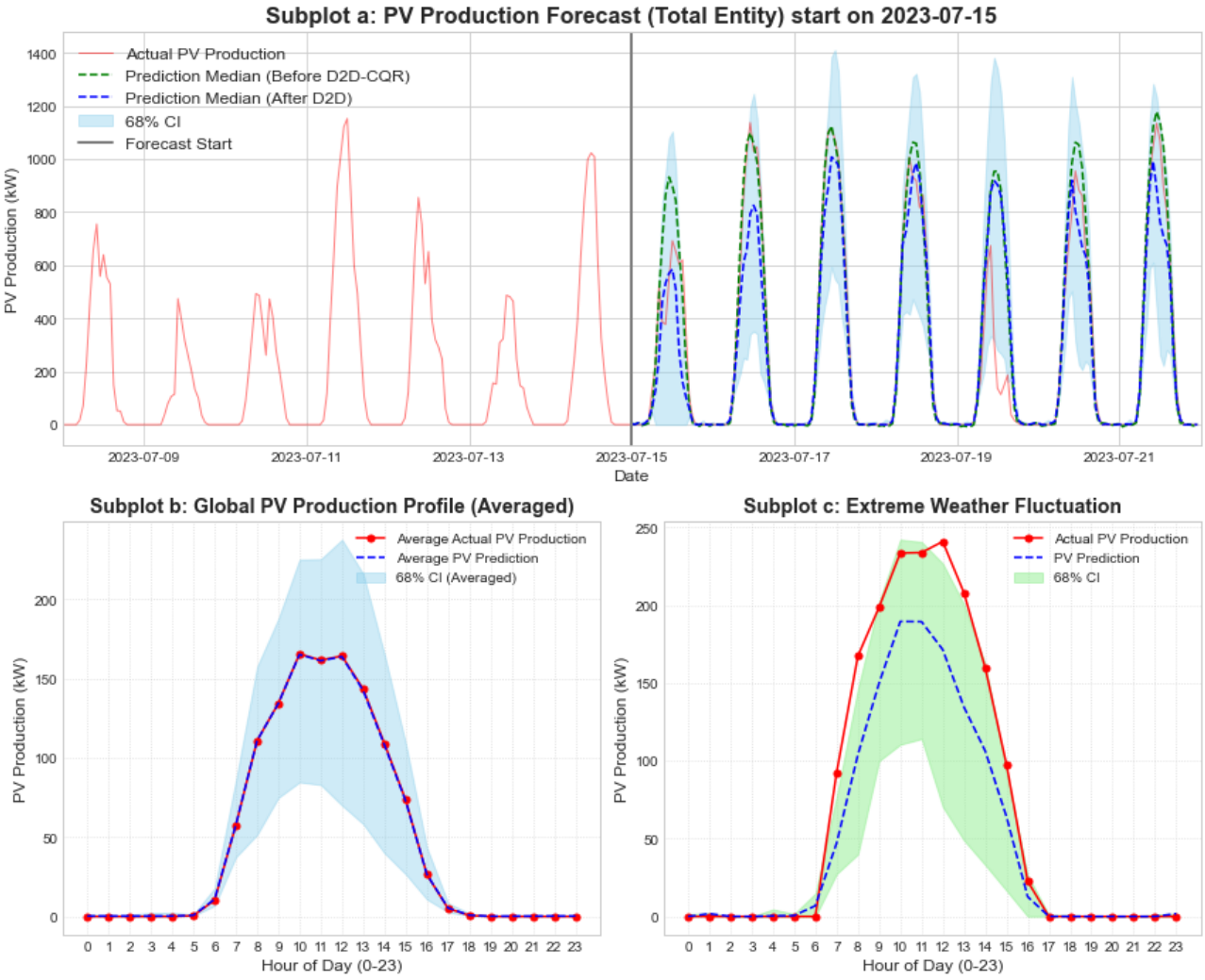

Fig. 8 Day-ahead PV production prediction results

4. Optimization models

4.1. Problem formulation overview

This section briefly describes the overall objective and structure of the model. The objective function is to minimize the total CO2 emissions of Building no.3 resulting from the operation of its facilities. The scope of the model covers the electricity and heat demand of Building no.3. The controllable facilities include the factory's shared BT, output from PV panels, the power grid, and three heat pumps dedicated to Building no.3. Inputs include outside temperature, PV generation, the CO2 emission factor for the power grid, and the performance and efficiency of the HPs and BT, as well as demand data. We used the actual recorded values from the Japan Meteorological Agency [14] for the outside temperature. Outputs consist of facility operation plans, such as BT charging/discharging, HPs operation, and the purchase of the power grid.

