IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 4 March 2024, accepted 1 April 2024, date of publication 10 April 2024, date of current version 17 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3387071

RESEARCH ARTICLE

Machine Learning-Based Japanese Spot Market Price Forecasting Considering the Solar Contribution

XUE FANG [1], JINDAN CUI [2], TAKASHI OOZEKI [3], AND YUZURU UEDA [2], (Member, IEEE)
[1]Research Institute for Science and Technology, Tokyo University of Science, Tokyo 125-8585, Japan
[2]Department of Electrical Engineering, Tokyo University of Science, Tokyo 125-8585, Japan
[3]National Institute of Advanced Industrial Science and Technology (AIST), Fukushima 963-0298, Japan

Corresponding author: Xue Fang (fangxue@rs.tus.ac.jp)

This work was supported by the New Energy and Industrial Technology Development Organization (NEDO) "Demonstration Study of Photovoltaic Power Generation Technology to Create Flexibility" under Grant 21501396-0.

ABSTRACT The increasing reliance on photovoltaic (PV) generation as a cornerstone of carbon neutrality has led to transformative changes in the energy structure, further impacting electricity market trading mechanisms and price volatility. The electric power system reform also promoted wholesale trading in the Japan Electric Power Exchange (JEPX) spot market. This study explores an effective JEPX spot market price forecasting model that enables PV power suppliers to make informed production decisions and ensure revenue optimization. We found that understanding the net demand (total demand minus PV generation) is crucial for accurate price forecasting, as it allows for a more precise reflection of the gradual evolution of the solar-dominated energy structure of the dynamic electricity demand. We also conducted parameter tests and a comparative analysis of different training loops and periods in the basic form using an artificial neural network (ANN) and support vector regression (SVR) algorithms. The results indicated that the narrow ANN and SVR models with a linear kernel function and training in the continuous loop method performed better in spot market price forecasting than other model settings. Our proposed approach can provide essential insights into future price trends, facilitating informed sustainable energy planning and resource allocation for power generation to guarantee the benefits of achieving solar promotion and net-zero transition.

INDEX TERMS Japan electric power exchange (JEPX) spot market, electricity price forecasting (EPF), net demand, PV generation, support vector regression (SVR), neutral network regression (ANN).

I. INTRODUCTION

The global transition towards renewable energy sources, driven by environmental concerns and the need for sustainable energy solutions, has reshaped the electricity generation landscape. In this context, Japan has embarked on a significant transition towards a more sustainable and diversified energy mix. Solar power has been the primary driver of renewable energy deployment, offering a clean and abundant energy source that has garnered substantial attention and investment since the late 1990s. By the end of 2022, the cumulative capacity reached 78.8 GW, and may reach up to 370 GW in the future, which is approximately 10% of the total estimated domestic primary energy demand by 2050 [1]. However, the output of renewable energy sources, such as solar and wind power, is variable and uncertain, posing integration challenges for grid stability and reliable power supplies. Next-generation renewable energy infrastructure must be built with cybersecurity to protect the power grid from potential attacks [2], [3], [4], [5]. Moreover, expensive maintenance and supervision fees, extended returns on investment periods, and barriers such as limited access to financing and inconsistent government incentives create financial challenges for potential investors and slow down introduction rates. Despite earlier initiatives, such as the introduction of feed-in tariffs (FIT) to incentivize renewable energy projects, FIT ended in 2019, making power producers reconsider the benefits of selling electricity [1]. With the approval of Japan's

The associate editor coordinating the review of this manuscript and approving it for publication was Behnam Mohammadi-Ivatloo.

Sixth Strategic Energy Plan for 2030 and the introduction of the Feed-in-Price Subsidy (FIP) plan, the discourse on electricity price forecasting has gained significant traction and relevance, making it a crucial and timely topic.

Electricity price forecasting plays an important role in planning operations, optimizing buying and selling strategies, and integrating renewable energy sources into the grid for utilities, energy traders, and other market participants [6]. For electricity suppliers, it helps well-balance their costs and benefits by providing more accurate estimates of future prices by considering factors such as market trends, technological advancements, and evolving energy policies. This, in the whole power grid industry, as a decision or preference factor, offers guidance or recommendations for more informed activities and decisions or preference factors for the entire power grid industry, which can lead to a more efficient allocation of resources when navigating the evolving energy landscape and align it with emerging market needs [7]. It has been proven that considering preferences (such as technological, financial, and even political factors) can help develop practical and effective cyber defense strategies to protect the safety of the power grid [3]. In the attacker-defender game, effective defense is not only determined by finding the critical sections [5] or simulating different attack-defense scenarios of a power grid [4] but also needs to enhance incident response capabilities during critical periods. Defenders can use price data for predicting potential attack times and weigh the cost of implementing security measures to prioritize and strategically allocate resources more effectively during periods of heightened risk. Moreover, supervisory control and data acquisition systems for solar PV have recently become some of the main targets. Real-time monitoring of day-ahead price forecast data to identify abnormal fluctuations facilitates the detection of anomalous data in renewable energy generation that may be caused by cybersecurity threats or unauthorized access, ensuring prompt responses and mitigation measures. Additionally, policymakers need to have prior knowledge of electricity price trends to evaluate or coordinate supply-demand plans, boost the introduction of renewable energy infrastructure, and design and run multiple electricity markets like capacity and balancing markets [1]. Understanding electricity price trends is essential for the long-term sustainability of solar power projects to ensure that the expansion of solar PV capacity reaches the net-zero target [8].

The Japan Electric Power Exchange (JEPX) is a wholesale electricity market in Japan established in 2005 as part of the Japanese government's efforts to deregulate the electricity market and introduce competition [9]. The purpose of the JEPX is to facilitate an efficient and stable supply of electricity in Japan by creating a transparent and competitive marketplace. It operates as an exchange, similar to a stock exchange, in which electricity and related products such as futures contracts and options are bought and sold between power generators, suppliers, and consumers. The spot market in the JEPX is a real-time market for buying and selling electricity for immediate delivery, which plays a critical role for efficient and effective operations [9], [10]. It operates in half-hour intervals, with market participants submitting their bids and offers for electricity during each trading interval. The market clearing price is then determined based on the balance between the supply and demand at any given moment [11]. Day-ahead forecasting, which is one of the short-term forecasting, can provide the forecast price for the delivery day for reference one day before the bidding day to match the day-ahead trading mechanism of the JEPX spot market. Typically, a combination of historical data such as previous day-ahead prices and real-time data are used [12], [13]. This information is then used to allocate resources such as generators and storage facilities to ensure that the expected demand can be met while maintaining the reliability and stability of the electricity grid [14], [15]. If the demand is underestimated, there may be a shortage in the supply, leading to blackouts or disruptions [16]. Conversely, overestimating the demand may result in an excess supply of electricity leading to wastage and increased costs. To accomplish this, there are two prerequisites for deriving the optimal forecasting model: (1) selecting effective predictor variables that influence market mechanisms by considering the context of Japan's energy situation and (2) choosing suitable algorithm configurations and conducting a specific logic in the training and forecasting process, as explained below.

A short-term electricity price model can be developed using only four label variables – week, month, year, and hour – as well as historical data of the same day in the previous week [17] To achieve more accurate predictions, it was confirmed that a high accuracy can be achieved through proper feature selection by investigating the relevant attributes of the data among the redundant features [18]. In Ref. [19], a sensitivity analysis was performed to select the most appropriate input features that considered the lagging electricity prices and demand loads. The forecasting results of Ref. [20] led to the same conclusion: models with system load demand as the exogenous variable generally performed better than those with only input price information. A state-space model was used to forecast the daily JEPX spot market price by adding the explanatory variables of buy and sell bid volume, daily highest and lowest temperatures, and electricity demand [21] Although many scholars have provided similar insights into the use of fundamental drivers (total demand, weather conditions, historical prices, etc.) through complex algorithms, only a few have explored the relationship between the usage of renewable energy sources and changes in the energy structure of the spot market price. The effects of wind generation and weekdays on Spanish electricity spot price forecasting were studied [22] and found to facilitate the fitting and forecasting processes. This conclusion confirms the influence of the energy situation on the electricity market. However, due to different data and complicated conditions, it is difficult to reach a consensus on countries that have solar-driven renewable energy targets. The selection of effective predictor variables must consider the prevalence of solar power, positioning Japan as one of the world's largest solar energy markets.

Several model algorithms are typically used for forecasting electricity prices. The choice of method depends on various factors such as the availability of data, time horizon of the forecast, and specific requirements of the object [23], [24]. Statistical forecasting is the simplest and most basic method based on an analysis of factors that can influence electricity prices with the goal of predicting future prices. Its limitations have always been criticized due to its inability to accurately model the typical nonlinear behavior of electricity prices and predictor variables. Time-series analysis can be performed using various statistical techniques such as the autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroscedasticity (GARCH), and exponential smoothing methods (ESMs) [12], [20], [25], [26]. However, this method relies solely on a fixed set of variables and assumptions, which limits its flexibility and adaptability to changing market conditions and data availability. Thus, it may not be possible to capture new or emerging trends that could affect electricity prices.

