WILEY | Hindawi

Research Article

Zero-Shot Photovoltaic Power Forecasting Scheme Based on a Deep Learning Model and Correlation Coefficient

Sungwoo Park [iD], **Dongjun Kim** [iD], **Jaeuk Moon** [iD], **and Eenjun Hwang** [iD]

School of Electrical Engineering, Korea University, Seoul 02841, Republic of Korea

Correspondence should be addressed to Eenjun Hwang; ehwang04@korea.ac.kr

Accurate photovoltaic (PV) power forecasting is essential for the stable and reliable operation of PV power generation systems. Recently, various deep learning- (DL-) based forecasting models have been proposed for accurate forecasting, but newly built systems cannot benefit from them due to the absence of PV power data. Although zero-shot methods based on single site can be used for PV power forecasting, they suffer from performance degradation problems when the characteristics of the source data and target data are different. To address this issue, we propose a novel zero-shot PV power forecasting scheme that leverages historical data from multiple PV generation systems at different sites. The proposed scheme constructs an individual forecasting model using historical data from each PV generation system. Then, two correlation coefficients are calculated for each forecasting model: one based on the correlation between the input variables of the source data and target data and the other on the correlation between the input variables and output variables of the source data. Lastly, the final forecasting value is calculated as a weighted sum of the predicted values of the constructed forecasting models for the input variables of the target data. In the extensive experiments for diverse DL models for forecasting, correlation coefficient types for weights, and data time intervals, the combination of recurrent neural network, Pearson's correlation coefficient, and solar-noon time yielded the best prediction performance, with an improvement of up to 34.47% in mean absolute error and up to 15.94% in root mean square error compared to the best single-site zero-shot prediction. In addition, in experiments on PV power data from 9 cities in Korea using this combination, the proposed scheme achieved the best predictive performance in almost all cases and the second-best performance with a very narrow margin only in a few cases.

1. Introduction

As the global population increases from 7.8 billion in 2020 to 9.9 billion by 2050, global energy demand is expected to increase by about 50% [1]. If such energy demand is met by thermal power generation using fossil fuels, the problem of environmental pollution can become even more severe [2]. To address this issue, countries worldwide have signed the Paris Climate Agreement, aiming to keep the rise in the average global temperature well below 2°C above preindustrial levels and preferably limit the increase to 1.5°C. Since then, many efforts have been made to increase the share of renewable energy in power generation worldwide [3].

Renewable energy derived from natural sources replenished at a higher rate than consumed includes solar, wind, hydro, geothermal, and biomass energy [4]. These energy types have low environmental influence and reduce dependence on fossil fuels, leading to a more sustainable future [5]. Among them, solar energy has gained more attention and popularity for several reasons, such as no risk of energy source depletion, fewer restrictions on solar panel installation, and various panel capacities [6]. Thus far, many photovoltaic (PV) generation systems have been built, with an average annual growth rate of 48% [7]. Moreover, the recent price reduction of PV panels has made PV generation more cost-effective and accessible to more individuals and businesses [8].

However, weather dependence remains a significant challenge for PV generation. The performance of the PV panel and output of the PV generation system is greatly affected by weather conditions, such as cloud cover, rain,

snow, and dust, reducing the stability and reliability of the PV generation system [9]. Accurate PV power forecasting is necessary to solve this problem [10].

Recently, various forecasting models based on machine learning (ML) and deep learning (DL) have performed better than traditional statistical models, such as autoregressive and moving average models [11, 12]. However, these models require vast historical data for training [13]. If the data are insufficient for training, the model may become too sensitive to noise in the training data, leading to overfitting. This cold-start problem can significantly degrade the forecasting performance [14], especially in newly built PV generation systems [15].

Zero-shot learning has been proposed to produce results, even without historical data [16]. However, there are very few cases where zero-shot learning has been applied to forecasting problems, especially in PV power forecasting [17]. Single-site-based zero-shot PV power forecasting (SZF) trains a model using source data and then makes predictions on target data without any fine-tuning. Here, source data and target data refer to a sufficient amount of data collected from an existing PV power generation system and data collected from a newly built PV power generation system, respectively. The most serious problem with this approach is that if the data distribution, pattern, and trend of the source data and the target data are different, the predictive performance may decrease significantly. To solve this problem, we propose a zero-shot PV power forecasting scheme that leverages historical data from multiple PV generation systems at different sites. In this scheme, an individual forecasting model is constructed using historical data from each PV generation system. Then, weights for each forecasting model are determined based on the correlation between the input variables of the source data and target data and the input variables and output variables of the source data. Lastly, the final prediction value is calculated as a weighted sum of the predicted values of the source forecasting models for the input variables of the target data.

In our experiments, we first tested different combinations of DL models, correlation coefficients, and data time options to find the best one and then used it across multiple data to evaluate the effectiveness of the proposed scheme. Experimental results showed that the combination of the recurrent neural network (RNN) model, Pearson's correlation coefficient (PCC), and solar-noon time option performed best. Figure 1 illustrates the overall process of the proposed scheme.

The main contributions of this paper are as follows:

(1) We proposed a novel zero-shot PV power forecasting scheme based on the DL model and correlation coefficient. To the best of our knowledge, this is the first effort to forecast the PV generation of newly built PV systems without historical PV power data

(2) We analyzed the effects of one-dimensional (1D) and two-dimensional (2D) forms of time data to determine the most effective form for PV power forecasting

(3) We conducted extensive experiments to find the most effective combination of the DL model, correlation coefficient, and time options for zero-shot PV power prediction

(4) We showed that the proposed scheme can guarantee good prediction performance, unlike the SZF model. This can contribute to the effective and accurate evaluation of the feasibility of PV generation projects or construction

The paper is organized as follows. Section 2 introduces several studies on the ML- and DL-based PV power forecasting models. Section 3 details the proposed scheme, and Section 4 describes input variable configurations for constructing the forecasting model. Next, Section 5 discusses the experimental results of evaluating the proposed scheme's performance. Finally, Section 6 summarizes this study and presents the conclusions.

2. Related Works

So far, many studies have been conducted to perform PV power forecasting efficiently. In particular, various ML and DL algorithms have been proposed recently to build more accurate PV power forecasting models. For instance, Pan et al. [18] proposed a forecasting model based on the support vector machine (SVM). In particular, they optimized the selection policies for the SVM hyperparameters by incorporating max–min ant colony optimization (ACO), differential evolutionary algorithm, and adaptive learning factor into the ACO algorithm. In addition, an ensemble filter algorithm was employed to remove abnormal data, and L2-norm regularization was applied to address standardization issues in the data. They demonstrated that the proposed model is significantly better than the comparative models and exhibits significantly improved performance in nighttime and peak power prediction. Meng and Song [19] proposed a daily PV power forecasting model based on random forest (RF). They constructed the RF model using data collected from the Zhonghe PV station in North China's central region. They used PM 2.5 as an additional input variable due to the severe winter air pollution in North China. The proposed model categorizes winter days into three types based on climate characteristics and creates an RF model for each type. They demonstrated that the proposed model outperforms the SVM, elastic net, and gradient-boosting decision tree for nearly all error evaluation metrics. Li et al. [20] proposed a very short-term PV power forecasting model based on the RNN. The PV power data were divided into inter- and intraday data, and the RNN was used to discover nonlinear features and invariant structures. The proposed model was extensively compared with various ML and DL models using PV power data collected in Flanders, Belgium. The results indicated that the proposed model outperforms the comparative models (e.g., persistence and SVM) in 15 and 30 min forecasting horizons. Hossain and Mahmood [21] proposed a day-ahead PV power forecasting model based on long short-term memory

$$Corr_i = CC_{Time}(X_{Test}, Y_{Train,i}) \times CC_{Time}(X_{Train,i}, Y_{Train,i})$$