4.2. Objective function

The objective function is to minimize CO2 emission of one day. The amount of daily CO2 generated by facility operation is calculated by multiplying the power grid by its CO2 emission factor. The hourly CO2 emission factor is calculated from the electricity supply and demand data published by power companies [15] and the CO2 emission factor for each power generation method [16], and is expressed by the following Eq. (13):

$$\text{objective function} = \text{minimize}(\Sigma_{t\in \boldsymbol{T}} C_{co2}(t)\dot{e}^{g}(t)\Delta t) \tag{13}$$

4.3. Solution

In this study, the operational plan for the target system is formulated as a Mixed Integer Linear Programming (MILP) problem. MILP is a subclass of Mixed Integer Programming, which refers to optimization problems that include both integer and continuous decision variables. In MILP, both the objective function and all constraints are linear. In optimal operational planning problems for energy systems, integer variables are used as 0-1

variables (binary variables) to formulate the start/stop operations of equipment within the system and variables representing their status. The COP is defined as the ratio of energy input to energy output for the HP, expressed by the following Eq.(16) using input electricity $\dot{e}^{hp}$ and heat supplied $\dot{q}^{hp}$:

$$COP = \frac{\dot{q}^{hp}}{\dot{e}^{hp}} \tag{16}$$

Here, the load factor is the output heat quantity normalized by the rated value $\dot{Q}^{hp}$. COP is a nonlinear function of the load factor, and its functional form varies with the cooling water inlet temperature. To express this COP characteristic and formulate it as a MILP problem, this COP curve is approximated by a piecewise linear function. In this study, the load factor range from 0.1 to 1 was divided into three intervals.

For the optimization calculation solver, Gurobi version 11.0.3, a general-purpose solver, was used.

5. Results and analysis

5.1. BT capacity design problem

As stated previously, in this study, since the BT and PV panels are shared throughout the factory, evaluations were conducted by simultaneously varying the capacity and charging speed while maintaining a relationship where the charging speed is $\boldsymbol{1/6\ C}$ relative to the battery capacity $\boldsymbol{C}$.

First, We will present the results regarding the BT capacity design problem. The actual PV generation data for the 177-day test period used for PV generation forecasting was classified into 3 patterns—"high, medium, and low"—based on the generation volume. The results showed that there were 104 days (58.8 %) with high power generation, 58 days (32.8 %) with medium generation, and 15 days (8.4 %) with low generation. Using the average value of each generation volume as a representative day, we performed optimization calculations while varying the BT capacity and compared the daily CO2 emissions. Fig. 9 shows the average power generation for each classification, and Fig. 10 shows the relationship between BT capacity and CO2 emissions, with the CO2 emissions in the case without a BT set to 100%. These results confirm that the higher the PV power generation, the greater the CO2 reduction effect from introducing a BT. Furthermore, even if the BT capacity is increased, it was found that CO2 emissions saturate at approximately 1200 kWh for days with high generation, 900 kWh for days with average generation, and 600 kWh for days with low generation, indicating that increasing the capacity beyond these points yields almost no additional reduction effect. Based on these results and the frequency of power generation occurrences, it can be determined that the optimal BT capacity for Building no. 3 is approximately 900 to 1,200 kWh. This corresponds to about 25–33 % of the total daily electricity demand. For a building like Building no. 3, which has a demand structure where nighttime demand is about 1/2 to 2/3 of daytime demand, this capacity range can be considered the most rational.