In recent years, as machine-learning techniques have become more sophisticated, they are increasingly being applied to forecasting problems and have a significant impact on various industries [12], [14], [24], [27], [28]. Machine learning can flexibly handle complex and non-linear tasks and has shown promising results in electricity price forecasting, particularly for short-term forecasts. Neural networks (ANN) and support vector regression (SVR) are among the most popular machine learning algorithms and are good at modeling data features used for forecasting. An ANN-based model combined with the weighted least squares (WLS) technique was proposed for forecasting locational marginal prices [29]. It enhanced the prediction accuracy primarily for spring and winter, but evidence for the model's performance across all seasons is lacking, particularly during high electricity demands in summer. A hybrid (Levenberg–Marquardt) model for day-ahead price forecasting in the electricity markets of the Indian and Austrian energy markets is presented in Ref. [30]. This highlights the effectiveness of the model in specific market environments for the current year and raises doubts regarding its accuracy in other scenarios [19]. In Ref. [31], a multiblock ANN was optimized using an intelligent algorithm to enhance both the training mechanism and price and load prediction abilities. However, these studies focused exclusively on optimizing the model through method combination development or parameter optimization. Some of these, with a more complex approach, hinder clear understanding and consensus. Additionally, the study in Ref. [32] was conducted using both SVR and ANN models with one unique solution and found that SVR requires less time, and its accuracy is almost the same as that of ANN. This indicates that machine learning requires large amounts of high-quality data and the careful selection and tuning of algorithms to achieve accurate results. Under such circumstances, the performance of the models depends on the representativeness of the reference dataset used for training [33], [34]. It should accurately reflect the complexities and fluctuations present in the target system, ensuring that the model generalizes well to unseen data and provides reliable predictions under different conditions. Existing research lacks sufficient attention to the crucial aspects of regular updates and adjustments to the training period of a dataset, which are essential for ensuring its continued representativeness as market dynamics evolve.

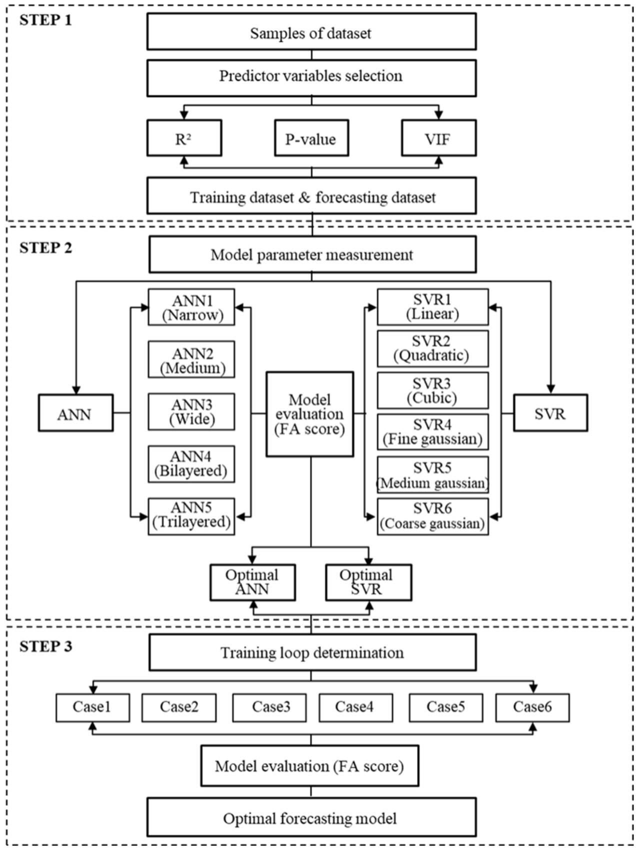


FIGURE 1. The process of empirical research.

This study presents a case study of the Tokyo area that explores a better option for forecasting JEPX spot market prices (Fig. 1). In this study, we recognize the increasing significance of solar power in Japan's electric power demand portfolio and the ongoing renewable energy revolution. When developing a forecasting model, in addition to commonly used input variables, we considered the impact of the energy supply structure (such as the influence of photovoltaic power generation) and the physical relationships that exist in electricity trading. This consideration adds a crucial layer to the price-forecasting model, which goes beyond the standard ones, and showcases a timely approach tailored to the specific challenges posed by renewable energy sources. Different settings of the model parameters and training loop were tested using efficient ANN and SVR algorithms to improve the prediction accuracy and make it easier to understand.

This series of comparative and empirical validations not only substantiates the model's effectiveness, but also contributes to a broader understanding of the applicability and performance of machine learning techniques in electricity price forecasting within the context of renewable energy. Our research aligns with the energy policy development goals and industry requirements of countries such as Japan, for which solar energy plays a pivotal role in shaping the future energy landscape. The findings have practical implications for stakeholders, policymakers, and other relevant entities of renewable energy in the power industry by informing decision-making with regards to market behavior or energy allocation. The structure of this paper is as follows: Section II presents information on the object, dataset, tool, ANN, and SVR methods, and the established forecasting process using machine learning. In Section III, we discuss a series of analyses, forecasting, and evaluation results, including the predictor variables used, considering the impact of solar contribution, model parameter measurement results, training loop determination, and forecast evaluation results. Finally, the conclusions are presented in Section IV.

II. MATERIALS AND METHODS

A. OBJECT, DATASET, AND TOOL

In this study, the spot market price in the Tokyo area is considered, which is controlled by the Tokyo Electric Power Company (TEPCO). TEPCO is one of the largest electric utilities in Japan, and participates in the JEPX spot market as a generator and supplier of electricity [35], [36]. It is subject to the same market forces as the other participants in the spot market, which means that TEPCO's participation in the JEPX spot market is an essential part of its strategy for balancing its supply and demand requirements and ensuring efficient production and delivery of electricity to its customers. Trading in the spot market consists of 48 timeframe items every 30 minutes, similar to other areas in the JEPX.

The dataset of all variables in the forecasting process comprises 30-minute average values, including the electricity spot price published on TEPCO [37] and other datasets of selected variables collected from the public authority websites of the Tokyo Area. It is divided into a training dataset for the fiscal year 2020 (FY2020:2020.4.1∼2021.3.31) for model learning and a forecasting dataset for the fiscal year 2021 (FY2021:2021.4.1∼2022.3.31) for model evaluation. To verify the validity of the model, we used actual data for testing. If this works, the predicted values will be used in practical projects in the future.

The entire forecasting process was performed in MATLAB (R2021b) platform, a widely used programming language and numerical computing software. It provides an interactive environment for algorithm development, data visualization, and data analysis, making it a powerful tool for researchers and engineers in various fields [38]. "fitrnet" and "fitrsvm" are the main function codes in the model used to perform the entire forecasting process based on the ANN regression algorithm and the support vector regression (SVR) [39].

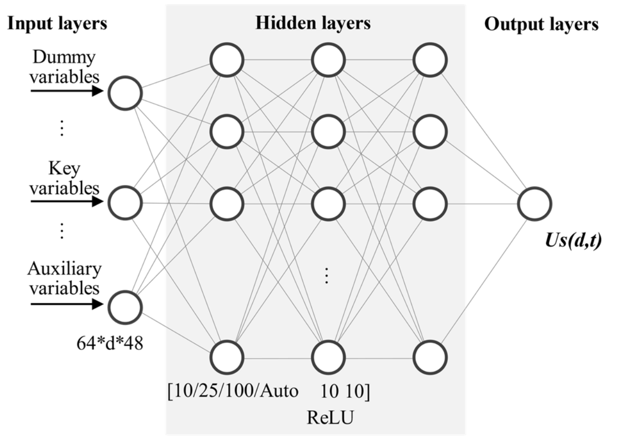


FIGURE 2. The structure of ANN models.

B. ARTIFICIAL NEUTRAL NETWORKS FOR REGRESSION (ANN)

The layer size of an ANN model (which refers to the number of neurons) is the most important hyperparameter that can significantly affect network performance in regression tasks. This determines the capacity of the neural network to model complex relationships between the input and output variables [40]. It is often necessary to experiment with different layer sizes to determine the appropriate hyperparameters [12].

