





A Novel Forecasting Model for Solar Power Generation by a Deep Learning Framework With Data Preprocessing and Postprocessing

Quoc-Thang Phan , Yuan-Kang Wu , *Member, IEEE*, Quoc-Dung Phan , and Hsin-Yen Lo 

Abstract—Photovoltaic power has become one of the most popular forms of energy owing to the growing consideration of environmental factors; however, solar power generation has brought many challenges for power system operations. With regard to optimizing safety and reducing the costs of power system operations, an accurate and reliable solar power forecasting model would be a significant step forward. This study proposes a deep learning method to improve the performance of short-term one-hour-ahead solar power forecasting, which includes data preprocessing, feature engineering, kernel principal component analysis, a gated recurrent unit network training model based on time-of-day classification, and postprocessing with error correction. Both historical solar power, solar irradiance, and numerical weather prediction (NWP) data, such as temperature, irradiance, rainfall, wind speed, air pressure, and humidity, were used as the input dataset in this work. As a case study, the measured power from ten PV sites in Taiwan were collected and predicted with a one-hour resolution. The normalized root mean squared error and normalized mean absolute percent error were chosen to evaluate the performance of the forecasting models. Compared with other benchmark models, including ANN, LSTM, XGBoost, and single GRU, the experimental results showed the proposed model's superior performance. Furthermore, the importance of data preprocessing and postprocessing based on error correction was demonstrated.

Index Terms—Deep learning, feature engineering, forecasting, gated recurrent unit (GRU), numerical weather prediction (NWP), post processing, preprocessing, solar power.

I. INTRODUCTION

RENEWABLE energy resources, especially wind and solar power as the most common forms, have been used

Manuscript received 12 April 2022; revised 23 July 2022 and 25 September 2022; accepted 4 October 2022. Date of publication 10 October 2022; date of current version 19 January 2023. Paper 2022-ESC-0493.R2, presented at the 2022 IEEE/IAS 58th Industrial and Commercial Power Systems Technical Conference, Las Vegas, NV, USA, May 02–05, and approved for publication in the IEEE TRANSACTIONS ON INDUSTRY APPLICATIONS by the Energy Systems Committee of the IEEE Industry Applications Society. This work was supported in part by the Ministry of Science and Technology (MOST) of Taiwan under Grant MOST 111-3116-F-008 -003. (*Corresponding author: Yuan-Kang Wu.*)

Quoc-Thang Phan, Yuan-Kang Wu, and Hsin-Yen Lo are with the National Chung-Cheng University, Chia-Yi 62102, Taiwan (e-mail: thangphan-quoc2407@gmail.com; allenwu@ccu.edu.tw; haruyan911@gmail.com).

Quoc-Dung Phan is with the Ho Chi Minh City University of Technology, Ho Chi Minh City 70000, Vietnam, and also with the Vietnam National University Ho Chi Minh City, Ho Chi Minh City 70000, Vietnam (e-mail: pqdung@hcmut.edu.vn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TIA.2022.3212999>.

Digital Object Identifier 10.1109/TIA.2022.3212999

worldwide for electricity generation to overcome the issues associated with greenhouse emissions and climate change [1]. However, renewable energy integration has created many challenges for the operation of power systems because some of its inherent characteristics [2]. Therefore, it is necessary to forecast renewable energy output since this would provide valuable information about the expected changes in the energy to be generated in the near future [3]. For instance, the stability and quality of power grids could be improved if the accuracy of solar power output forecasting increased. Photovoltaic solar power is reaching a high level of penetration in smart grids [4]; however, PV solar power forecasting can be very challenging because of the variety of weather conditions and several other uncertainties such as temperature, humidity, wind speed, cloud cover, rainfall, and so on.