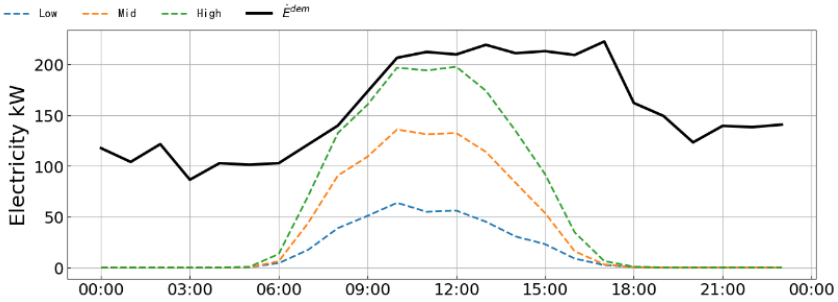

Fig. 9 Each average PV power generation and demand

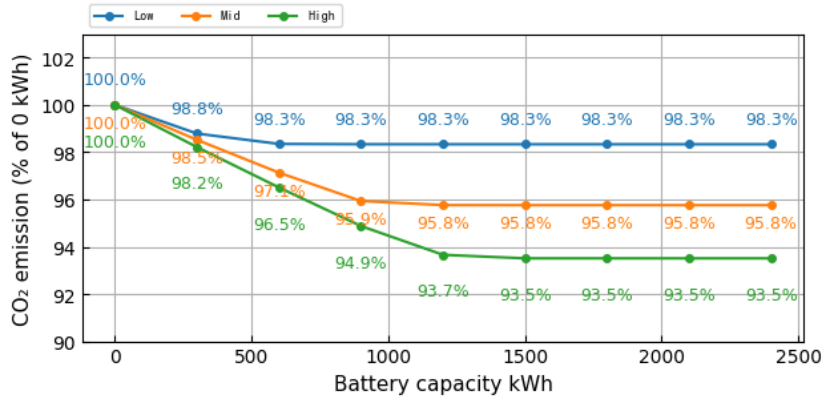

Fig. 10 Relationship between BT capacity and CO2 emissions for different power generation levels

5.2. Accuracy of PV power generation forecasts and problems with equipment operation planning

Next, We will describe the relationship between the accuracy of PV power generation forecasting and facility operation planning.

First, for a 177-day test period, we extracted the days where the MAE was at its best. When applying a 20% tolerance range to that error and the prediction interval (MAE x 1.2, width x 1.2), 39 days fell within this range. This indicates that while the current forecasting model has a certain level of accuracy, there are not a few days where the error becomes significant. Furthermore, an analysis of meteorological conditions for the top 10% of days with the largest forecasting errors revealed clear differences in precipitation, cloud cover, and solar radiation. From this, it is considered that the current model is prone to decreased accuracy under unstable weather conditions such as rain or cloudiness, and that forecasting errors tend to increase sharply on such days. Additionally, comparing the forecasted values that resulted in the minimum MAE against actual values over the 177 days, the median was closest on 102 days, the upper limit was closest on 46 days, and the lower limit was closest on 29 days. While days where the median was closest to the actual results account for the majority, it was confirmed that for over 40% of the days, either the upper or lower limit was closer, indicating that forecasting errors exist within a certain range.

Next, we developed facility operation plans using the forecasted values and compared the CO2 emissions when recalculating based on actual values using those plans. The BT capacity was set to 1,200 kWh, and data from days where the median forecast was the best was used. When an operation plan was created using actual values assuming 100% prediction accuracy, CO2 emissions were 714.5 kg, whereas when planned using median predictions, emissions were 753.8 kg, indicating that prediction errors increased CO2 emissions. However, when recalculated with actual values based on the facility operation of this plan, CO2 emissions were 714.8 kg, which was reduced to a nearly ideal value. Fig. 11–13 show the operation plans. This is because, in the re-planning based on actual values, the amount of electricity stored during the day increased by considering the day's weather and PV generation, and by reaching full charge earlier, priority could be given to discharging during time periods when the CO2 emission factor was high in the evening and beyond.