In this study, five types of ANN models were used for a comparative analysis (Fig. 2). Narrow, medium, and wide neural networks refer to the number of neurons in the hidden layers of a network, whereas bilayer and trilayer neural networks refer to the number of hidden layers in a network [41]. The advantages of each type of neural network and the layer size depend on the specific problem, data, and architecture of the network [42]. A rectified linear unit (ReLU) [43] function was used and performed as follows, where x is the input for the neuron:

$$f(x) = x^{+} = \max(0,x) = \begin{cases} x & \text{if } x > 0, \\ 0 & \text{otherwise}. \end{cases} \tag{1}$$

$$f'(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{2}$$

C. SUPPORT VECTOR REGRESSION (SVR)

SVR requires careful parameter tuning, especially the choice of kernel function and scale, to ensure that the model does not overfit the training data. Kernel function selection is the key point because SVR uses it to transform input features into a higher-dimensional space, allowing for better separation of data points. The choice of the kernel function significantly affects the model's ability to capture complex relationships in the data. Selecting an inappropriate kernel may lead to a poor generalization of unseen data.

FIGURE 3. Training and forecasting process using machine learning.

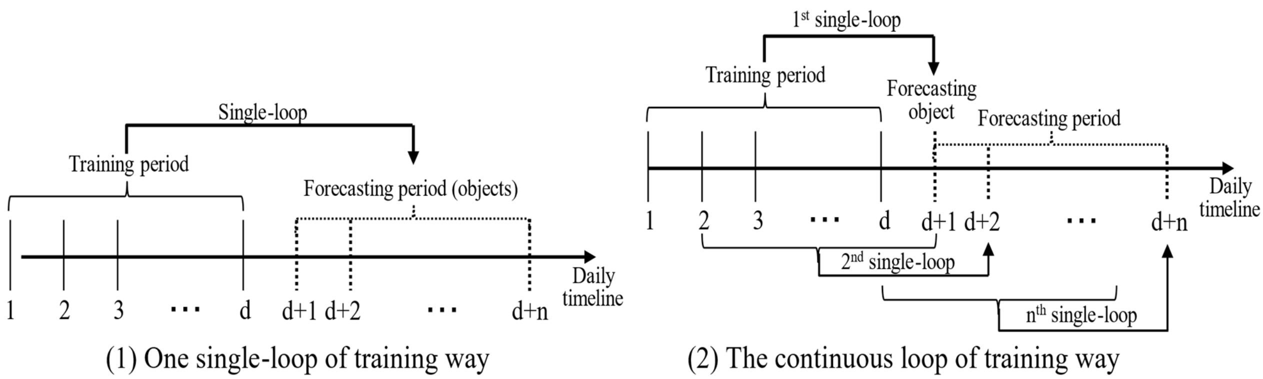


FIGURE 4. Training forecasting models in two loop ways.

In this study, six types of SVR models were created using different kernel functions with other automatic hyperparameter settings. Linear, quadratic, and cubic SVR models have the advantages of simplicity, ease of understanding, computational efficiency, and easy interpretation of the relationship between the input features and target variable [44]. Gaussian kernel function SVR Models, namely fine Gaussian, medium Gaussian, and coarse Gaussian models, are capable of modeling complex nonlinear relationships [45]. They have a high degree of generalization and are robust to noise and outliers in the data, making them suitable for real-world applications. Model flexibility using the Gaussian function decreased with the kernel-scale setting. The kernel function is given by:

Linear kernel:

$$K(x_i, x_j) = x_i x_j^T \tag{3}$$

Quadratic and cubic kernel:

$$K(x_i, x_j) = (x_i x_j^T + c)^d \tag{4}$$

Gaussian kernel:

$$K(x_i, x_j) = exp\left(-\left\|x_i - x_j\right\|^2 / 2\sigma^2\right) \tag{5}$$

where x_i and x_j are input vectors, and c is a constant. $\left\|x_i - x_j\right\|$ denotes the Euclidean distance between the two vectors, and σ is a bandwidth parameter that controls the kernel scale for different types of Gaussian models.

D. TRAINING AND FORECASTING PROCESS USING MACHINE LEARNING

Our short-term forecasting is a one-day-ahead process; that is, we must predict the spot market price on the delivery day one day before the bidding date to maximize benefits. The basic process was the same regardless of the machine-learning technique (Fig. 3). Before training, we standardized the inputs and outputs so that they always fell within a specified range using the min-max scaling method [-1 1] [46]. However, there was a difference between the two types of models in the internal training paradigm. While both ANN and SVR involve iterative processes for model training, ANN focuses on adjusting the weights through multiple training epochs to minimize forecasting errors and ensure the capture of complex relationships in the data. By contrast, SVR emphasizes finding the hyperplane that best fits the data in a transformed space defined by the chosen kernel function.

We then implemented the model with two training logics, which helped optimize the forecasting model and ensure its accuracy (Fig. 4):

(1) One single-loop of training is the common process of inputting training observations for batch learning to produce the network's output once [47], and the export model is applied for the forecasting period.

(2) The continuous loop of training is the process of maintaining a certain fixed training period, ranging from automatically retraining the model each day by adding the 48 observations of the next day to using the newly derived model to complete a one-day prediction loop until the end of the prediction period.

Subsequently, the forecasting model was extracted by identifying the relevant features from the historical data. We input the normalized values of the forecast data into the model. The output data were denormalized to obtain the final predicted value. In addition, the commonly used methods of *MAE* and *RMSE* [48], [49] are chosen to determine the extent of

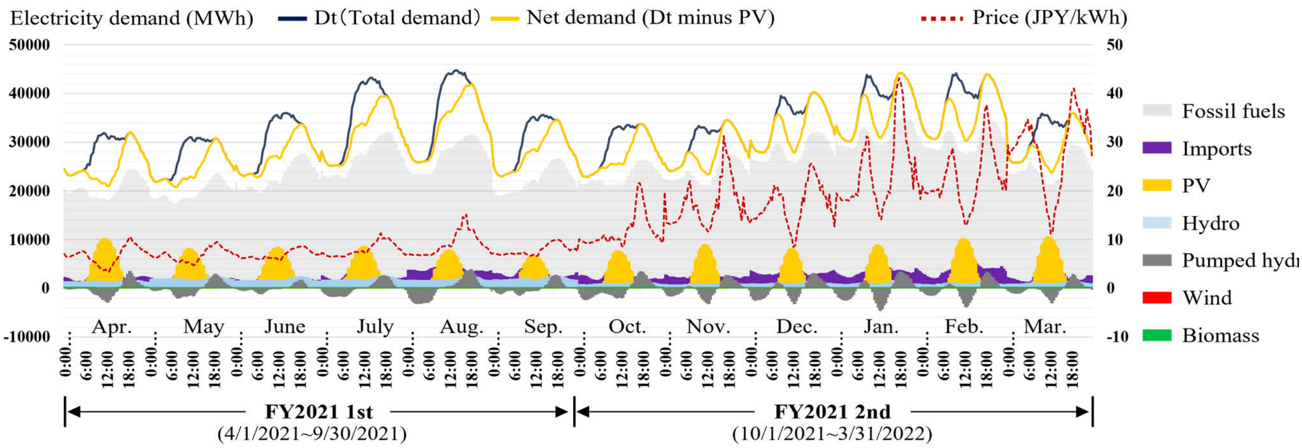


FIGURE 5. Average timeframe electricity demand profile and system price of Spot Market, Tokyo Area.

spread between the forecast and actual values. Meanwhile, we provided a combined *FA score* for performance ranking, calculated as follows:

$$MAE = \frac{1}{N}\sum\nolimits_{i=1}^{N} |\widehat{Us_i} - Us_i| \tag{6}$$

$$RMSE = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N} (\widehat{Us_i} - Us_i)^2} \tag{7}$$

$$FA\ score = 2 - \frac{MAE - \min(MAE)}{\max(MAE) - \min(MAE)} - \frac{RMSE - \min(RMSE)}{\max(RMSE) - \min(RMSE)} \tag{8}$$

where, $\widehat{Us_i}$ is the forecast, Us_i is the actual spot market price for the timeframe i. N is the total number of timeframes over the forecasting period.

III. RESULTS AND DISCUSSIONS

A. PREDICTOR VARIABLES SELECTION CONSIDERING THE SOLAR CONTRIBUTION

In the Tokyo area, solar PV power generation is expected to account for 7.6% of the total demand by 2021, which is much higher than the share of other renewable electricity sources (Fig. 5). PV generation is speculated to be a key factor that may play an important role in dominating spot market prices in the JEPX, particularly as solar energy use continues to grow. This is because solar PV power has very low marginal costs, meaning that the electricity generated by solar panels is essentially free from other sources such as fossil fuels. Conversely, when there is a shortage in electricity from PV power generation, other sources may need to increase their output to meet demand, which may lead to an increase in spot market prices. Moreover, the spot market price is determined by the supply of and demand for electricity; thus, the total power demand is another key factor that can directly influence the price at any given time.