In practice, numerous solar power forecasting methods have been proposed over the past few decades that can provide very short-term to long-term predictions. The forecasting techniques are divided into different categories: physical methods, statistical methods, artificial intelligence-based methods, and ensemble methods [5]. AI methods in particular could become useful tools for solar power forecasting since they can solve the nonlinear problem between input and output data. Of the many methods for short-term forecasts, the main ones include linear regression, support vector machine (SVM), autoregressive moving average (ARMA), time series modeling, back-propagation neural network (BPNN), among others. It should be noted that linear regression requires a large data set and pathological data can affect the fitting results [6]. The ARIMA models only use historical power outputs, which can result in large forecasting errors [7]. The SVM method cannot deal with large amounts of data in terms of training speed and forecasting accuracy [8]. In addition, it is difficult to select the kernel functions in that process since it is more suited for classification [9]. The algorithm of a traditional BPNN needs to be improved to obtain a better convergence rate [10], [11], [12]. Furthermore, the Markov chain is also reliant on a big dataset, although it still has good performance even if there are missing data [13]. A machine learning method, k-nearest neighbor (kNN), has been chosen for solar power forecasting in the past [14]. Other machine learning methods such as boosting, bagging, and regression trees have proven to be accurate and effective.

Deep learning has been an active field of study owing to its applications in renewable power forecasting, particularly for

wind power forecasts; however, it is observed that numerous ensemble models used in previous publications lack deep learning techniques, such as long short-term memory (LSTM) or gated recurrent unit (GRU) networks [15], [16], [17]. Thus, this paper overcomes this problem by integrating GRU into the proposed model, which is then applied to three time-of-day clusters of the dataset to precisely capture the relationship between solar power generation and selected weather parameters.

In recent years, some research works have promoted deep learning-based models on account of their regression, data mining, and feature extraction capabilities [18]. The common deep learning models for forecasting solar power generation include the deep neural network (DNN), deep belief network (DBN), Boltzmann machines, and recurrent neural network (RNN). The RNN has become the preferred choice for forecasts in smart grids [19]. The proposed RNN model in [20] only used weather data as inputs to predict solar output. LSTM and GRU are special types of RNN, and some research works applied LSTM to improve forecasting performance compared to standard RNN [21], [22]. The goal of both GRU and LSTM is to solve the vanishing and exploding gradient problems that exist for RNN [23], and Reference [24] proposed a deep LSTM-RNN model to accurately forecast solar power output. Although LSTM provides a high degree of prediction accuracy, its training time is long. Thus, GRU can be applied to reduce the parameters and training time. When dealing with a large dataset, GRU can be used to save more time in the experimental process. In [25], the input features were extracted using the Pearson correlation coefficient, then K-mean was used to cluster the data into groups, and finally each group was trained with each GRU model. The forecasting results showed that GRU outperformed SVM, ARIMA, and BPNN. GRUs have been seldom used in the field of forecasts but their ability to learn the trend of time series is vastly superior [26]. In [27], a forecasting model based on a modified multiple GRU was proposed to improve the training time. Thus, it appears that this kind of ensemble model can surpass any pure type of model by overcoming the drawbacks of each model type [28].

In the process of deterministic solar power forecasting, data preprocessing is considered one of the most important steps since the input data combines historical measurement solar power, solar radiation, and data from numerical weather prediction (NWP). Data preprocessing includes cleaning, normalization, transformation, dimensional reduction, feature extraction, and feature selection [29]. Research works [30], [31] used K-means to classify the similar weather types of a day from training data as sunny, cloudy, or rainy. There are numerous factors that affect PV power. If all of them are included as input indicators for a forecasting model, there will be a large amount of redundant data. Thus, feature selection is essential [32]. PCA, a multivariate statistical technique that integrates several variables into a smaller number of variables, is more effective when dealing with individual indicators that have a strong linear connection. However, PCA is a linear approach that cannot obtain higher-order data features and ignores the nonlinear data information when reducing dimensionality. Therefore, KPCA is also utilized in the proposed model to reduce the dimensionality

of the input dataset while keeping the nonlinear information [33]. Another study [34] demonstrated the effectiveness of including a feature selection method for solar power forecasting. However, the forecasting accuracy of these models can still be enhanced through postprocessing.

Postprocessing refers to any method that refines the forecast results. In PV forecasts, there are two well-accepted postprocessing approaches: model output statistics (MOS) [35] and Kalman filtering [36]. However, the use of Kalman filtering is not a “genuine” postprocessing technique because it modifies the prediction horizon. Thus, MOS is the only approved method in the literature. This paper applied a new postprocessing technique that adjusts deterministic predictions by removing outliers of daily fitting curves at one standard deviation, which covers 68% of the data.