Similarly, we verified days when the median prediction was not the best, i.e., days when actual results were significantly lower than predicted. CO2 emissions were 1173.1 kg assuming 100% prediction accuracy and 754.2 kg when planned using median predictions; however, when recalculated with actual values, the emissions became 1173.1 kg, achieving the same result as the ideal operation. On this day, the plan predicted reaching a full charge during the day, but because PV generation was actually insufficient, the operation shifted to using all daytime PV for electricity storage, which consequently minimized CO2 emissions.

Based on the above results, it has become clear that even on days classified as having "best" predictions, prediction errors affect facility operation and CO2 emissions in the current prediction model. However, it has been shown that CO2 emissions can be brought closer to the ideal value by switching to an operation that prioritizes full charging during the day according to the day's weather and PV generation, while using the plan created with predicted values as a baseline.

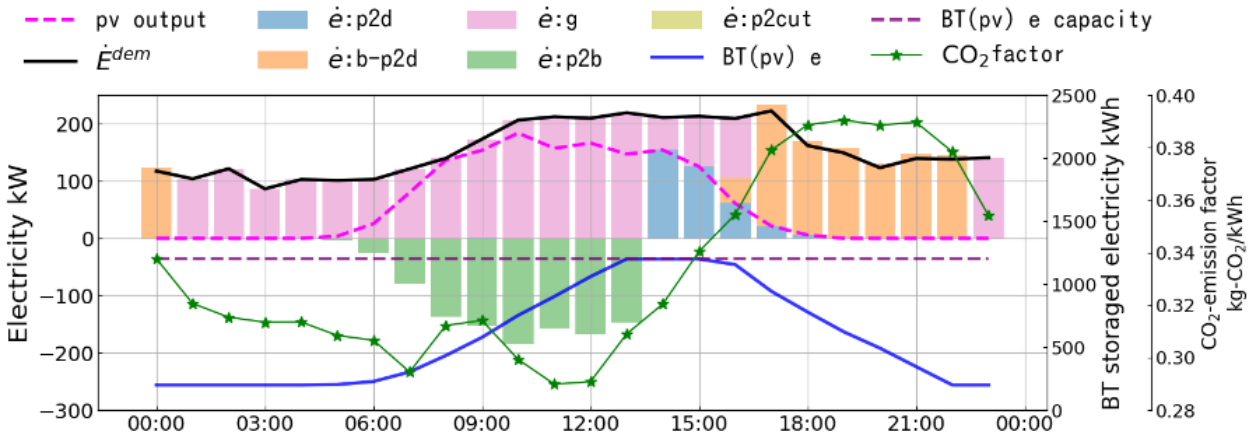

Fig. 11 Operational plan using actual PV generation data

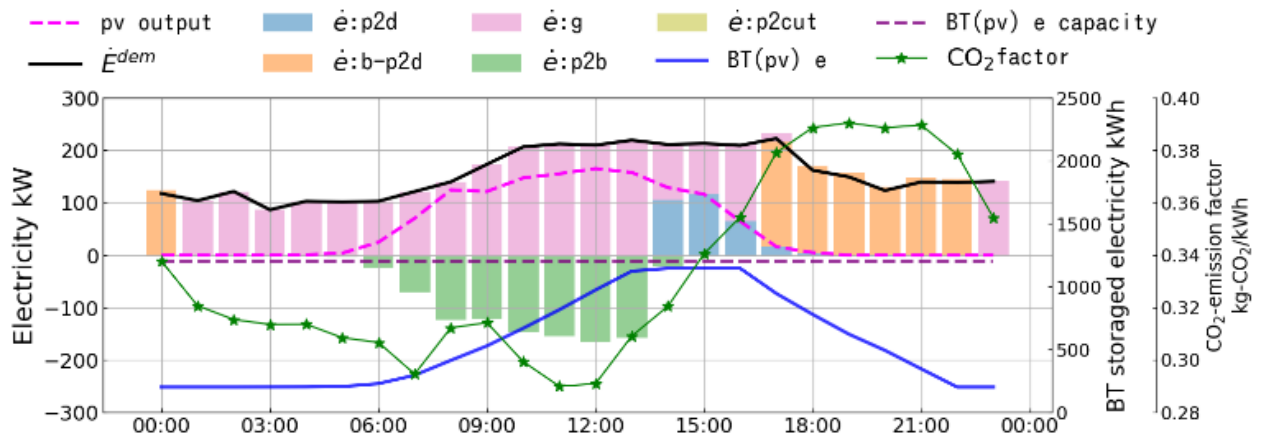

Fig. 12 Operational plan using predicted median PV generation data

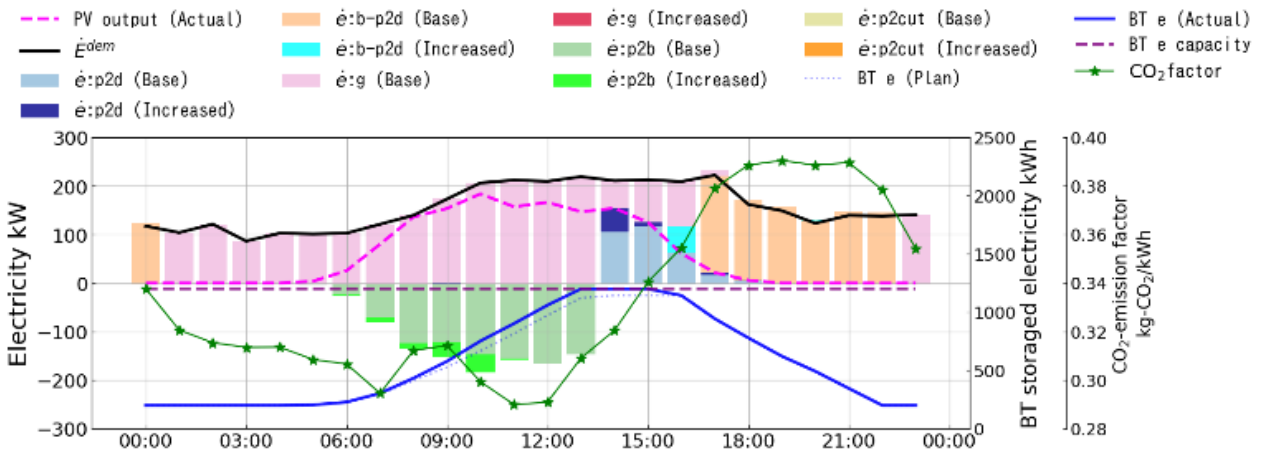

Fig. 13 Recalculated operation plan based on predicted values

6. Conclusion and future prospects

6.1. Conclusion

The purpose of this study was to target a factory equipped with PV generation facilities and storage batteries shared among multiple buildings, and to (1) clarify the appropriate BT capacity for a specific building within the factory, and (2) analyze how the accuracy of PV power generation forecasts affects CO2 emissions through the following day's facility operation plan. As a result, the following findings were obtained:

- This research proposes a quantile regression prediction model based on the TFT model architecture. This model performs D2D statistical correction on the prediction results and uses the CQR algorithm to scale the prediction intervals during peak hours. While this design can estimate the next day's PV power generation with high reliability, prediction errors increase under unstable weather conditions such as cloudy or rainy days.
- For 24-hour operational buildings that have about half the electricity demand at night as they do during the day, the optimal BT capacity is approximately 33% of the total daily electricity demand.
- In the current prediction model, prediction errors exist, and it does not necessarily capture actual values precisely. However, it was revealed that even when prediction errors are present, CO2 emissions can be suppressed to almost the same level as the ideal value by establishing an operational policy that prioritizes full charging of the BT during the day and discharging during the evening hours when the CO2 emission factor is high, based on the operational plan formulated using the predicted values, while adjusting for the day's weather and actual PV power generation.

6.2. Future prospects

Regarding future prospects, two main directions can be identified. First, it is necessary to enhance operational flexibility to address uncertainties in demand and PV generation. Even if a day-ahead operational plan is

formulated based on forecast bounds, deviations in actual conditions can make fixed plans suboptimal. Therefore, establishing adaptive operational rules that revise the plan according to real-time PV generation is essential. For example, when PV generation is lower than forecast, prioritizing battery charging for nighttime peak shifting can improve CO2 performance, whereas when it is higher than forecast, increasing both battery charging and direct supply to demand is effective. In addition, it is important to analyze the relationship between the required reserved battery capacity and forecast accuracy, as prior studies have shown the effectiveness of pre-charging in small-scale systems, although such analyses remain limited in industrial applications.