To verify the inference that the dominance of solar PV makes it a pivotal and influential factor in shaping electricity prices, we designed a series of comparative verifications of

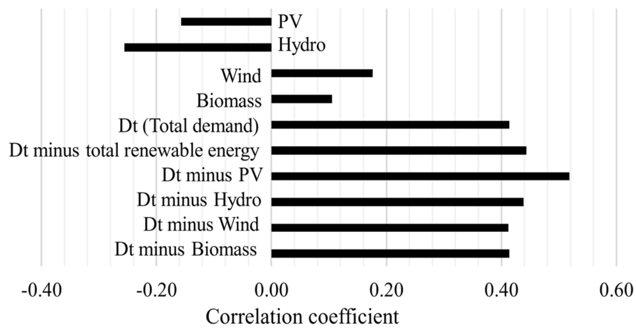


FIGURE 6. The correlation analysis between variables of renewable energy generation and spot market price in FY2021.

the correlation analysis, as shown in Fig. 6. This involved assessing the correlation between the electricity generation from each renewable energy source (including PV, hydro, wind, biomass, whereby geothermal and nuclear power were excluded due to minimal or negligible generation) and spot market price, as well as examining the correlation between the net demand (total demand minus each renewable source's generation) and price. The results showed that except for solar PV and hydropower, all other variables exhibited a positive correlation with the prices. The relationship between the net demand and the price surpasses that of the individual energy generation correlations. Remarkably, the positive relationship between the net demand (total demand minus PV generation) and the price was the most substantial among all combinations.

Moreover, we conducted a thorough examination of the interactions between the optimal variable and price over time (Fig. 7). Our investigation revealed that during periods of abundant solar energy production (from 5:30 to 17:30 in FY2021), the impact of the net demand on the price experienced a notable increase, playing a significant role in enhancing spot market stability. However, this correlation weakens during periods of low solar power generation, resembling the behavior of the total demand. This can be

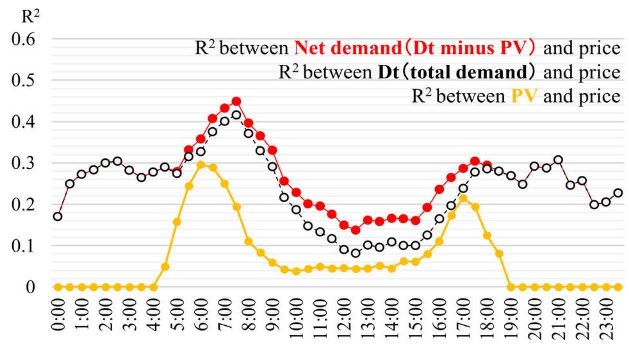


FIGURE 7. The coefficient of determination (R^2) between Net demand / Dt (total demand) / PV and spot market price in FY2021.

attributed to the reliance on more expensive power generation resources, such as fossil fuels, coupled with heightened nighttime demand, which leads to frequent price surges. Consequently, the dual dimensions of solar contribution – its overwhelming share in the renewable energy portfolio and its dynamically driven net demand were strongly correlated with electricity prices, affirming our decision to designate it as the key predictor variable in the forecasting model.

In addition, prices can vary depending on various factors, including weather conditions, historical prices, and other seasonal and temporal characteristics. We used a series of potentially related factor data through correlation analysis and collinearity detection to screen out the composite temperature in the Tokyo Area (the composite value provided by seven meteorological observatories in the Tokyo Area) and four historical prices variables with a high coefficient of determination (R^2) of over 0.7. The external labels for the seasonal and temporal characteristics were converted into dummy variables for the forecasting model. The monthly label variable was used only when the training time exceeded six months. All variables matched the standard with an R^2 value greater than 0.3 and passed significance (p-value $<$ 0.05) and VIF testing (VIF $<$ 10), categorizing them into dummy, key and auxiliary variables (Table 1).

B. MODEL PARAMETER MEASUREMENT RESULTS

The performance of any model is highly dependent on the specific dataset and problem at hand, and it is important to evaluate multiple models to determine which performs best in spot market price forecasting. To test the different model parameters, we used one year of data (FY2020) for training and the next year's data (FY2021) as the forecasting set in a single-loop training method, using both SVR and ANN models. The predictive accuracy of the models was ranked based on a comprehensive evaluation of the *FA score*.

The comparative results indicated that (1) In ANN series models, the narrow ANN1 model with one hidden layer of ten neurons was the best choice for prediction, followed by the medium- and wide-type models (Table 2). The range of the forecasting error, calculated by subtracting the actual value from the forecasting value, was almost the same for all

TABLE 1. The variable list of ANN and SVR model.

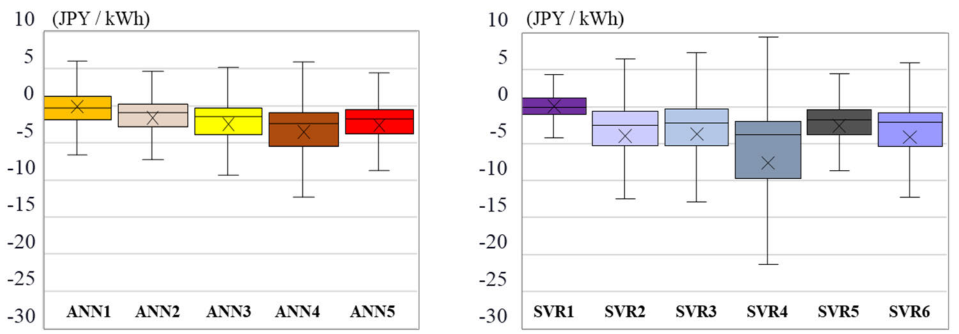
Variables			Description	Unit	Source
X	Dummy variables	***t(d,t)***	Timeframe (labeled 1 to 48)	-	-
		d(d,t)	Days of the week (labeled 1 to 7)	-	-
		m(d,t)	Months of the year (labeled 1 to 12)	-	-
	Key variables	***Dn (d,t)***	Net demand of electricity (total demand minus PV generation)	(MWh)	[37]
	Auxiliary variables	***T(d,t)***	The composite temperature of Tokyo Area	(°C)	[51]
		Us(d-1,t)	Spot market price during the same timeframe for one day before	(JPY/kWh)	[52]
		Us(d-2,t)	Spot market price during the same timeframe for two days before	(JPY/kWh)	[52]
		Us(d-3,t)	Spot market price during the same timeframe for three days before	(JPY/kWh)	[52]
		Us(AveW,t)	Average spot market price during the same timeframe for the previous week	(JPY/kWh)	[52]
Y		***Us(d,t)***	Spot market price during the same timeframe for the delivery date	(JPY/kWh)	[52]

five models, with the average range between the first quartile and the third quartile being approximately 4.00 JPY/kWh. However, the bilayered ANN4 model shows a larger error range and its error range, approximately 1.5 times of the other models. The error of the optimal ANN1 model ranges within 0.26 JPY/kWh and -2.74 JPY/kWh, respectively (Fig.8-1). (2) From among the SVR series models, the SVR1 model with a linear kernel function performed the best, followed by the medium and coarse Gaussian kernel function models (Table 3). Compared with the ANN models, the range of the forecasting error was significantly different for the six SVR models (Fig. 8-2) The linear SVR1 model had the smallest margin of error and the upper limit for the overestimation and underestimation of most timeframes is within 1.2 JPY/kWh and -0.95 JPY/kWh excluding outliers, respectively. The quadratic SVR2 model had the lowest margin of error, with the range from the first quartile to the third quartile being nearly 3.00 JPY/kWh, followed by the linear,

TABLE 2. The model parameter measurement results of ANN models.

Model		ANN1	ANN2	ANN3	ANN4	ANN5
Parameter	Types	**Narrow**	**Medium**	**Wide**	**Bilayered**	**Trilayered**
	Layer sizes [a]	[10]	[25]	[100]	[10 10]	[10 10 10]
	Activations			ReLU [b]		
FY2021 (4/21-3/22)	*RMSE*	8.24	8.48	9.60	9.59	11.01
	MAE	4.06	4.17	5.08	5.68	6.69
	FA score	2.00	1.87	1.12	0.89	0.00
	Ranking	1	2	3	4	5
FY2021_1st (4/21-9/21)	*RMSE*	2.22	2.56	2.33	2.49	2.81
	MAE	1.35	1.28	1.41	1.89	2.00
	FA score	1.90	1.42	1.63	0.70	0.00
	Ranking	1	3	2	4	5
FY2021_2nd (10/21-3/22)	*RMSE*	11.45	11.74	13.39	13.36	15.33
	MAE	6.78	7.07	8.78	9.49	11.40
	FA score	2.00	1.86	1.07	0.92	0.00
	Ranking	1	2	3	4	5

[a] Rectified linear unit (ReLU) function performs a threshold operation on each element of the input, where any value less than zero is set to zero.



(1) The range of forecasting error using ANN models. (2) The range of forecasting error using SVR models.