$$W_i = Corr_i / \sum_{i=1}^{N}(Corr_i)$$
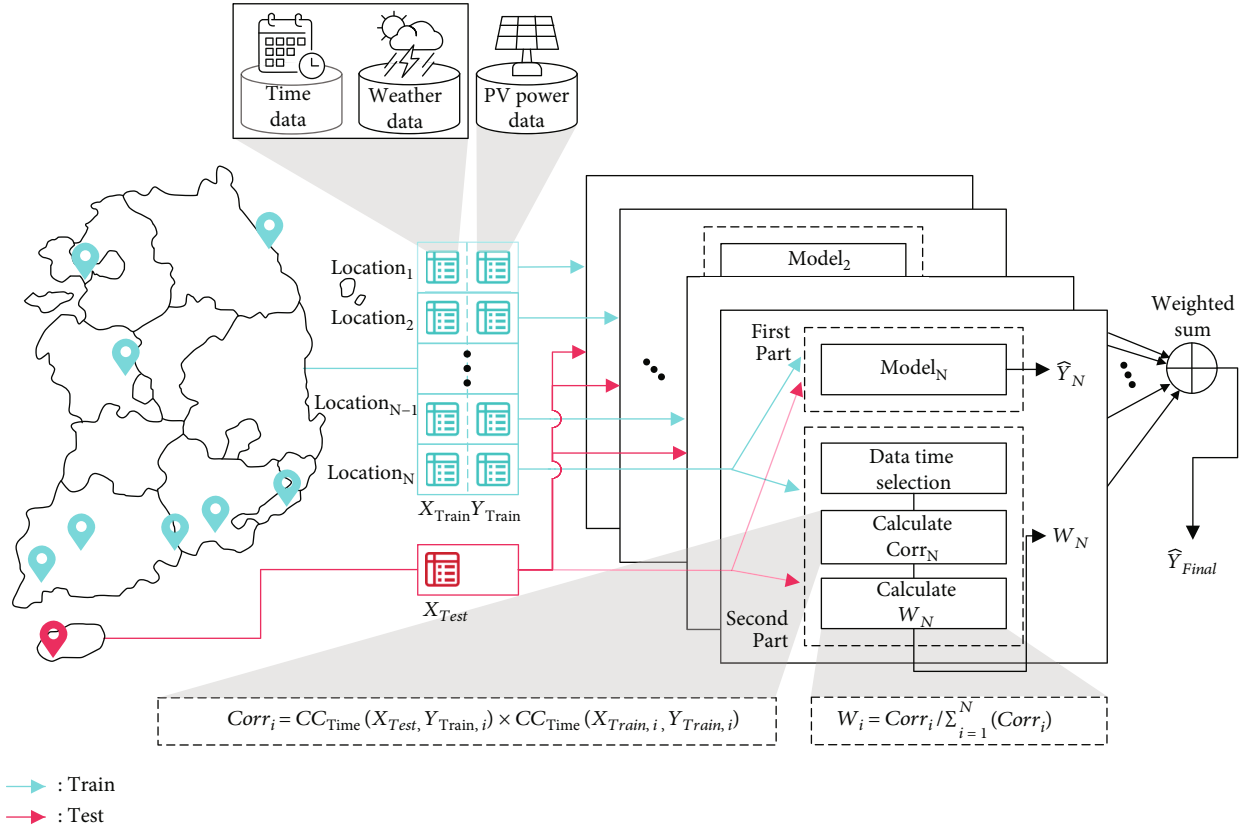
→ : Train
→ : Test

FIGURE 1: Overall structure of the proposed scheme.

(LSTM). They integrated the statistical knowledge from historical irradiance data and the sky forecast to create a synthetic weather forecast to use as input for the LSTM. The performance of the LSTM model was analyzed using a synthetic weather forecast and real sky forecast as input variables. They revealed that synthetic weather forecasts significantly improve accuracy by more than 30% compared to original weather forecasts. Liu et al. [22] constructed an ensemble model using SVM, multilayer perceptron (MLP), and multivariate adaptive regression spline (MARS) as stand-alone models to improve the weights of sub-stand-alone models through a recursive arithmetic average process. The proposed ensemble model generally performed better than stand-alone models using historical data from the Australian technology demonstration facility. Li et al. [23] proposed a hybrid DL approach for PV power forecasting that combines a convolutional neural network (CNN) with an LSTM model. The CNN structure extracts nonlinear features and invariant patterns in past PV output data, and the LSTM structure models the temporal changes in the recent PV power data to make predictions for the next time step. In the comprehensive evaluation using PV power data from Limburg, Belgium, the proposed model performed better than other comparative models with significantly less prediction error.

ML and DL models exhibit good performance in PV power forecasting. Table 1 summarizes some of the related studies on PV power forecasting using ML and DL models.

However, these models cannot be used for newly built PV generation systems because they require data covering at least one year (12 months) for training. One possible way to solve this problem is to train a model using data collected from an existing PV power system and use the trained model on a newly built PV power system. However, if the two sites have different characteristics for PV generation, the predictive performance may be poor. This can be solved by effectively utilizing data collected from PV power systems at various sites.

3. Proposed Scheme

In this section, we first briefly explain the theoretical background of the DL model and correlation coefficient and then present the details of the proposed scheme. In general, state-of-the-art models such as transformer and temporal convolutional network-based models perform well in long-term time-series forecasting due to their strong long-term dependencies. However, in PV power forecasting, short-term dependencies play a more important role than long-term dependencies. Moreover, these models typically take a long training time due to complex structures or have low convergence stability. Hence, in this paper, we consider five popular artificial neural network (ANN) models: shallow neural network (SNN), deep neural network (DNN), RNN, LSTM, and gated recurrent units (GRU).

Table 1: Summary of related work.

Author (year)	Data resolution	Data duration	Methodology	Description
Pan et al. (2020) [18]	5 min	23 months	SVM	The proposed model uses an ensemble filter algorithm, L2-norm regularization, and SVM optimized using ACO to improve the accuracy of PV power forecasting.
Meng and Song (2020) [19]	1 day	25 months	RF	The proposed model divides data into three categories considering climatic characteristics and uses RF for each categorized data.
Li et al. (2019) [20]	15 min	12 months	RNN	The proposed model divides data into inter- and intraday data and applies the RNN to discover nonlinear features and invariant structures.
Hossain and Mahmood (2020) [21]	1 h	36 months	LSTM	The proposed model creates synthetic weather forecast data using the K-means algorithm and uses it as an input variable to the LSTM model.
Liu et al. (2019) [22]	5 min	58 months	SVM, MLP, MARS	The proposed model uses an ensemble modeling technique integrating SVM, MLP, and MARS within a recursive arithmetic average model.
Li et al. (2020) [23]	15 min	12 months	CNN, LSTM	The proposed model combines the CNN with the LSTM model for extracting nonlinear features and invariant patterns and modeling the temporal changes in data.

The correlation coefficient indicates the strength of the relationship between the two variables. If the product of the correlation coefficient between the input and output variables of the source data and between the input variables of the source data and the input variables of the target data is large, the model trained using the source data is more likely to be effective on the target data as well. In this study, the weights of forecasting models trained using different source data are determined based on the correlation coefficients so that the prediction values of forecasting models with higher correlations can be more reflected in the final prediction value. Among various correlation coefficients, PCC best expresses the relationship between two continuous variables. However, PCC can only detect linear relationships. Therefore, in addition to PCC, we also consider the distance correlation coefficient (DCC) to detect nonlinear relationships.

3.1. *Deep Learning Models.* Among the aforementioned models, the SNN usually refers to an ANN with only one hidden layer [24]. It is called shallow because it does not have multiple layers to learn representations from the input data. In contrast, the DNN refers to an ANN with multiple hidden layers, usually more than two [25]. DNNs are designed to learn complex representations from the input data by leveraging the depth of the network. Each layer of a DNN is trained independently and learns to recognize different features from the input data. Due to its simplicity and computational efficiency, the SNN is primarily used to solve simple problems, while the DNN solves more complex problems that the SNN cannot solve.