Thus, a combination of preprocessing and postprocessing with deep learning models can significantly improve short-term PV power forecasting.

The main contributions of this paper are summarized as follow:

- 1) A new correction method was proposed for data postprocessing to improve forecasting accuracy. This was accomplished by adding a daily error-correction curve constructed using a fourth-degree polynomial in which an average curve with outliers above one standard deviation removed was obtained. Such a process has rarely been considered in previous publications. Reference [37] applied postprocessing to an NWP model that forecasted solar irradiance; however, the postprocessing technique was different. Reference [38] discussed postprocessing technologies for solar irradiance forecasting. Reference [39] summarized postprocessing technologies for probabilistic forecasts. The above-mentioned methods leveraged advances in statistical learning to handle NWP postprocessing, and they applied regression methods for solar irradiance, NWP models, and probabilistic forecasts. The main purpose of this paper, however, is to improve deterministic PV power forecasting based on a new error correction method that removes outliers beyond one standard deviation, and not by using a regression model.
- 2) Proposed a complete deep learning framework by combining the preprocessing technique based on Pearson correlation analysis, feature selection using XGBoost, KPCA dimension reduction, and postprocessing based on a daily numerical fitting curve. Such a complete model has rarely been considered in previous publications as most of them focused only on data preprocessing, and only a handful of works established a complete model that combines data preprocessing, feature selection, dimension reduction, and postprocessing simultaneously. The forecasting results demonstrate that the proposed “complete process” forecasting model achieves superior performance.
- 3) Proposed three GRU-based forecasting engines for three time periods throughout each day, including morning, noon, and afternoon.

- 4) Combined historical PV power and radiation with NWP data from the Central Weather Bureau (CWB) of Taiwan.
- 5) Analyzed the contributions of each step in the proposed forecasting model. That is, the performance of the prediction process with both pre- and postprocessing, with only postprocessing, with only preprocessing, and without pre- or postprocessing was analyzed. The results revealed that the combination of pre- and postprocessing methods (our proposed model) outperforms other benchmark models.
- 6) All results were obtained and validated using real data recorded over ten different PV sites.

The rest of the paper is organized as follows: Section II provides the methodologies of the proposed model. Section III presents the detailed framework for solar power forecasting including preprocessing, the forecast engine, and postprocessing. Section IV provides the case study and analysis of the forecasting results. Finally, the conclusions are drawn in Section V.

II. METHODOLOGY

A. Kernel Principal Component Analysis (KPCA)

In KPCA, the computations are performed in a feature space that is nonlinearly related to the input space. Owing to this nonlinear relationship, the KPCA is considered to be a nonlinear algorithm. However, unlike other nonlinear PCAs, KPCA depends on linear algebra by mapping original inputs into a high-dimensional feature space via a kernel map, which makes the data structure more linear. This paper utilized KPCA to extract key factors, which can effectively deal with the nonlinear relationship between the different variables and achieve more reasonable results than other PCAs.

Set group of random vectors $X = \{x_1, x_2, \dots, x_n\}^T$, where $x_k \in R^N$ ($k = 1, 2, \dots, N$); N is the number of input samples. If maps N points to an N -dimension space by:

$$\Phi(x_i) \text{ where } \Phi : \mathbb{R}^d \rightarrow \mathbb{R}^N \quad (1)$$

The covariance matrix is estimated as:

$$C = \frac{1}{N} \sum_{i=1}^N \Phi(x_i) \Phi^T(x_i) \quad (2)$$

Assuming that the feature vectors are centered, the eigenvalues can be obtained as:

$$\lambda_i \phi_i = C \phi_i \quad (3)$$

where λ_i is non-zero eigenvalues of C , ϕ_i is corresponding eigenvectors. Then, the eigenvector is calculated by:

$$v = \sum_{j=1}^N \alpha_j \Phi(x_j) \quad (4)$$

where α_j is coefficient. From (2), (3), (4) and matrix $Q(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$. The following equation can be obtained:

$$\sum_{i=1}^N \sum_{j=1}^N \alpha_j \Phi(x_i) Q(x_i, x_j) = N \lambda \sum_{j=1}^N \alpha_j \Phi(x_j) \quad (5)$$

Finally, the principal component is calculated by:

$$y_i = \sum_{j=1}^N \alpha_j \Phi(x_i, x_j), \quad i = 1, 2, \dots, N \quad (6)$$

Gaussian Radial Basis Function was employed in this work. As a result, the kernel principal component can be calculated by referring to the extraction technique of traditional PCAs.