FIGURE 8. The range of forecasting error (forecasting value minus actual value).

trilayered, cubic, and bilayered models. The performance of the fine Gaussian SVR4 model in price forecasting was the worst, with an error of up to 9.65 JPY/kWh for underestimation, which was ten times higher than that of the best SVR1 model.

Both the ANN1 and SVR 1 models were favored because of their effectiveness in capturing limited complexity data characteristics. Data with the key variable net demand have a relatively regular curve change pattern dominated by PV generation, such as a sine wave-like shape, morning and afternoon ramps, and seasonal variations. The ANN1 model, characterized by a single hidden layer and a restrained number of neurons, was strategically chosen to mitigate the risk of overfitting. This architectural choice prevented the model from becoming overly complex, thereby reducing the number of parameters that must be learned. Deeper networks with more layers often face challenges, such as vanishing or exploding gradients, which hamper effective training. Furthermore, the ANN1 model efficiently filters instantaneous fluctuations in seconds of solar data, enhancing the computational efficiency by minimizing the number of weights and biases that require training. This streamlined approach not only optimizes efficiency, but also bolsters the model's ability to generalize effectively to diverse datasets. In the case of the SVR algorithm, the selection of the linear SVR model was motivated by its resistance to overfitting and superior capability to capture the underlying relationships within the data, outperforming other more complex models such as non-linear and Gaussian kernel functions in this task. This makes it easier to interpret the model and understand the relationship between the input variables and spot market prices.

In terms of the ANN algorithm (Table 2), the narrow ANN1 model with one hidden layer of ten neurons was the best choice for prediction, followed by the medium- and wild-type models. A narrower network can reduce the risk of overfitting by limiting the number of parameters that must be learned. It is computationally efficient with fewer weights and biases that must be trained. However, deeper networks with more layers are prone to vanishing or exploding gradients, making it difficult to train the network effectively. Compared with the SVR model, the ANN model had a smaller forecasting error, and the error ranges of all five ANN models were similar, with a larger margin of underestimation (Fig.8-1). The quadratic SVR2 model had the lowest margin of error, with the range from the first quartile to the third quartile

TABLE 3. The model parameter measurement results of SVR models.

Model		SVR1	SVR2	SVR3	SVR4	SVR5	SVR6
Parameter	Kernel function	**Linear**	**Quadratic**	**Cubic**	**Fine gaussian**	**Medium gaussian**	**Coarse gaussian**
	Kernel Scale [a]		Automatic		0.75	3.00	12.00
FY2021 (4/21-3/22)	*RMSE*	8.46	9.45	11.60	12.83	8.15	9.48
	MAE	3.53	5.40	6.53	7.73	4.32	5.18
	FA score	1.93	1.28	0.55	0.00	1.81	1.32
	Ranking	1	4	5	6	2	3
FY2021_1st (4/21-9/21)	*RMSE*	2.13	2.60	2.82	2.94	2.20	2.32
	MAE	1.04	1.90	2.00	2.20	1.52	1.58
	FA score	2.00	0.68	0.32	0.00	1.49	1.30
	Ranking	1	4	5	6	2	3
FY2021_2nd (10/21-3/22)	*RMSE*	11.79	13.12	16.18	17.93	11.32	13.21
	MAE	6.03	8.91	11.08	13.28	7.13	8.81
	FA score	1.93	1.33	0.57	0.00	1.85	1.33
	Ranking	1	3	5	6	2	4

[a] Fine gaussian model allows rapid variations in the response function with a kernel scale set to sqrt(P)/4. Medium gaussian model gives a less flexible response function with a kernel scale set to sqrt(P). Coarse gaussian model gives a rigid response function with a kernel scale set to sqrt(P)*4, where P means the number of predictor variables.

being nearly 3.00 JPY/kWh, followed by the linear, trilayered, cubic, and bilayered models. The error of the optimal ANN1 model ranges from 0.26 JPY/kWh to -2.74 JPY/kWh, respectively.

In terms of the SVR algorithm (Table 3), the empirical results show that the SVR1 model with a linear kernel function performed best in both 1st and 2nd half of FY2021, followed by the medium- and coarse-Gaussian kernel function models. The linear SVR is less prone to overfitting and can better capture the underlying relationships, outperforming other more complex models such as non-linear and Gaussian kernel functions in this task. This makes it easier to interpret the model and understand the relationship between the input variables and spot market prices. Moreover, the range of the forecasting error, which was calculated by subtracting the actual value from the forecasting value, differed significantly among the six SVR models (Fig.8-2). All the models tended to underestimate the errors, which were relatively large. The linear model has the smallest margin of error and the upper limit for the overestimation and underestimation of most timeframes is within 1.2 JPY/kWh and -0.95 JPY/kWh excluding outliers, respectively. However, the performance of the fine Gaussian SVR4 model in price forecasting was the worst, with an error of up to 9.65 JPY/kWh for underestimation, which is 10 times higher than that of the best SVR1 model.

C. TRAINING LOOP DETERMINATION

Based on these results, we identified the best-performing model parameter settings, which were the narrow ANN1 and linear SVR1 models for each algorithm. In this phase, we used these two models as the base settings and tested the most accurate option for JEPX spot market price forecasting by setting two training loop methods (a single loop and a continuous loop of training) and different training periods (one, half, quarter, one month, and one week).

We found that the process in the continuous training loop can lead to significantly better forecast accuracy than a single-loop training process, regardless of the machine learning algorithm. With continuous training, real- or near-real-time data can be incorporated seamlessly into a model, allowing it to leverage the most recent forecasting information. This adaptability is crucial in markets or scenarios for which the latest data significantly influence predictions. The continuous-loop approach allows the model to adapt to dynamic patterns in the underlying data over time, so the input variable data can match the changing trends in both the energy structure and electricity supply-demand relationship well. By regularly retraining the model with the addition of new observations, the model captures evolving patterns more effectively. Moreover, in environments in which data exhibit nonstationary behavior, especially for net demand data with seasonal characteristics, continuous retraining can help learn and adjust the model to evolving patterns, thereby improving the accuracy. In contrast, the single-loop method relies on a fixed training period, which may become less representative over time. It could potentially miss emerging patterns or shifts in the data that occur after the initial training, particularly during the accelerating transition towards cleaner energy sources.

Regarding the detailed forecast results (1) from the ANN1 model (Table 4), case 3 was the most suitable for a full-year prediction with a half-year training period. In monthly evaluations, a higher prediction accuracy was achieved for the first half of FY2021 with a training period of a quarter in case 4 or shorter. However, a longer training period of six months or more was required for the second half of FY2021. This is due to a significant increase in prices and large fluctuations in the spot market prices under more complex market conditions. Consequently, neural networks require considerable computational resources and must spend time identifying patterns and optimizing models. For the SVR1 model (Table 5),

TABLE 4. **The forecast evaluation results of six cases by using the narrow ANN1 model.**

Advanced cases		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Parameter	Training period	one-year **(365 days)**	one-year **(365 days)**	half-year **(24*7 days)**	a quarter **(12*7 days)**	one-month **(4*7 days)**	one-week **(7 days)**
	For loop [a]	No			Yes		
	Operation time	5s	1700s	821s	441s	126s	95s
FY2021 (4/21-3/22)	*RMSE*	8.24	6.66	6.18	7.76	7.69	8.22
	MAE	4.06	3.34	2.94	3.43	3.38	3.74
	FA score	0.00	1.41	2.00	0.80	0.87	0.29
	Ranking	6	2	1	4	3	5
FY2021_1st (4/21-9/21)	*RMSE*	2.22	2.06	2.18	1.69	1.94	1.99
	MAE	1.35	1.14	1.18	0.90	0.94	1.02
	FA score	0.00	0.78	0.46	2.00	1.46	1.16
	Ranking	6	4	5	1	2	3
FY2021_2nd (10/21-3/22)	*RMSE*	11.45	9.20	8.47	10.85	10.72	11.47
	MAE	6.78	5.56	4.72	5.96	5.84	6.46
	FA score	0.01	1.35	2.00	0.60	0.71	0.15
	Ranking	6	3	1	5	4	2

[a] For loop represents the different model building ways, that is, "No" means one single-loop of training way, and "Yes" means the continuous loop of training way.