Unlike general ANN structures, the RNN, LSTM, and GRU are specifically designed to handle sequence data. These models primarily process sequence data, such as time series, text, and speech. RNN has a loop structure, allowing the retainment of information from previous steps [26]. This feature enables the RNN to process data sequences by considering the context from previous time steps. The RNN is used for various applications, including sentiment analysis, language translation, and speech recognition. However, the RNN has difficulty retaining information over long sequences due to the vanishing gradient problem. To address this problem, the LSTM model was introduced as a variant of the RNN [27]. The LSTM has a unique structure called memory cells that enable the LSTM to retain information over long sequences. The memory cells are controlled by gates determining when information should be stored, forgotten, or used for predictions. This structure enables the LSTM to handle sequence data effectively. However, the complexity of its structure and the number of parameters can make it computationally inefficient. The GRU approach was introduced as a simpler alternative to the LSTM. The GRU model has a structure similar to the LSTM but has fewer parameters and, therefore, is computationally more efficient [28]. The GRU method uses two gates, the update and reset gates, to control the flow of information in the network. The update gate determines how much past information should be retained, whereas the reset gate determines how much past information should be forgotten. Figure 2 illustrates a simplified structure of the ANN models.

3.2. *Correlation Coefficient.* An input variable, also known as an independent variable, is used to predict the value of an output variable. Understanding the relationship between input and output variables is essential for accurate forecasting
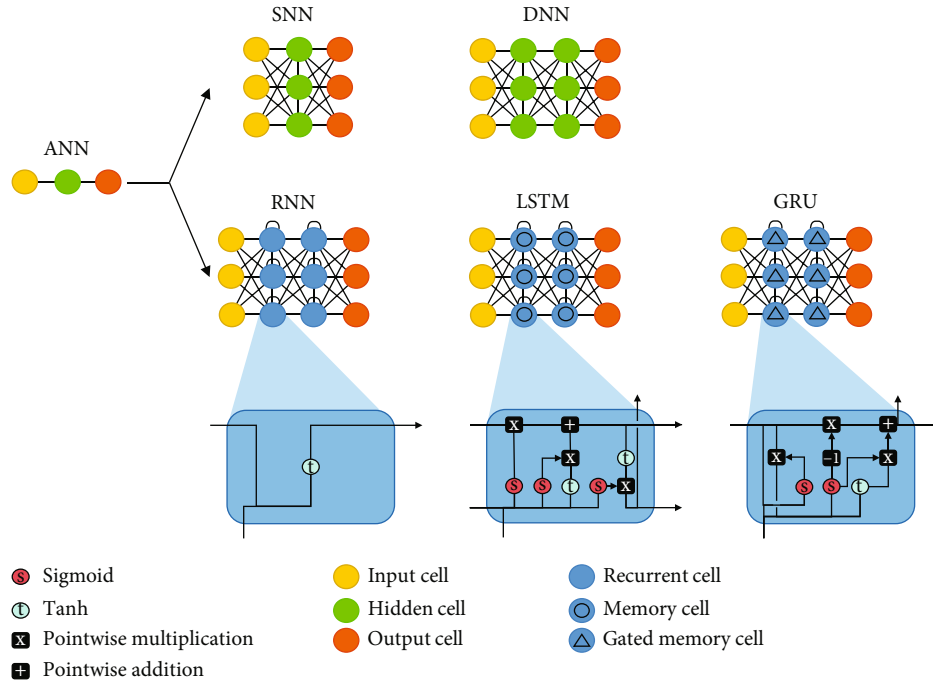
FIGURE 2: Simplified structure of the artificial neural network.

models, enabling accurate forecasting of the future by correctly selecting the appropriate input variables and model structure. Correlation coefficients can be used to determine the relationship between an input and output variable. The most commonly used correlation coefficient is PCC [29]. It measures the strength and direction of the linear association between two variables, with values ranging from -1 to 1. The PCC is defined by Equation (1), where n is the number of observations, X_i and Y_i denote the values at time i, and $\bar{X}$ and $\bar{Y}$ represent the mean values of X and Y, respectively.

$$\text{PCC}(X, Y) = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) \bullet (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \bullet \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}. \quad (1)$$

The PCC is calculated as the covariance of the two variables divided by the product of their standard deviations. Covariance measures how two variables change together, whereas the standard deviation measures the spread of a single variable. A positive PCC indicates that as the value of the input variable increases, the value of the output variable also increases. A negative PCC indicates that as the value of the input variable increases, the value of the output variable decreases. A PCC of 0 indicates no linear relationship between the two variables. The PCC can detect a linear relationship between two variables well, but many nonlinear relationships exist in actual data.

Unlike the PCC, the distance correlation coefficient (DCC) can capture nonlinear relationships between variables [30]. The DCC is the square root of the distance covariance between two variables. The distance covariance measures the combined variability of two variables, consid-

ering the linear and nonlinear dependence between the variables. The DCC ranges from 0 to 1, with values close to 0 indicating independence between the two variables and values close to 1 indicating strong dependence. The DCC is defined by the following equations.

$$\text{DCC}(X, Y) = \frac{d\text{Cov}(X, Y)}{\sqrt{d\text{Var}(X) \times d\text{Var}(Y)}}, \quad (2)$$

$$d\text{Cov}(X, Y) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} D(X_i, X_j) \bullet D(Y_i, Y_j), \quad (3)$$

$$d\text{Var}(X) = d\text{Cov}(X, X) = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} D(X_i, X_j)^2. \quad (4)$$

In the equations, dCov denotes the distance covariance, dVar denotes the distance variance, and $D(X_i, X_j)$ represents the Euclidean distance between X_i and X_j, respectively.

The PCC and DCC have strengths and weaknesses. The choice between them depends on the specific analysis needs and relationship characteristics between variables. It is best to consider both methods and choose the one that best fits the data and research question.

3.3. Forecasting Process. This section describes the forecasting process of the proposed scheme in more detail. Algorithm 1 shows the process of calculating the final forecasting value.

A brief description of the algorithm is as follows.

Input: No. of sites n, Source input data $X[1...n]$, Source output data $Y[1...n]$, Target input data X_{Test}
Output: Final forecast value $\widehat{Y}_{Final}$
1   Select a DL model type M
2   Select correlation coefficient type CC
3   Select data time interval T
4   **for each** site i **do**
5   Train M_i using $(X[i], Y[i])$
6            $\widehat{Y}_i \longleftarrow M_i(X_{Test})$
7   **end for**
8   **for each** site i **do**
9            $Corr_i \longleftarrow CC(T(X_{Test}), T(X[i])) * CC(T(X[i]), T(Y[i]))$
10   **end for**
11     $Corr\_Sum \longleftarrow \sum_{i=1}^{n}(Corr_i)$
12   **for each** i **do**
13           $W_i \longleftarrow Corr_i/Corr\_Sum$
14   **end for**
15     $\widehat{Y}_{Final} \longleftarrow \sum_{i=1}^{n}(\widehat{Y}_i \times W_i)$
16   **return** $\widehat{Y}_{Final}$

ALGORITHM 1: Calculating final forecasting value.

(i) Lines 4-7 construct a source forecasting model for each site and calculate its predictive value for the target data

(ii) Lines 8-10 first calculate the two correlation coefficients for each site and then compute their product

(iii) Lines 11-14 calculate the weight of each site based on the correlation coefficient. This is done by summing the correlation coefficients of all sites and dividing each site's correlation coefficient by this sum. This makes the sum of the correlation coefficients for all sites 1

(iv) Line 15 calculates the final forecasting value based on the predicted values of all forecasting models for the target data and their weights

4. Input Variable Configuration

For the experiments in this study, we used public PV power data from the Korean Open Data Portal (KODP). The KODP provides 1 h resolution PV power data collected from PV systems in various locations. To demonstrate the robustness of the proposed scheme, we considered PV generation systems in regions with diverse characteristics, such as inland, coastal, and island regions in Korea. Hence, we collected data from nine regions from January 1, 2017, to December 31, 2019. The statistics of the collected PV power data are summarized in Table 2.

From the table, we can see that the capacities of the PV power systems vary greatly. In order to utilize such data in the proposed method, it is necessary to scale it according to the capacity of the power system. This is because if the scale of the source power generation system of the data used for training and the scale of the newly constructed power generation system are significantly different, the predicted value of the forecasting model trained with the source data is not suitable for the new system due to the difference in scale. For this purpose, we performed scaling that divides the PV power data by the capacity of the corresponding PV power system.