B. XGBoost

In this study, XGBoost was used for selecting the most important features. XGBoost selects a decision tree as its base classifier and adds a new decision tree at each iteration to improve the objective function. The final prediction values are equal to the summation of all base learners. The XGBoost algorithm applies the greedy algorithm to build the decision tree. Through the continuous process of building its decision trees, a complete XGBoost model can be established. The final prediction of XGBoost is calculated as follows:

$$Y_i = \sum_{m=1}^M f_m(x_i), \quad f_m \in F \quad (7)$$

Where f is a decision tree, and F represents the function of all decision trees. Y_i is prediction for the i -th instance at the m -th boost. The idea of XGBoost is to minimize the loss function L_m :

$$L_m = \sum_{i=1}^n l(y_i, Y_i) \quad (8)$$

where $l(y_i, Y_i)$ measures the difference between forecasting and real value. Adding a regularization term as (9) to loss function, the objective function can be archived in (10):

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (9)$$

The regularization term defines the model complexity. To reduce the over-fitting and enhance model performance, sub-sample, max leaves, max depth are added. Therefore, as the t^{th} step learning happens, the predicted value of x_i is $Y_i^t = Y_i^{t-1} + f_t(x_i)$. Therefore, the object function is expressed by:

$$\begin{aligned} Obj(\theta)^t &= \sum_{i=1}^n l(y_i, Y_i^t) + \sum_{m=1}^t \Omega(f_m) \\ &= \sum_{i=1}^n l(y_i, Y_i^{t-1} + f_t(x_i)) + \sum_{m=1}^t \Omega(f_m) \end{aligned} \quad (10)$$

C. Gated Recurrent Unit Network (GRU)

GRU is a special type of RNN, introduced in 2014 by Cho et al. [40]. Different to traditional ANNs or convolutional neural network (CNN), the inputs between the hidden layers of an RNN have the output of the upper layer and output of the previous cell on the same layer. Many studies have pointed out that RNN has several weak points when learning long-term dependencies [41], and GRU was created as the solution to short-term memory and is an attempt to simplify LSTM. As shown in Fig. 1, GRU only has

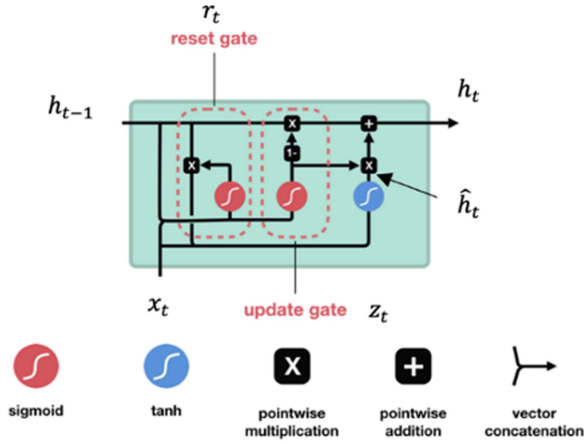


Fig. 1. Structure of one cell in a GRU network.

two gates, whereas LSTM has three gates that are controlled by inputs and outputs. There are two input features at each time step: the previous output vector h_{t-1} and input vector x_t . The output of each gate is calculated based on nonlinear transformation and logical operation. A standard GRU network structure is defined by the following equations:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \quad (11)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \quad (12)$$

$$\hat{h}_t = \sigma_h(W_h x_t + U_h (h_{t-1} \odot r_t) + b_h) \quad (13)$$

$$h_t = (1 - z_t) \odot \hat{h}_t + z_t \odot h_{t-1} \quad (14)$$

where z_t is update gate vector, r_t is reset gate vector, h_t denotes the state vector for current time step t , W , U are parameter matrices and vector. \odot is element-wise multiplication. σ is sigmoid activation function and σ_h is a hyperbolic tangent.