Second, with the further expansion of RE, surplus generation—particularly during daytime—is expected to increase, leading to a growing need for upward demand response. In this context, integrating not only BT but also thermal storage systems can provide additional flexibility by shifting energy consumption and effectively utilizing surplus RE. Evaluating the impact of such multi-energy storage systems in industrial settings is an important direction for future research.

Acknowledgments

This work was supported by the Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), and the third period of the SIP “Smart energy management JPJ012207 system” Grant Number JPJ012207 (funding agency: JST). We express our gratitude to the operators and related organizations.

References

[1]Ministry of Economy, Trade and Industry of Japan. Green Growth Strategies Associated with 2050 Carbon Neutrality. Available at:< https://www.meti.go.jp/policy/energy_environment/global_warming/ggs/pdf/green_honbun.pdf>.
[2]Agency for Natural Resources and Energy, Ministry of Economy, Trade and Industry. The 7th Strategic Energy Plan. 2025. Available at:<https://www.enecho.meti.go.jp/>.
[3]International Electrotechnical Commission. IEC 63376: Industrial facility energy management system (FEMS) – Functions and information flows. Geneva, Switzerland: IEC; 2023.
[4]Kansai Electric Power Company. Denki Yohou (Electricity Forecast). Available at:<https://www.kansai-td.co.jp/denkiyoho/area-performance/index.html> [accessed 24.3.2026].
[5]Japan Meteorological Agency. Past Weather Data Download. Available at:<https://www.data.jma.go.jp/gmd/risk/obsdl/> [accessed 24.3.2026].
[6]Mollasalehi A, Farhadi A. Solar and Wind Power Forecasting: A Comparative Review of LSTM, Random Forest, and XGBoost Models. arXiv preprint arXiv:2509.24059. 2025.
[7]Lim B, Arık SÖ, Loeff N, Pfister T. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. International Journal of Forecasting. 2021;37(4):1748-1764.
[8]Miguel López Santos, Xela García-Santiago, Fernando Echevarría Camarero, Gonzalo Blázquez Gil, Pablo Carrasco Ortega. Application of Temporal Fusion Transformer for Day-Ahead PV Power Forecasting. Energies, MDPI, vol. 15(14), pages 1-22, July,2022
[9]Yang D., Kleissl J., Solar Irradiance and Photovoltaic Power Forecasting. Boca Raton, USA: CRC Press; 2024.
[10]Romano Y., Patterson E., Candès E.J., Conformalized quantile regression. In: Wallach H., Larochelle H., Beygelzimer A., d'Alché-Buc F., Fox E., Garnett R., editors. Advances in Neural Information Processing Systems 32 (NeurIPS 2019); 2019 Dec 8-14; Vancouver, Canada. Curran Associates, Inc.:3538-48.
[11]European Commission Joint Research Centre. Photovoltaic Geographical Information System (PVGIS) – Available at: https://re.jrc.ec.europa.eu/pvg_tools/en/#api
[12]Open-Meteo. Open-Meteo: Open-Source Weather API – Available at: https://open-meteo.com/
[13]Optuna. Optuna: A hyperparameter optimization framework – Available at: https://optuna.org/
[14]Japan Meteorological Agency. Past Weather Data Download. Available at:https://www.data.jma.go.jp/gmd/risk/obsdl/ [accessed 24.3.2026].
[15]Kansai Electric Power Company. Denki Yohou (Electricity Forecast). Available at:https://www.kansai-td.co.jp/denkiyoho/area-performance/index.html [accessed 24.3.2026].
[16]Central Research Institute of Electric Power Industry (CRIEPI). Comprehensive evaluation of life-cycle CO2 emissions of power generation technologies in Japan. 2016.
[17]Rashid L.B., Shuja S.Z., Rehman S. Machine Learning Forecasting of Direct Solar Radiation: A Multi-Model Evaluation with Trigonometric Cyclical Encoding. Forecasting 2025; 7(4): 58.