The PV generation is directly affected by solar radiation that reaches the PV panel surface. The actual amount of solar radiation reaching the panel is influenced by various weather data, such as cloud cover, humidity, and atmospheric conditions. Therefore, we considered time and weather data for input variables in this study. The time data determine the maximum solar radiation on the PV panel.

In addition, we use month, day, and hour to represent time information. These time data were recorded as 1D data, making it difficult to reflect their periodic nature in ML and DL algorithms [31]. To address this issue, we transformed the 1D time data into 2D data using periodic functions and used them together. Additionally, we apply a proper time delay to the time data to improve the prediction performance. This can be an effective preprocessing technique for forecasting models because it can facilitate capturing the relationship between the input and output variables more accurately. Equations (5) and (6) represent periodic functions to transform 1D time data into 2D time data.

$$\text{Time}_X = \sin\left((360/\text{cycle}) \times (\text{Time} - \text{delay})\right) \tag{5}$$

$$\text{Time}_Y = \cos\left((360/\text{cycle}) \times (\text{Time} - \text{delay})\right). \tag{6}$$

Here, cycle and delay represent data period and time delay, respectively. For example, if Time represents the month, cycle would be 12, and if Time represents the hour, it would be 24. We calculated the PCC between time and PV power data to determine the optimal time delay. Figure 3 shows a heatmap representing the PCC between time delay and PV power data. The upper heatmap in the figure displays the PCC between the monthly time delay and PV power data, whereas the lower heatmap displays

TABLE 2: Statistics of the PV power data.

| Statistics | Location | | | | | | | | |
	Busan	Gangneung	Gwangju	Hadong	Incheon	Jeju	Jinju	Mokpo	Sejong
Mean	28.413	164.325	112.693	6.093	153.439	24.312	136.762	2099.293	250.782
Standard error	0.267	1.574	1.079	0.059	1.465	0.243	1.259	19.571	2.385
Standard deviation	43.34	255.216	174.906	9.516	237.565	39.427	204.119	3172.688	386.58
Kurtosis	0.41	0.57	0.533	0.639	0.722	1.197	0.174	0.462	0.473
Skewness	1.348	1.404	1.392	1.423	1.426	1.565	1.279	1.356	1.368
Observation	26280	26280	26280	26280	26280	26280	26280	26280	26280
Min value	0	0	0	0	0	0	0	0	0
Max value	163.44	980.832	689.04	38.004	968.928	161.288	733.68	12384	1485.48
Capacity	187	1065	806.4	41	1000	196	905	13300	1628
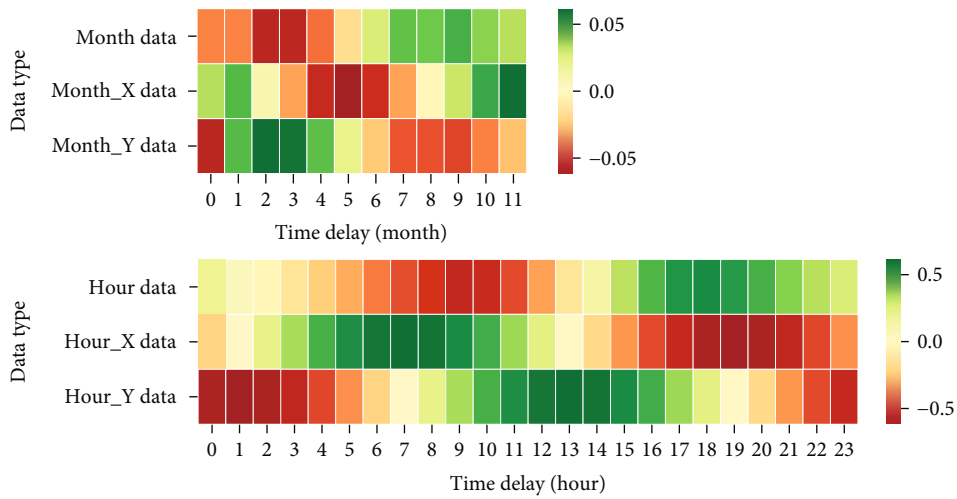
FIGURE 3: Correlation heatmap between the time and photovoltaic power data.

the PCC between the hourly time delay and PV power data. The Y-axis of the heatmap represents the Time, Time_X, and Time_Y from top to bottom, whereas the X-axis represents the amount of delay.

For 1D data, PCC showed the largest absolute value when a time delay of 2 months was applied to month data and when a time delay of 18 h was applied to hour data. In the case of 2D data, the PCC showed the largest absolute value when a 5-month delay was applied to month data and a 13-hour delay was applied to hour data. We also investigated the time delay at which the sum of the absolute values of the PCC of each 2D time data was maximum and found that a 4-month delay for month data and a 4-hour delay for hour data achieved the largest sum. Table 3 summarizes the regression statistics for various time data settings, where bold values indicate the best performance for each regression statistic.

The table shows that using 2D time data is more effective than using 1D time data, and using both 1D and 2D time data is most effective. In addition, using the time delay that achieved the highest absolute value of the PCC for 1D and 2D data achieves better performance than other settings.

As mentioned earlier, weather conditions are closely related to PV power generation. Hence, we collected weather data from the Korea Meteorological Administration (KMA),

which provides weather data at various resolutions. To construct a forecasting model, we used nine types of 1 h resolution weather data related to PV power generation, including temperature, precipitation, relative humidity, sunshine hours, and solar radiation. Thus, we configured nine datasets consisting of 18 input variables and one output variable. Input variables consist of nine types of time data and nine types of weather data. Figure 4 represents a correlation heatmap of the datasets. The figure shows that for PV power data, hour, hour_Y, sunshine hours, and solar radiation have a strong positive correlation, and relative humidity has a strong negative correlation.

5. Experimental Results

As metrics to evaluate the performance of the forecasting model, we used the mean absolute error (MAE) and root mean square error (RMSE), defined by Equations (7) and (8), respectively. Here, n, Y_i, and $\widehat{Y}_i$ represent data amount, actual data, and forecasted data, respectively.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left| Y_i - \widehat{Y}_i \right|, \tag{7}$$

TABLE 3: Regression statistics for time data settings.