After building the network structure, an appropriate training method for GRU should be considered. For instance, real-time recurrent learning or back propagation through time. Reference [42] proved that back-propagation has a high computational efficiency and short training time compared to real-time recurrent learning. In addition, Adam optimizers have been used to train GRU systems, which demonstrated better performance than those trained by other algorithms such as Adagrad, stochastic gradient descent, and RMSProp [43].

D. Measurement of Forecasting Errors

For short-term solar power forecasts, several statistical indices have been proposed to evaluate the forecasting performance. The evaluation indices show how well the data fit the models and quantify the forecasting performance. In this paper, Normalized Root Mean Squared Error (NRMSE), and Normalized Mean Absolute Error (NMAPE) were selected to calculate the forecasting errors, which is represented by:

$$NMAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{capacity} * 100 \quad (15)$$

 TABLE I
DIFFERENT FEATURES MEANING

Feature	Meaning
power	Measured solar power (kW)
rad	Measured solar irradiance (Wm2)
NWP_rad	NWP solar irradiance (Wm2)
NWP_rain	NWP rainfall precipitation (mm)
NWP_press	NWP air pressure (Pa)
NWP_WS	NWP wind speed (m/s)
NWP_temp	NWP temperature at 2 meters (K)
NWP_hum	NWP humidity

$$NRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}}{capacity} * 100 \quad (16)$$

where n is the amount of data in the testing set, y_i is the real PV power generation, \hat{y}_i is the forecasted PV power generation, and capacity is the installed capacity of a PV site.

III. PROPOSED MODEL FRAMEWORK FOR SOLAR PV POWER FORECASTING

In this section, the proposed framework of the solar power forecasting model based on the abovementioned theories, called KPCA-XGB-GRU, is developed. Fig. 2 illustrates the flowchart of the developed method with detailed steps. There are three important steps in the proposed model: preprocessing, building the forecast engine, and postprocessing. First, the model uses two types of input data- historical measured solar power and solar irradiance, and NWP data including the meteorological values that are summarized in Table I. Min-max normalization was used to scale the raw dataset to the range of [0,1]. Next, feature engineering was applied to transform the reduced data into training data for the model with the use of the Pearson correlation and XGBoost. Then, feature dimensionality reduction on the impact factor was performed using KPCA. Second, the preprocessed data were separated into a training set and a testing set, and three GRU networks were trained on each group of data based on time-of-the-day: morning set, noon set, and afternoon set. The solar power output was calculated by averaging the three trained GRU networks. Finally, an error correction method was used to improve forecasting accuracy.

A. Data Preprocessing and Feature Engineering

Feature selection relies on a correlation analysis to determine the best features, and the selected features (independent variables) have the most statistical influence on determining the target variable (dependent variable). A feature in a dataset is a column of data. When working with any dataset, the user must understand which feature is going to have a statistically significant effect on the output variable. Therefore, the authors combined the two methods.