TABLE 5. **The forecast evaluation results of six cases by using the linear SVR1 model.**

Advanced cases		Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Parameter	Training period	one-year **(365 days)**	one-year **(365 days)**	half-year **(24*7 days)**	a quarter **(12*7 days)**	one-month **(4*7 days)**	one-week **(7 days)**
	For loop [a]	No			Yes		
	Operation time	58s	17635s	6331s	924s	129s	19s
FY2021 (4/21-3/22)	*RMSE*	8.46	7.87	6.82	7.14	6.57	7.46
	MAE	3.53	3.48	2.84	2.98	2.81	3.48
	FA score	0.00	0.38	1.83	1.46	2.00	0.60
	Ranking	6	5	2	3	1	4
FY2021_1st (4/21-9/21)	*RMSE*	2.13	2.13	1.91	1.78	1.76	1.87
	MAE	1.04	1.03	0.95	0.85	0.85	0.95
	FA score	0.00	0.04	0.91	1.67	2.00	1.04
	Ranking	6	5	4	2	1	3
FY2021_2nd (10/21-3/22)	*RMSE*	11.79	10.94	9.47	9.95	9.14	10.39
	MAE	6.03	5.95	4.74	5.12	4.83	6.03
	FA score	0.00	0.38	1.88	1.40	1.93	0.53
	Ranking	6	5	2	3	1	4

[a] For loop represents the different model building ways, that is, "No" means one single-loop of training way, and "Yes" means the continuous loop of training way.

a one-month training period (case 5) is optimal for spot market price forecasting in both the first and second halves of the year. In contrast to the computational logic of neural networks, SVRs are well-suited for handling limited training data because of their ability to accurately identify complex relationships using a separating hyperplane, regularization to prevent overfitting, kernel functions, and convex optimization to find the global optimum solution.

D. FORECASTING EVALUATION AND DISCUSSIONS

The optimized forecasting solutions were then selected for an accuracy evaluation. Specifically, the ANN1 model with a continuous-loop training period of one quarter (case 4) was employed to forecast the first half of the period (case 3), which was characterized by relatively stable price fluctuations. Subsequently, the same model, with a continuous training period of six months, was used to forecast the second half of the period, marked by high and frequently fluctuating prices. The results (Figs. 9 and 10) indicated that: (1) the forecast accuracy is better for the first half of the period and higher than that for the second half. This is because the electricity market price in that period was affected by a variety of complicated and unexpected factors leading to difficulties in forecasting, such as an imbalance in the crude oil supply and demand caused by the pandemic, sudden changes in market regulations, and climate change. (2) The R^2 of the correlation coefficient exceeds 0.5, compared with the predicted and measured values displayed in the scatter plots. However, the approximate curves deviate from the diagonal line because of the difficulty in predicting extreme prices, such as drastically high values of over 40 yen/kWh or much lower values close to 0 yen/kWh. In addition, we selected typical weeks for each of the four seasons and found that (3) the model perfectly matched the variance trends of the

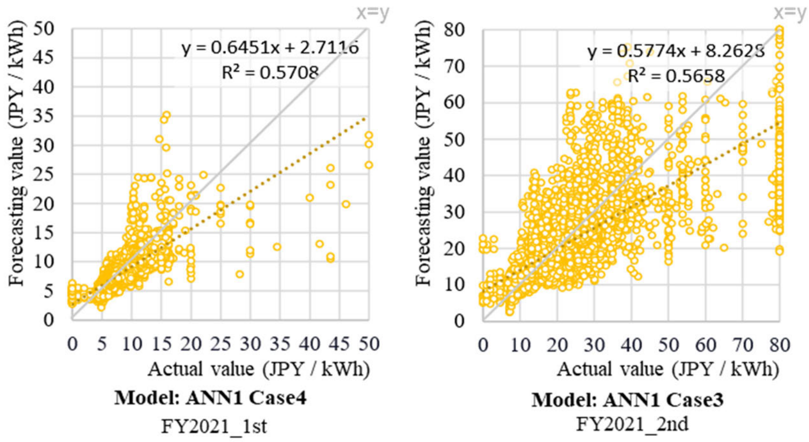


FIGURE 9. The scatter plot between actual and forecasting value using the optimal ANN1 (case 3 and case4) model.

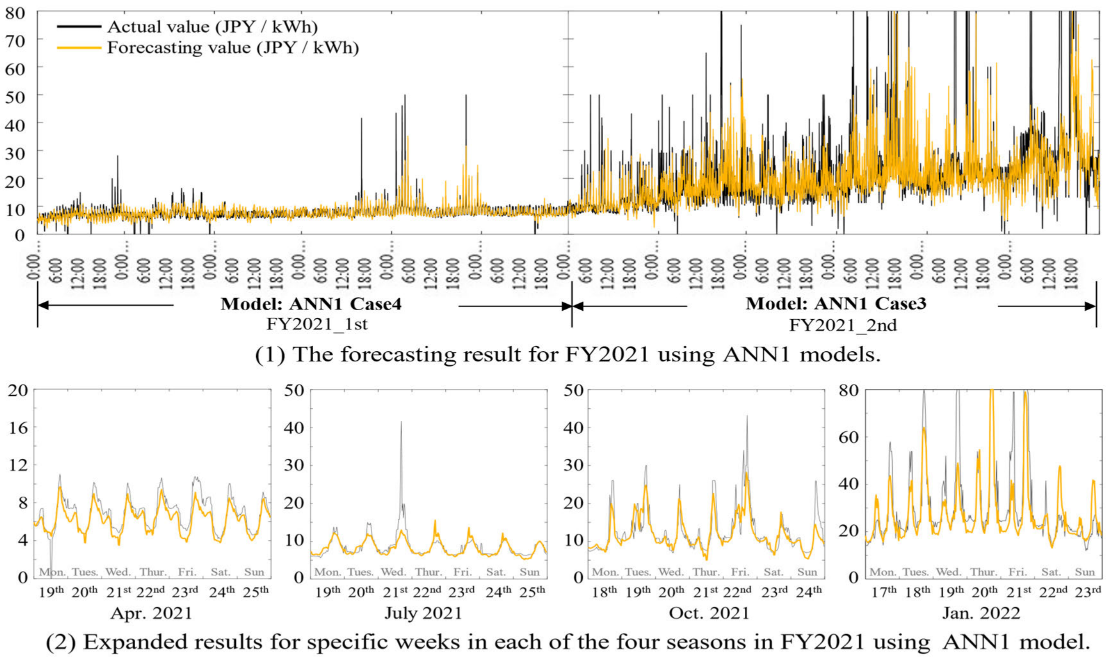


FIGURE 10. The forecasting results using the optimal ANN1 (case 3 and case4) model.

spot market prices, especially during the daytime from 8:00 to 17:00, because the changes in net demand were mainly driven by the increased utilization of solar power.

Despite employing two distinct machine learning architectures, we observed remarkably similar performances in the spot market price forecasting for both architectures. The error ranges for the more accurate predictions generated by the ANN1 and SVR1 models were not significantly different (Fig. 11). For most timeframes in the first half of the year, the upper limits for overestimation and underestimation were within 0.6 JPY/kWh, which was considerably better than the margin observed for the second half of the year, which was almost within 3 JPY/kWh. The distinction between forecasting accuracy under stable and fluctuating market conditions highlights these limitations. The differences in accuracy can be attributed to the dynamics of the electricity supply-demand relationship during fluctuations, whereby sudden and unpredictable events introduce additional challenges for predictive models. The training data utilized for model development may inherently harbor biases or limitations that contributes to both models exhibiting sub-optimal performances when confronted with extreme price scenarios. The ANN1 model lacks the complexity required to capture and represent the intricacies of the data, especially when dealing with scenarios involving fluctuating market conditions. The narrow ANN1 model may be sensitive to

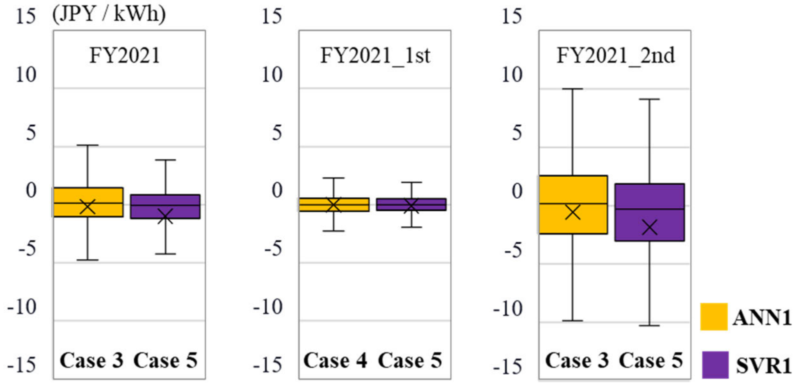


FIGURE 11. The range of forecasting error using ANN1 and SVR1 models.

outliers or extreme values, causing less weight to be assigned to these instances during training. Similarly, the linear SVR model is limited in accommodating non-linear relationships, particularly in extreme scenarios. Consequently, the model may struggle to adapt to these outliers, resulting in prediction difficulties.