| Time data settings		Regression statistics			
1D	2D	Multiple R	R^2	Adjusted R^2	Standard error
Original	—	0.1533	0.0235	0.0235	0.2268
$\mathrm{Max}(\mathrm{Corr}_{Time})$	—	0.6510	0.4238	0.4238	0.1743
—	Original	0.7496	0.5619	0.5619	0.1519
—	$\mathrm{Max}(|\mathrm{Corr}_{Time_X,PV}|)$ or $\mathrm{Max}(|\mathrm{Corr}_{Time_Y,PV}|)$	0.7496	0.5619	0.5619	0.1519
—	$\mathrm{Max}(|\mathrm{Corr}_{Time_X,PV}| + |\mathrm{Corr}_{Time_Y,PV}|)$	0.7496	0.5619	0.5619	0.1519
Original	Original	0.7621	0.5808	0.5808	0.1486
Original	$\mathrm{Max}(|\mathrm{Corr}_{Time_X,PV}|)$ or $\mathrm{Max}(|\mathrm{Corr}_{Time_Y,PV}|)$	0.7621	0.5808	0.5808	0.1486
Original	$\mathrm{Max}(|\mathrm{Corr}_{Time_X,PV}| + |\mathrm{Corr}_{Time_Y,PV}|)$	0.7621	0.5808	0.5808	0.1486
$\mathrm{Max}(\mathrm{Corr}_{Time})$	$\mathrm{Max}(|\mathrm{Corr}_{Time_X,PV}|)$ or $\mathrm{Max}(|\mathrm{Corr}_{Time_Y,PV}|)$	**0.7709**	**0.5943**	**0.5942**	**0.1462**
$\mathrm{Max}(\mathrm{Corr}_{Time})$	$\mathrm{Max}(|\mathrm{Corr}_{Time_X,PV}| + |\mathrm{Corr}_{Time_Y,PV}|)$	0.7709	0.5943	0.5942	0.1462
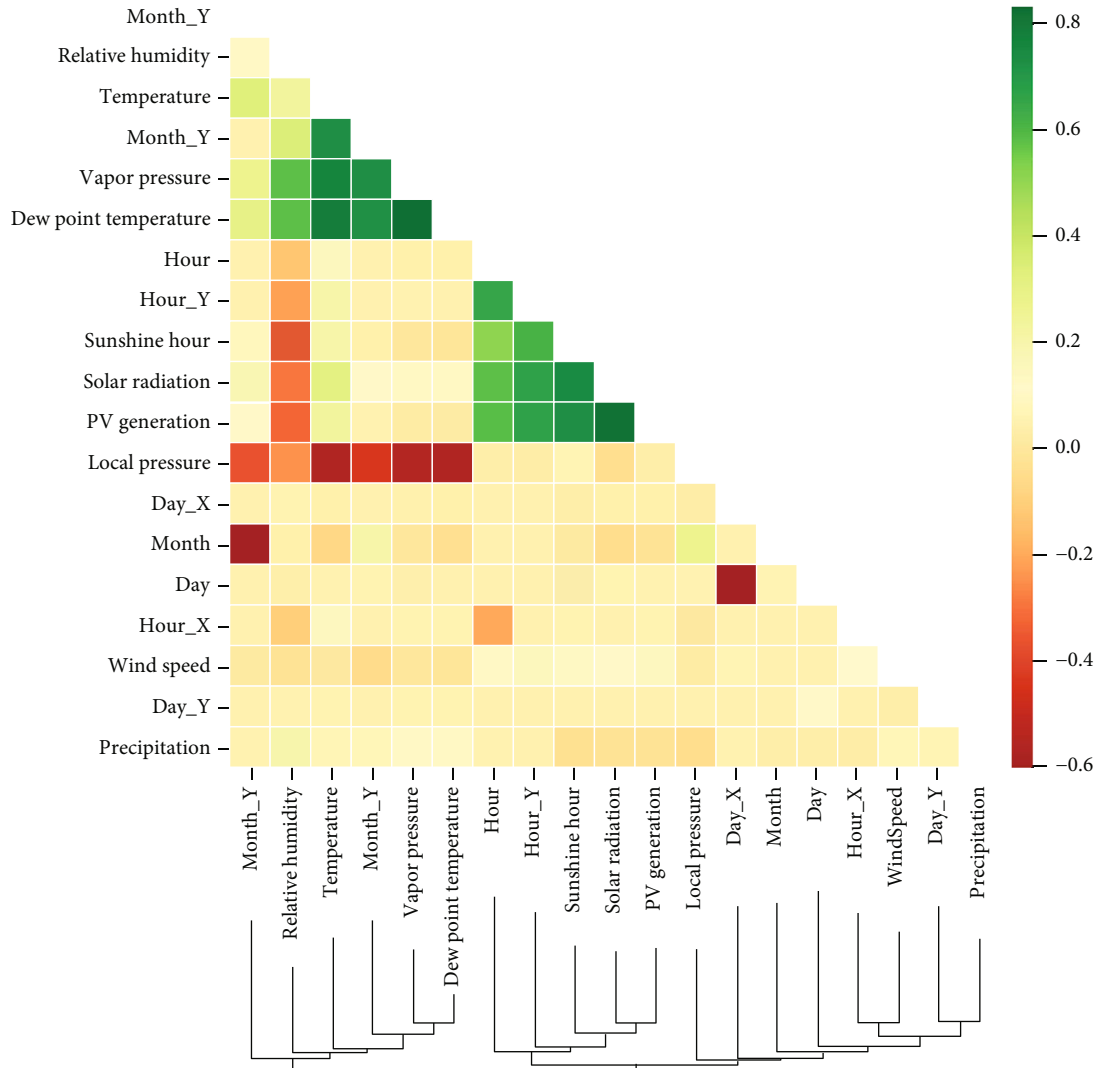
FIGURE 4: Correlation heatmap of the dataset.

Table 4: Hyperparameters used in each deep learning model.

Hyperparameter	SNN	DNN	RNN	LSTM	GRU
Window size	—	—	24	24	24
No. of hidden nodes	289	289	289	289	289
No. of hidden layers	1	5	2	2	2
Activation function	ReLU	ReLU	ReLU	ReLU	ReLU
Optimizer	Adam	Adam	Adam	Adam	Adam
Learning rate	0.000001	0.000001	0.000001	0.000001	0.000001
No. of epochs	1000	1000	1000	1000	1000

Table 5: Mean absolute error comparison of the proposed scheme using long short-term memory.

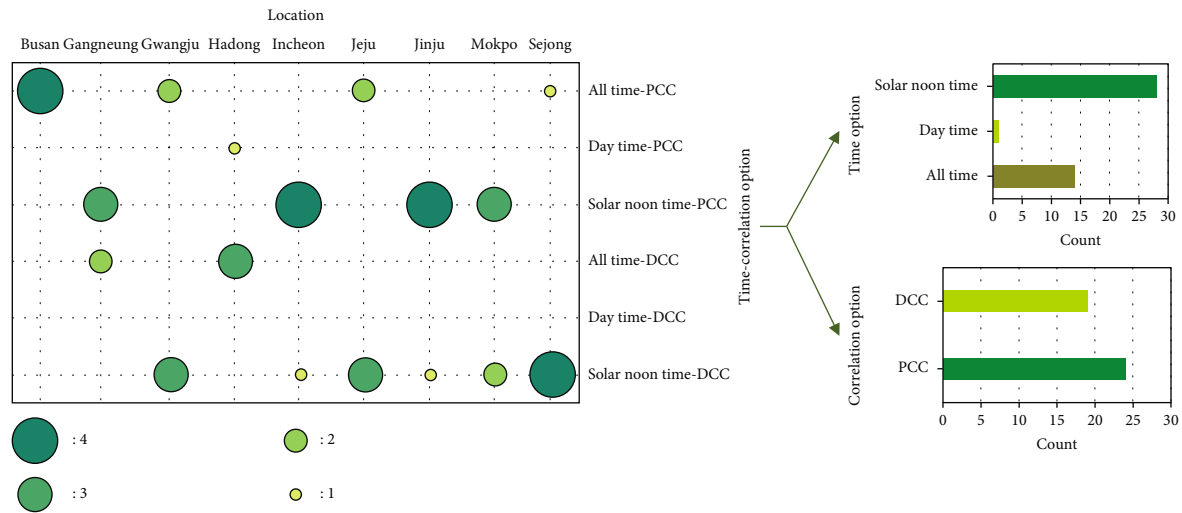|	Location								
	Busan	Gangneung	Gwangju	Hadong	Incheon	Jeju	Jinju	Mokpo	Sejong
$Scheme_{LSTM,PCC,All}$	0.061	0.050	**0.045**	0.050	0.041	**0.047**	0.037	0.041	0.042
$Scheme_{LSTM,PCC,Day}$	0.061	0.050	0.045	**0.049**	0.040	0.048	0.036	0.041	0.042
$Scheme_{LSTM,PCC,SolarNoon}$	0.062	**0.050**	0.045	0.049	**0.040**	0.048	**0.036**	**0.041**	0.042
$Scheme_{LSTM,DCC,All}$	0.061	0.050	0.045	0.049	0.041	0.047	0.037	0.041	0.042
$Scheme_{LSTM,DCC,Day}$	0.061	0.050	0.045	0.049	0.040	0.047	0.037	0.041	0.042
$Scheme_{LSTM,DCC,SolarNoon}$	0.062	0.050	0.045	0.050	0.040	0.047	0.036	0.041	**0.042**

Note. Bold: best performance result.

Figure 5: Bubble chart of the best forecasting cases and bar chart of the frequency according to the option.

Table 6: MAE comparisons of DL models.