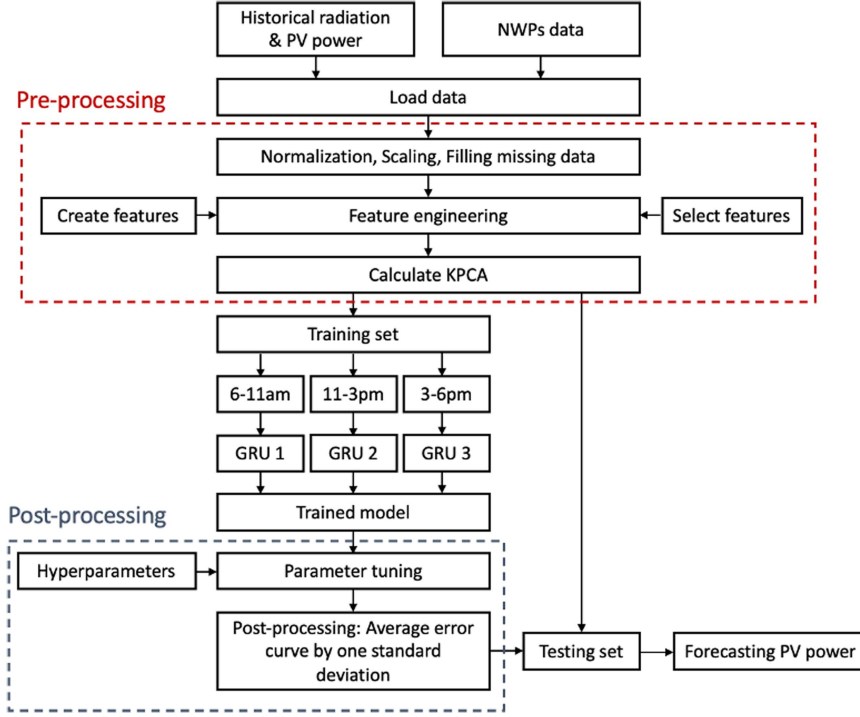


Fig. 2. The proposed framework of KPCA-XGB-GRU model with Corrected Error method.

Using the Pearson correlation, the linear relationship between the inputs and target variables was investigated. This correlation is only suitable for continuous inputs and target variables such as time-series data. If two variables have a low correlation coefficient, it means that there exists a weak linear relationship between them; however, there may exist a nonlinear relationship between them, i.e., they could have a quadratic or cubic relationship. This nonlinear relationship can achieve a high score when calculated using a tree-based model such as XGBoost. Thus, this study employed the Pearson correlation coefficient to remove similar and redundant features and utilized XGBoost to identify and sort important features.

There are several factors that could directly affect solar power generation, such as wind speeds, air pressure, cloud cover, and humidity. That is, it is necessary to improve forecasting accuracy by considering various external features. In this study, real data from nine solar sites in Taiwan and NWP data from the CWB were collected as raw input data. For data preprocessing, the first step is data normalization since different variables have different units. All data values were transformed into the range of [0,1].

$$D_{i,n} = \frac{D_i - \min(D_{train})}{\max(D_{train}) - \min(D_{train})} \quad (17)$$

where D_i is the value of data at hour i , $\min(D_{train})$ is the minimum values of training data, $\max(D_{train})$ is maximum values of training data, and $D_{i,n}$ is the normalized value.

By contrast, if included features did not have a relationship with the level of solar power generation, the model's complexity would increase, affecting the forecasting stability. Hence, the process of extracting features is important. Table I describes the different input features. Feature engineering is the process

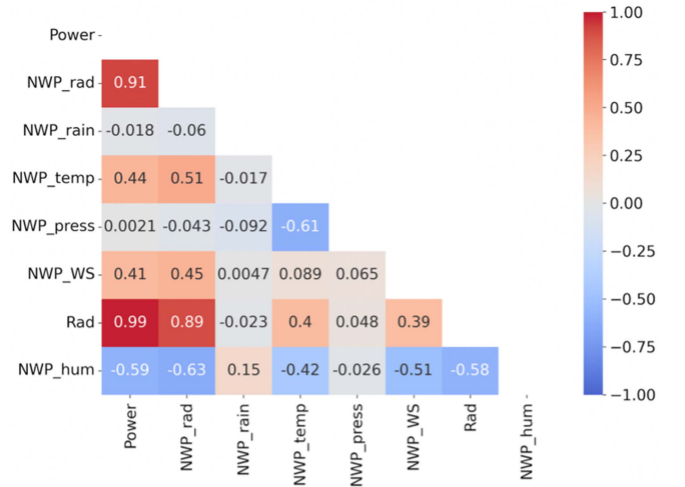


Fig. 3. Heatmap of Pearson correlation matrix.

of transforming input data into better quality training data by extracting more features.

A Pearson correlation coefficient matrix was established to reflect the linear correlation between two features, represented by a heatmap as shown in Fig. 3. The mathematical formular of Pearson correlation is:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (18)$$

where r_{xy} is the Pearson coefficient between x and y . \bar{x} is the mean of x , and \bar{y} is the mean of y . The range of r_{xy} is from -1 to 1.