Although our current models prove applicable in most timeframes, their performance may be constrained in highly volatile conditions, which needs to be further explored using the following potential methods or approaches. Enhancements can be made from the model mechanism perspective by (1) incorporating data transformation or specialized treatment to identify and handle outliers more appropriately, and (2) implementing ensemble methods (such as bagging or boosting) that combine the strengths of different models to enhance the robustness and better capture extreme price movements. Another avenue for improvement is factor exploration by (1) exploring external data sources in a broader context, including economic indicators, geopolitical events, and major policy changes. Additionally, (2) collecting more granular or real-time data allows for dynamic adjustments to the unfolding market situation.

IV. CONCLUSION AND FUTURE WORK

This study advances the field by addressing the specific challenges regarding production, investment, and energy allocation for solar power suppliers and government authorities through the application of machine learning, consideration of the key demand and energy supply factors, and comparative validation, thereby contributing useful insights to the broader understanding of electricity price forecasting in the context of a solar-dominated renewable energy structure. We established a systematic predictive analysis and evaluation workflow. Through this process, an optimal solution for forecasting the spot market price of the wholesale electricity exchange market in Japan is proposed, including the selection of predictor variables, measurement of model parameters, and determination of the training period. Our results reveal that the model can accurately reflect price change trends, especially during steady periods, and demonstrates a high level of forecast accuracy. The main conclusions from the JEPX spot market price forecast for Tokyo can be summarized as follows:

(1) For the selection of predictive model variables, the net demand (the total demand minus PV generation) is a key one besides time dummy variables, temperature, and auxiliary variables of historical prices. The widespread introduction of solar energy is anticipated to stabilize spot market prices. It should be noted that a reduction in net demand owing to the increasing introduction of solar power has a direct downward effect on spot market prices. Therefore, the bidding strategy should carefully consider the multiparty electricity market to maximize benefits, with particular attention paid to the trading mechanism of the supply-demand adjustment market.

(2) For the model parameter decision, the narrow ANN1 model with one hidden layer of 10 neurons and the SVR1 model with a linear kernel function performed better in spot market price forecasting than the other models. The forecasting accuracies of the two models are very similar in their capacity to efficiently handle the limited complexity of key variable data, with the ANN1 model designed to combat overfitting challenges, and the linear SVR model excelling at capturing and interpreting the underlying relationships within the dataset. Both methods can achieve a high predictive accuracy in periods with stable price variations, thereby ensuring a robust performance in most market stability scenarios. However, it is relatively difficult to forecast periods of sudden price changes or frequent fluctuations owing to complex or unexpected factors.

(3) With regards to the training logic of forecasting, training using the continuous-loop method significantly improved the forecast accuracy compared with the commonly used single-loop training process. This is because regular data updates can better match changes in the energy structure caused by the continuous introduction of solar energy. Owing to the different calculation mechanisms of the models, ANN1 requires a training period of more than six months for optimization, whereas SVR1 only needs to find support vectors on the decision boundary, making it possible to reach the

same level of predictive accuracy with only one month of training data.

However, the performance of the current model is unsatisfactory because of the difficulty in predicting the occurrence of extreme prices in unexpected situations. This is expected to improve by addressing model constraints through data transformation and ensemble methods, leveraging external data sources for a broader context and collecting more granular or real-time data for dynamic adjustments to unfolding market situations. In the real application environment of our forecasting model, given the increasing reliance on digital technologies and data in energy markets, it is critical to address potential cybersecurity risks and mitigation strategies. We also need to establish a real-time monitoring system to maintain continuous availability for decision-making in the electricity market. Adversarial training techniques were imported during the model development phase to enhance the resilience of the machine-learning models against manipulative attacks. Additionally, reinforcement approaches such as Q-learning are incorporated to develop specific asset optimization and attack resilience strategies. By combining advancements in modeling techniques with cybersecurity considerations, our future work will focus on comprehensively improving the accuracy, reliability, and security of the current approach to provide a more holistic view of the challenges in modern electricity markets.

REFERENCES

[1] *The Electric Power Industry in Japan 2022*, Japan Electr. Power Inf. Center, Inc. (JEPIC), Tokyo, Japan, 2022.

[2] A. Walker, J. Desai, D. Saleem, and T. Gunda, "Cybersecurity in photovoltaic plant operations," Dept. National Renewable Energy Lab., Golden, CO, USA, Tech. Rep. NREL/TP-5D00-78755, 2021.

[3] M. Moradi, Y. Weng, J. Dirkman, and Y.-C. Lai, "Preferential cyber defense for power grids," *PRX Energy*, vol. 2, no. 4, p. 43007, Oct. 2023, doi: 10.1103/prxenergy.2.043007.

[4] M. Moradi, Y. Weng, and Y.-C. Lai, "Defending smart electrical power grids against cyberattacks with deep *Q*-learning," *PRX Energy*, vol. 1, no. 3, Nov. 2022, Art. no. 033005.

[5] D. Ye, M. Zhang, and D. Sutanto, "A hybrid multiagent framework with Q-learning for power grid systems restoration," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2434–2441, Nov. 2011.

[6] T. Ahmad, R. Madonski, D. Zhang, C. Huang, and A. Mujeeb, "Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of smart grid paradigm," *Renew. Sustain. Energy Rev.*, vol. 160, May 2022, Art. no. 112128.

[7] A.-H. Mohsenian-Rad and A. Leon-Garcia, "Optimal residential load control with price prediction in real-time electricity pricing environments," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 120–133, Sep. 2010.

[8] J. E. T. Bistline and G. J. Blanford, "The role of the power sector in net-zero energy systems," *Energy Climate Change*, vol. 2, Dec. 2021, Art. no. 100045.

[9] J. B. Kucharski and H. Unesaki, "An institutional analysis of the Japanese energy transition," *Environ. Innov. Societal Transitions*, vol. 29, pp. 126–143, Dec. 2018.

[10] Y. Li, W. Xu, X. Zhang, Z. Wang, W. Gao, and Y. Xu, "System value and utilization performance analysis of grid-integrated energy storage technologies in Japan," *J. Energy Storage*, vol. 63, Jul. 2023, Art. no. 107051.

[11] *Japan Electric Power EXchange Guide*, JEPX, Tokyo, Japan, Jan. 2019.

[12] R. Weron, "Electricity price forecasting: A review of the state-of-the-art with a look into the future," *Int. J. Forecasting*, vol. 30, no. 4, pp. 1030–1081, Oct. 2014, doi: 10.1016/j.ijforecast.2014.08.008.

[13] B. Neupane, K. S. Perera, Z. Aung, and W. L. Woon, "Artificial neural network-based electricity price forecasting for smart grid deployment," in *Proc. Int. Conf. Comput. Syst. Ind. Informat.*, Dec. 2012, pp. 1–6.

[14] D. Singhal and K. S. Swarup, "Electricity price forecasting using artificial neural networks," *Int. J. Elect. Power Energy Syst.*, vol. 33, no. 3, pp. 550–555, Mar. 2011.

[15] E. Ceperic, V. Ceperic, and A. Baric, "A strategy for short-term load forecasting by support vector regression machines," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 4356–4364, Nov. 2013.

[16] N. O'Connell, P. Pinson, H. Madsen, and M. O'Malley, "Benefits and challenges of electrical demand response: A critical review," *Renew. Sustain. Energy Rev.*, vol. 39, pp. 686–699, Nov. 2014.

[17] D. Saini, A. Saxena, and R. C. Bansal, "Electricity price forecasting by linear regression and SVM," in *Proc. Int. Conf. Recent Adv. Innov. Eng. (ICRAIE)*, Dec. 2016, pp. 1–7.

[18] A. Pourdaryaei, H. Mokhlis, H. A. Illias, S. H. A. Kaboli, S. Ahmad, and S. P. Ang, "Hybrid ANN and artificial cooperative search algorithm to forecast short-term electricity price in de-regulated electricity market," *IEEE Access*, vol. 7, pp. 125369–125386, 2019.

[19] H. Yamin, S. Shahidehpour, and Z. Li, "Adaptive short-term electricity price forecasting using artificial neural networks in the restructured power markets," *Int. J. Electr. Power Energy Syst.*, vol. 26, no. 8, pp. 571–581, Oct. 2004.

[20] R. Weron and A. Misiorek, "Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models," *Int. J. Forecasting*, vol. 24, no. 4, pp. 744–763, Oct. 2008.

[21] K. Ofuji and S. Kanemoto, "Price forecasting of Japan electric power exchange using time-varying AR model," in *Proc. Int. Conf. Intell. Syst. Appl. Power Syst.*, Dec. 2007, pp. 448–453.

[22] A. Cruz, A. Muñoz, J. L. Zamora, and R. Espínola, "The effect of wind generation and weekday on Spanish electricity spot price forecasting," *Electr. Power Syst. Res.*, vol. 81, no. 10, pp. 1924–1935, Oct. 2011.

[23] S. K. Aggarwal, L. M. Saini, and A. Kumar, "Electricity price forecasting in deregulated markets: A review and evaluation," *Int. J. Electr. Power Energy Syst.*, vol. 31, no. 1, pp. 13–22, Jan. 2009.