Model type	Location								
	Busan	Gangneung	Gwangju	Hadong	Incheon	Jeju	Jinju	Mokpo	Sejong
SNN	0.062	0.042	0.042	**0.044**	0.035	0.045	0.028	0.032	0.037
DNN	0.061	0.046	0.042	0.046	0.036	0.050	0.029	0.034	0.036
RNN	**0.058**	**0.041**	**0.040**	0.045	**0.029**	**0.038**	**0.027**	**0.030**	**0.034**
LSTM	0.062	0.050	0.045	0.049	0.040	0.048	0.036	0.041	0.042
GRU	0.059	0.048	0.043	0.044	0.035	0.048	0.031	0.037	0.038

Note. Bold: best performance result; PCC and solar-noon time used.

Table 7: RMSE comparisons of DL models.

Model type	Location								
	Busan	Gangneung	Gwangju	Hadong	Incheon	Jeju	Jinju	Mokpo	Sejong
SNN	0.103	0.081	0.082	0.080	0.064	0.090	0.061	0.057	0.067
DNN	0.099	0.083	0.082	0.081	0.064	0.092	0.063	0.057	0.066
RNN	**0.097**	**0.080**	**0.079**	**0.079**	**0.057**	**0.079**	**0.058**	**0.056**	**0.063**
LSTM	0.102	0.089	0.084	0.087	0.070	0.090	0.070	0.068	0.073
GRU	0.100	0.083	0.082	0.080	0.063	0.088	0.063	0.060	0.068

Note. Bold: best performance result; PCC and solar-noon time used.

Table 8: MAE comparisons of SZF models and the proposed scheme using RNN.

Model	Location								
	Busan	Gangneung	Gwangju	Hadong	Incheon	Jeju	Jinju	Mokpo	Sejong
RNN_{Busan}	—	0.073	0.057	0.079	0.061	0.060	0.054	0.056	0.053
$RNN_{Gangneung}$	0.075	—	0.057	<u>0.046</u>	<u>0.045</u>	<u>0.060</u>	0.043	0.047	0.051
$RNN_{Gwangju}$	<u>0.061</u>	0.063	—	0.064	0.051	0.061	0.045	0.052	<u>0.049</u>
RNN_{Hadong}	0.088	<u>0.055</u>	0.064	—	0.054	0.065	0.049	0.054	0.060
$RNN_{Incheon}$	0.081	0.070	0.070	0.068	—	0.073	0.058	0.053	0.063
RNN_{Jeju}	0.071	0.063	0.056	0.066	0.051	—	0.053	0.050	0.054
RNN_{Jinju}	0.076	0.056	0.060	0.058	0.053	0.071	—	0.056	0.053
RNN_{Mokpo}	0.079	0.065	0.064	0.066	0.053	0.066	0.054	—	0.056
RNN_{Sejong}	0.063	0.057	<u>0.051</u>	0.058	0.045	0.060	<u>0.043</u>	<u>0.045</u>	—
Proposed Scheme	**0.058**	**0.041**	**0.040**	**0.045**	**0.029**	**0.038**	**0.027**	**0.030**	**0.034**

Note. Bold: best performance result; underline: second-best performance result.

Table 9: RMSE comparisons of SZF models and the proposed scheme based on RNN.

Model	Location								
	Busan	Gangneung	Gwangju	Hadong	Incheon	Jeju	Jinju	Mokpo	Sejong
RNN_{Busan}	—	0.114	0.089	0.121	0.088	0.091	0.081	0.082	0.078
$RNN_{Gangneung}$	0.116	—	0.092	**0.070**	0.069	0.095	0.073	0.069	0.078
$RNN_{Gwangju}$	**0.086**	0.096	—	0.095	0.077	<u>0.089</u>	<u>0.069</u>	0.075	<u>0.074</u>
RNN_{Hadong}	0.129	<u>0.086</u>	0.099	—	0.077	0.098	0.074	0.078	0.087
$RNN_{Incheon}$	0.112	0.102	0.103	0.094	—	0.105	0.082	0.073	0.087
RNN_{Jeju}	0.098	0.094	0.085	0.093	0.075	—	0.079	0.071	0.078
RNN_{Jinju}	0.105	0.088	0.090	0.082	0.074	0.099	—	0.077	0.075
RNN_{Mokpo}	0.112	0.098	0.097	0.095	0.075	0.105	0.082	—	0.081
RNN_{Sejong}	<u>0.091</u>	0.090	<u>0.083</u>	0.087	<u>0.067</u>	0.096	0.070	<u>0.064</u>	—
Proposed Scheme	0.097	**0.080**	**0.079**	<u>0.079</u>	**0.057**	**0.079**	**0.058**	**0.056**	**0.063**

Note. Bold: best performance result; underline: second-best performance result.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \widehat{Y}_i\right)^2}. \qquad (8)$$

In the first experiment, we determined three important types: DL model type, correlation coefficient type, and data time type. For this purpose, we considered all possible combinations of these three types. Table 4 shows the hyperparameters used in each DL model.

In particular, the window size for RNN-type DL models (RNN, LSTM, and GRU) in the table specifies how much historical data should be used for prediction. On the other

TABLE 10: MAE improvement of each deep learning model.

| Model type		Location								
		Busan	Gangneung	Gwangju	Hadong	Incheon	Jeju	Jinju	Mokpo	Sejong
SNN	Best SZF	0.069	0.055	0.059	0.050	0.049	0.063	0.046	0.047	0.055
	Proposed scheme	0.062	0.042	0.042	**0.044**	0.035	0.045	0.028	0.032	0.037
	Improvement (%)	9.807	22.571	29.130	11.841	30.041	28.033	38.527	32.272	33.305
DNN	Best SZF	0.061	0.051	0.053	0.045	0.043	0.058	0.041	0.041	0.049
	Proposed scheme	0.061	0.046	0.042	0.046	0.036	0.050	0.029	0.034	0.036
	Improvement	-0.128	9.615	20.700	-3.418	16.546	14.038	29.906	17.674	25.608
RNN	Best SZF	0.061	0.055	0.051	0.046	0.045	0.060	0.043	0.045	0.049
	Proposed scheme	**0.058**	**0.041**	**0.040**	0.045	**0.029**	**0.038**	**0.027**	**0.030**	**0.034**
	Improvement	5.092	26.484	21.413	1.763	34.629	36.926	37.472	32.913	30.741
LSTM	Best SZF	0.060	0.054	0.053	0.050	0.047	0.053	0.044	0.046	0.049
	Proposed scheme	0.062	0.050	0.045	0.049	0.040	0.048	0.036	0.041	0.042
	Improvement	-3.735	8.015	14.228	0.585	14.389	10.947	18.055	10.899	14.884
GRU	Best SZF	0.058	0.050	0.052	0.046	0.043	0.052	0.041	0.041	0.049
	Proposed scheme	0.059	0.048	0.043	0.044	0.035	0.048	0.031	0.037	0.038
	Improvement	-1.579	5.036	17.918	2.629	17.469	7.918	24.452	9.386	22.561

Note. Bold: best performance result.
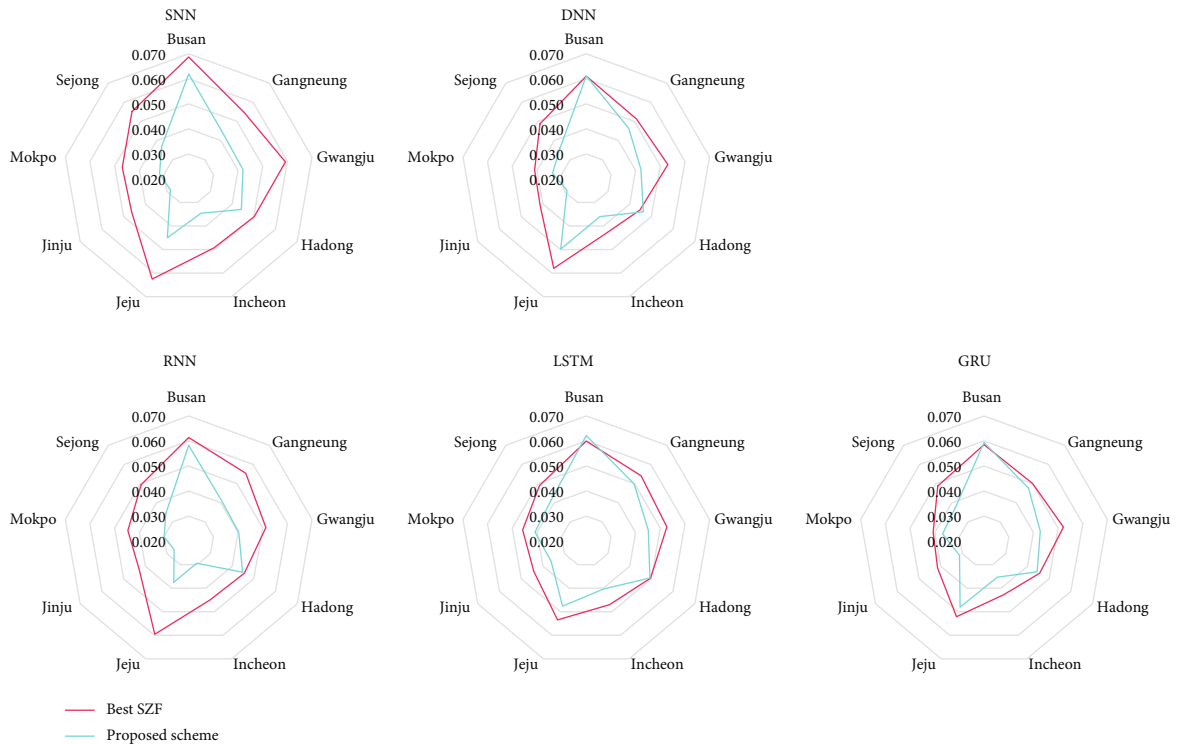
FIGURE 6: Radar graphs of MAE of locations according to deep learning model.

hand, SNN and DNN do not have this hyperparameter, so they cannot consider historical data in their predictions. Therefore, to ensure a fair comparison of prediction performance, the data used for SNN and DNN training were configured to include past data. After all, all models have 18 input variables but reflect all data from the past 24 points; that is, 432 input data are used in the model to generate one output. For instance, Table 5 presents the MAE comparison of all other combinations under the LSTM model.

In the table, the first column shows the type information of the proposed model. For instance, $Scheme_{LSTM,PCC,All}$ indicates that LSTM, PCC, and all time are used as the type
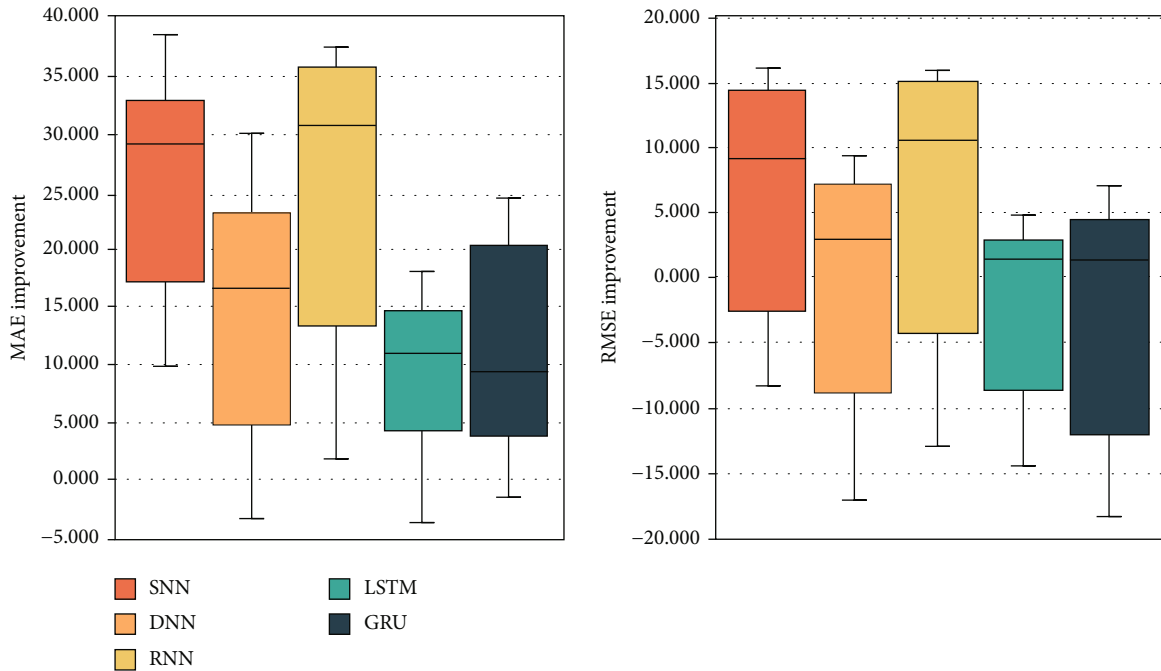
FIGURE 7: Box plot of MAE and RMSE improvement compared to the best SZF.

of DL model, correlation coefficient, and data time, respectively. In addition, values in bold indicate the best performance for each location. In the case of Busan, no bold values are listed because SZF performs better than the proposed scheme.

In the experiment, we analyzed all cases that showed the best results. Figure 5 shows a bubble chart of the number of best performing cases according to data time and correlation coefficient. From the figure, we can see that PCC is better than DCC in terms of correlation coefficient and noon time data is the best in terms of data time. Hence, we used PCC and solar-noon time to implement the proposed scheme using various DL models. Tables 6 and 7 present the MAE and RMSE values of five different DL models for each location, respectively. Here, the bold values indicate the best performance for each location.

The tables show that the proposed scheme achieves the best prediction performance when using RNN. On the other hand, DNN and GRU models showed the next best performance. To prove the effectiveness of the proposed scheme, we compared its performance with RNN-based SZFs trained using data from other areas. Tables 8 and 9 present the MAE and RMSE values, respectively. Values in bold indicate the best performance, and values with an underline indicate the second-best performance for each location.

The proposed scheme performed best in most cases. We calculated how much the performance of the proposed scheme improved compared to the best performing SZF. Compared to the best SZF, the proposed scheme improved MAE performance by an average of 25.3% and RMSE performance by an average of 4.7%. In the next experiment, to prove the robustness of the proposed scheme for predictive models, we compared the performance of all comparative DL models and the proposed scheme. Table 10 shows the

MAE of the best SZF for each model and the proposed scheme, and the improvement rate achieved by the proposed scheme. Figure 6 presents radar graphs for the MAE of the best SZF and the proposed scheme. In the figure, the smaller the radar size, the better the predictive performance. In summary, the proposed scheme performs better than the best SZF in almost all cases. Figure 7 presents the MAE and RMSE improvement compared to the best for each DL model type. We can see that the proposed scheme outperforms the best SZF in almost all cases. In only a few cases where the source and target data have very similar patterns, SZF gave slightly better forecasting performance than the proposed scheme. However, since we have to make predictions without knowing what pattern the target data will have, this is impractical for use in a real-world environment. Additionally, in terms of model construction time, the proposed scheme takes the time to construct a single SZF multiplied by the number of source data. However, if real-time updates to the model are not required, prediction accuracy is much more important than model construction time. Hence, the proposed scheme is more desirable for newly built PV power systems than SZF.

6. Conclusion

In this paper, we proposed a novel zero-shot PV power forecasting scheme based on DL model and correlation coefficient. We constructed an individual DL-based forecasting model for multiple locations and then calculated the final prediction value through the correlation coefficient-based weighted sum of the predicted values of these models. By using this approach, the proposed scheme can always guarantee good prediction performance even in newly built PV generation systems.

To evaluate the effectiveness and efficiency of the proposed scheme, we performed extensive experiments using historical data collected from nine source sites in Korea. We first compared the performance of the proposed schemes using different combinations of DL model, correlation coefficient, and data time option to find the best one. We considered five DL models (SNN, DNN, RNN, LSTM, and GRU) for forecasting model and two correlation coefficients (PCC and DCC) and three data time types (all time, day time, and solar-noon time) for weights. The experimental results revealed that the forecasting performance is best for the combination of RNN, PCC, and solar-noon data time data. The proposed scheme using this combination achieved an average MAE of 0.038 and an average RMSE of 0.072.