[24] J. Nowotarski and R. Weron, "Recent advances in electricity price forecasting: A review of probabilistic forecasting," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1548–1568, Jan. 2018.

[25] X. Zhang, "A model combining LightGBM and neural network for high-frequency realized volatility forecasting," in *Proc. 7th Int. Conf. Financial Innov. Econ. Develop. (ICFIED)*, 2022, pp. 2906–2912.

[26] T. Hong, P. Pinson, Y. Wang, R. Weron, D. Yang, and H. Zareipour, "Energy forecasting: A review and outlook," *IEEE Open Access J. Power Energy*, vol. 7, pp. 376–388, 2020, doi: 10.1109/OAJPE.2020.3029979.

[27] V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, "Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines," *Ore Geol. Rev.*, vol. 71, pp. 804–818, Dec. 2015.

[28] J. Lago, G. Marcjasz, B. De Schutter, and R. Weron, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Appl. Energy*, vol. 293, Jul. 2021, Art. no. 116983, doi: 10.1016/j.apenergy.2021.116983.

[29] S. S. Reddy, C.-M. Jung, and K. J. Seog, "Day-ahead electricity price forecasting using back propagation neural networks and weighted least square technique," *Frontiers Energy*, vol. 10, no. 1, pp. 105–113, Mar. 2016.

[30] S. Anbazhagan and B. Ramachandran, "Day-ahead electricity price forecasting model based on artificial neural networks for energy markets," *EAI Endorsed Trans. Energy Web*, vol. 8, no. 33, p. e13, Jul. 2021, Art. no. 167660.

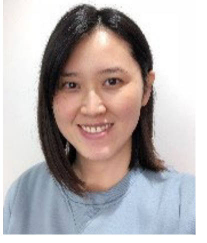
[31] W. Gao, A. Darvishan, M. Toghani, M. Mohammadi, O. Abedinia, and N. Ghadimi, "Different states of multi-block based forecast engine for price and load prediction," *Int. J. Electr. Power Energy Syst.*, vol. 104, pp. 423–435, Jan. 2019.

[32] D. C. Sansom, T. Downs, and T. K. Saha, "Evaluation of support vector machine based forecasting tool in electricity price forecasting for Australian national electricity market participants," *J. Elect. Electron. Eng., Aust.*, vol. 22, no. 3, pp. 227–233, 2003.

[33] A. Aldoseri, K. N. Al-Khalifa, and A. M. Hamouda, "Re-thinking data strategy and integration for artificial intelligence: Concepts, opportunities, and challenges," *Appl. Sci.*, vol. 13, no. 12, p. 7082, Jun. 2023, doi: 10.3390/app13127082.

[34] T. S. Talagala, F. Li, and Y. Kang, "FFORMPP: Feature-based forecast model performance prediction," *Int. J. Forecasting*, vol. 38, no. 3, pp. 920–943, Jul. 2022.
[35] L. R. Visser, M. E. Kootte, A. C. Ferreira, O. Sicurani, E. J. Pauwels, C. Vuik, W. G. J. H. M. Van Sark, and T. A. AlSkaif, "An operational bidding framework for aggregated electric vehicles on the electricity spot market," *Appl. Energy*, vol. 308, Feb. 2022, Art. no. 118280.
[36] S. Kawashima and F. Takeda, "The effect of the Fukushima nuclear accident on stock prices of electric power utilities in Japan," *Energy Econ.*, vol. 34, no. 6, pp. 2029–2038, Nov. 2012.
[37] Tokyo Electric Power Company (TEPCO). *TEPCO Electricity Forecast*. Japan. Accessed: Apr. 20, 2022. [Online]. Available: https://www.tepco.co.jp/forecast/
[38] D. J. Higham and N. J. Higham, *MATLAB Guide*. Philadelphia, PA, USA: SIAM, 2016.
[39] R. Collobert, K. Kavukcuoglu, and C. Farabet, "Torch7: A MATLAB-like environment for machine learning," in *Proc. NIPS Workshop*, 2011, pp. 1–6. [Online]. Available: https://www.academia.edu/2844564/Torch7_A_Matlab_like_Environment_for_Machine_Learning
[40] V. Asnaghi, D. Pecorino, E. Ottaviani, A. Pedroncini, R. M. Bertolotto, and M. Chiantore, "A novel application of an adaptable modeling approach to the management of toxic microalgal Bloom events in coastal areas," *Harmful Algae*, vol. 63, pp. 184–192, Mar. 2017.
[41] S. Navlakha, Z. Bar-Joseph, and A. L. Barth, "Network design and the brain," *Trends Cognit. Sci.*, vol. 22, no. 1, pp. 64–78, Jan. 2018, doi: 10.1016/j.tics.2017.09.012.
[42] E. D. Übeyli and I. Güler, "Neural network analysis of internal carotid arterial Doppler signals: Predictions of stenosis and occlusion," *Expert Syst. Appl.*, vol. 25, no. 1, pp. 1–13, Jul. 2003.
[43] A. Maier, C. Syben, T. Lasser, and C. Riess, "A gentle introduction to deep learning in medical image processing," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 86–101, May 2019.
[44] A. Dabouei, H. Kazemi, S. M. Iranmanesh, J. Dawson, and N. M. Nasrabadi, "Fingerprint distortion rectification using deep convolutional neural networks," in *Proc. IEEE Int. Conf. Biometrics*, Feb. 2018, pp. 1–8.
[45] L. N. Jahan, T. A. Munshi, S. S. Sutradhor, and M. Hashan, "A comparative study of empirical, statistical, and soft computing methods coupled with feature ranking for the prediction of water saturation in a heterogeneous oil reservoir," *Acta Geophysica*, vol. 69, no. 5, pp. 1697–1715, Oct. 2021.
[46] B. Lantz, *Machine Learning With R: Expert Techniques for Predictive Modeling*. Birmingham, U.K.: Packt Publishing, 2019.
[47] G. Fu, C. Liu, R. Zhou, T. Sun, and Q. Zhang, "Classification for high resolution remote sensing imagery using a fully convolutional network," *Remote Sens.*, vol. 9, no. 5, p. 498, May 2017.
[48] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, p. e623, Jul. 2021.
[49] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature," *Geosci. Model Develop.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014.
[50] X. Fang, J. Cui, B. Jie, Y. Tu, T. Oozeki, and Y. Ueda, "Relationship between the PV contribution to the electricity demand and spot market price," in *Proc. 19th Next-Generat. Photovoltaic Power Generat. Syst., Symp., 2nd Annu. Meeting Jpn. Photovoltaic Energy Soc.*, vol. 2, 2022, p. 40, doi: 10.57295/jpvsproc.2.0_40.
[51] Japan Meteorological Agency. *Weather, Climate & Earthquake Information*. Accessed: Apr. 20, 2022. [Online]. Available: https://www.data.jma.go.jp/risk/obsdl/index.php
[52] JEPX. *Trading Information: Spot Market/Intraday Market*. Accessed: Apr. 20, 2022. [Online]. Available: http://www.jepx.org/market/index.html

XUE FANG received the Ph.D. degree in environmental engineering from The University of Kitakyushu, Japan, in 2021. She is currently a Project Researcher with the Research Institute for Science and Technology, Tokyo University of Science, Japan. Her research interests include the electricity market, solar energy, and forecasting methodology.



JINDAN CUI received the Ph.D. degree in engineering from Yokohama National University, Japan, in 2013. From 2013 to 2014, she was a PD Researcher with Yokohama National University, where she focused on the operation plan of demand response. In 2015, she joined as a Project Researcher with the EMS Project CREST, JST, Tokyo University of Science, where she focused on the optimization operation of battery systems in residential groups with PV systems. Since 2021, she has been an Assistant Professor with Tokyo University of Science, where she focuses on creating reserve power to bid the electric power market to increase the value of PV power plants.

TAKASHI OOZEKI received the Ph.D. degree in engineering from Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2005. He has been with the National Institute of Advanced Industrial Science and Technology, Japan, since 2005. His research interest includes photovoltaic system technologies.

YUZURU UEDA (Member, IEEE) received the B.S. degree in physics from Shinshu University, Nagano, Japan, in 1995, and the Ph.D. degree in engineering from Tokyo University of Agriculture and Technology, Tokyo, Japan, in 2007. From 1995 to 2003, he was with Applied Materials Japan Inc., Tokyo, as a Process Engineer of semiconductor wafer plasma processing. From 2008 to 2014, he was an Assistant Professor with the Graduate School of Science and Engineering, Tokyo Institute of Technology, Tokyo. Since 2014, he has been with Tokyo University of Science, Tokyo, where he is currently a Professor. His research interests include photovoltaic systems, grid integration of distributed generators, and solar energy.