In addition, we compared the performance of the proposed scheme with SZF models trained using different datasets. We confirmed that the proposed scheme achieved the best predictive performance in almost all cases and the second-best performance with a very narrow margin only in a few cases. More specifically, it improved MAE performance by an average of 25.3% and RMSE performance by an average of 4.7% compared to the best performing SZF model. Although SZF showed the best performance in some cases, it is difficult to use in real-world environment because its performance varies greatly depending on the data used for training.

In the future, based on the forecasted value derived through the zero-shot PV power forecasting model and the electrical load data of the building or cluster where the PV system will be installed, we plan to research the optimal capacity of an energy storage system to be integrated with this grid. In addition, we intend to develop an energy operation scheduling algorithm that can guarantee the most economic benefits to users by analyzing how energy is most economically operated in a PV–energy storage system-integrated grid with optimal capacity.

Data Availability

The data that support the findings of this study are available from "https://www.data.go.kr/index.do" and "https://data.kma.go.kr/cmmn/main.do."

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgments

References

[1] H. Liu, J. D. Ampah, Y. Zhao et al., "A perspective on the overarching role of hydrogen, ammonia, and methanol carbon-neutral fuels towards net zero emission in the next three decades," *Energies*, vol. 16, no. 1, p. 280, 2023.

[2] A. Hassan, S. Z. Ilyas, A. Jalil, and Z. Ullah, "Monetization of the environmental damage caused by fossil fuels," *Environmental Science and Pollution Research*, vol. 28, no. 17, pp. 21204–21211, 2021.

[3] L. Li, J. Lin, N. Wu et al., "Review and outlook on the international renewable energy development," *Energy and Built Environment*, vol. 3, no. 2, pp. 139–157, 2022.

[4] O. Ellabban, H. Abu-Rub, and F. Blaabjerg, "Renewable energy resources: current status, future prospects and their enabling technology," *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 748–764, 2014.

[5] A. Qazi, F. Hussain, N. A. Rahim et al., "Towards sustainable energy: a systematic review of renewable energy sources, technologies, and public opinions," *IEEE Access*, vol. 7, pp. 63837–63851, 2019.

[6] A. H. Elsheikh, S. W. Sharshir, M. Abd Elaziz, A. E. Kabeel, W. Guilan, and Z. Haiou, "Modeling of solar energy systems using artificial neural network: a comprehensive review," *Solar Energy*, vol. 180, pp. 622–639, 2019.

[7] T. Qiu, L. Wang, Y. Lu et al., "Potential assessment of photovoltaic power generation in China," *Renewable and Sustainable Energy Reviews*, vol. 154, article 111900, 2022.

[8] A. Aslam, N. Ahmed, S. A. Qureshi, M. Assadi, and N. Ahmed, "Advances in solar PV systems; a comprehensive review of PV performance, influencing factors, and mitigation techniques," *Energies*, vol. 15, no. 20, p. 7595, 2022.

[9] S. Ghazi and K. Ip, "The effect of weather conditions on the efficiency of PV panels in the southeast of UK," *Renewable Energy*, vol. 69, pp. 50–59, 2014.

[10] P. C. Babu, D. G. Kumar, N. V. Sireesha et al., "Modeling and performance analysis of a grid-connected photovoltaic system with advanced controller considering varying environmental conditions," *International Journal of Energy Research*, vol. 2023, Article ID 1631605, 23 pages, 2023.

[11] Y. Li, Y. Su, and L. Shu, "An ARMAX model for forecasting the power output of a grid connected photovoltaic system," *Renewable Energy*, vol. 66, pp. 78–89, 2014.

[12] Y. Onal, "Gaussian kernel based SVR model for short-term photovoltaic MPP power prediction," *Computer Systems Science and Engineering*, vol. 41, no. 1, pp. 141–156, 2022.

[13] S. M. Jung, S. Park, S. W. Jung, and E. Hwang, "Monthly electric load forecasting using transfer learning for smart cities," *Sustainability*, vol. 12, no. 16, p. 6364, 2020.

[14] J. Moon, J. Kim, P. Kang, and E. Hwang, "Solving the cold-start problem in short-term load forecasting using tree-based methods," *Energies*, vol. 13, no. 4, p. 886, 2020.

[15] J. Bottieau, Z. De Grève, T. Piraux, A. Dubois, F. Vallée, and J. F. Toubeau, "A cross-learning approach for cold-start forecasting of residential photovoltaic generation," *Electric Power Systems Research*, vol. 212, article 108415, 2022.

[16] F. Pourpanah, M. Abdar, Y. Luo et al., "A review of generalized zero-shot learning methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4051–4070, 2022.

[17] W. Wang, V. W. Zheng, H. Yu, and C. Miao, "A survey of zero-shot learning," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–37, 2019.

[18] M. Pan, C. Li, R. Gao et al., "Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization," *Journal of Cleaner Production*, vol. 277, article 123948, 2020.

[19] M. Meng and C. Song, "Daily photovoltaic power generation forecasting model based on random forest algorithm for north China in winter," *Sustainability*, vol. 12, no. 6, p. 2247, 2020.

[20] G. Li, H. Wang, S. Zhang, J. Xin, and H. Liu, "Recurrent neural networks based photovoltaic power forecasting approach," *Energies*, vol. 12, no. 13, p. 2538, 2019.

[21] M. S. Hossain and H. Mahmood, "Short-term photovoltaic power forecasting using an LSTM neural network and synthetic weather forecast," *IEEE Access*, vol. 8, pp. 172524–172533, 2020.

[22] L. Liu, M. Zhan, and Y. Bai, "A recursive ensemble model for forecasting the power output of photovoltaic systems," *Solar Energy*, vol. 189, pp. 291–298, 2019.

[23] G. Li, S. Xie, B. Wang, J. Xin, Y. Li, and S. Du, "Photovoltaic power forecasting with a hybrid deep learning approach," *IEEE Access*, vol. 8, pp. 175871–175880, 2020.

[24] L. Ricci and D. Papurello, "A prediction model for energy production in a solar concentrator using artificial neural networks," *International Journal of Energy Research*, vol. 2023, Article ID 9196506, 20 pages, 2023.

[25] J. Moon, S. Park, S. Rho, and E. Hwang, "A comparative analysis of artificial neural network architectures for building energy consumption forecasting," *International Journal of Distributed Sensor Networks*, vol. 15, no. 9, 2019.

[26] H. Alghamdi, C. Maduabuchi, K. Okoli et al., "Smart optimization of semiconductors in photovoltaic-thermoelectric systems using recurrent neural networks," *International Journal of Energy Research*, vol. 2023, Article ID 6927245, 18 pages, 2023.

[27] M. J. J. Gul, M. H. Firmansyah, S. Rho, and A. Paul, "BI-LSTM-LSTM based time series electricity consumption forecast for South Korea," in *Advances in Artificial Intelligence and Applied Cognitive Computing. Transactions on Computational Science and Computational Intelligence*, H. R. Arabnia, K. Ferens, D. Fuente, E. B. Kozerenko, J. A. Olivas Varela, and F. G. Tinetti, Eds., pp. 897–902, Springer, Cham, 2021.

[28] B. Liu, J. Xu, and W. Xia, "State-of-health estimation for lithium-ion battery based on an attention-based CNN-GRU model with reconstructed feature series," *International Journal of Energy Research*, vol. 2023, Article ID 8569161, 13 pages, 2023.

[29] J. Park, S. H. Hong, S. H. Yeon, B. M. Seo, and K. H. Lee, "Predictive model for solar insolation using the deep learning technique," *International Journal of Energy Research*, vol. 2023, Article ID 3525651, 17 pages, 2023.

[30] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," *The Annals of Statistics*, vol. 35, no. 6, pp. 2769–2794, 2007.

[31] S. Park, J. Moon, S. Jung, S. Rho, S. W. Baik, and E. Hwang, "A two-stage industrial load forecasting scheme for day-ahead combined cooling, heating and power scheduling," *Energies*, vol. 13, no. 2, p. 443, 2020.