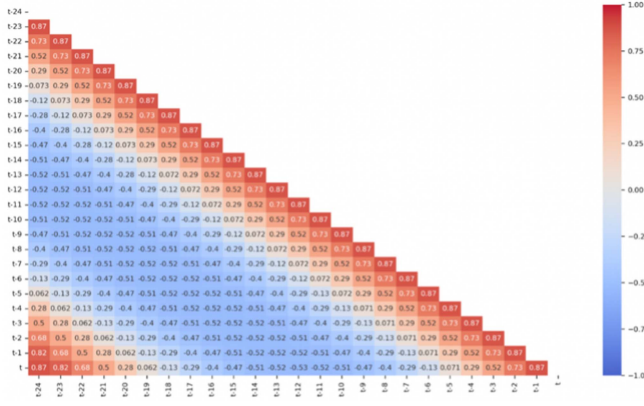


Fig. 4. Heatmap of Pearson correlation matrix of historical solar power and future solar power.

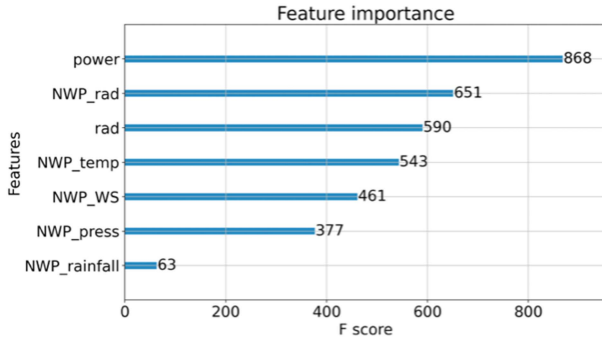


Fig. 5. Feature importance based on F-score.

The correlation coefficient is used to measure the linear relationship between two features, and it ranges between -1 and 1 . If one feature had a weak linear correlation with measured solar power or irradiance, it had to be removed. Taking the eight variables as the original features, the Pearson correlation was utilized to calculate the correlation between “power” and the other features. Based on the statistics, the “NWP_rain” and “power” correlation and the “NWP_press” and “power” correlation were the weakest. Therefore, both NWP_rain and NWP_press were removed as input variables; however, the other input variables were retained because they had a strong positive or negative correlation with power.

To analyze the relationship of the historical and future solar power time series, the solar power over the past 24 timesteps was set as $(t-1)$, $(t-2)$, ..., $(t-24)$. Based on the statistical analysis in Fig. 4, it can be observed that the solar power of time step “ t ” is mainly related to the historical power at time steps $t-1$ to $t-3$ and $t-22$ to $t-24$, because they have a strong positive correlation with “ t ”. Then, the corresponding historical powers at these lead times were selected as inputs for our forecasting model. The above selection is reasonable because previous publications [44] also used the Pearson correlation coefficient to find days that had the similar characteristics to the day of interest, indicating that historical PV power has a strong correlation with future PV power. Therefore, both the historical solar power with the

Contribution of variance of KPCA (%)

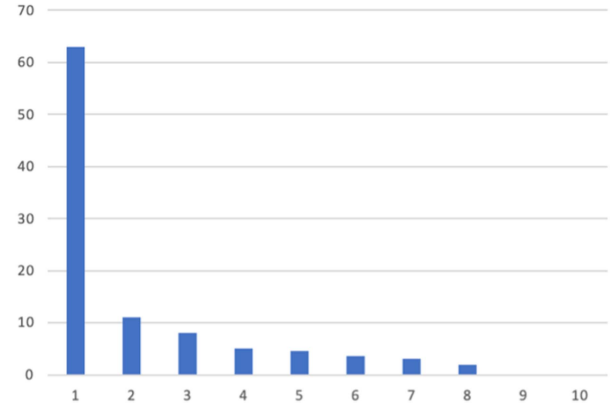


Fig. 6. Results of KPCA calculations.

strongest correlation and several other features obtained from Fig. 3 should be selected to improve the forecasting accuracy.

Feature importance (calculated by F-score) in an XGBoost-based model is more likely to identify which features are most influential. In a XGBoost-based model, after the boosted trees are constructed it is relatively straightforward to retrieve the scores for each attribute. Therefore, the authors used the feature importance technique to overcome the disadvantage of the Pearson correlation. Fig. 5 show the different features in order of their contributions from high to low. “NWP_rainfall” can be filtered out since its F score is considerably lower than that of the other features, indicating that it is relatively unimportant in the forecasting process. The most important features identified for the next steps were “NWP_rad”, “power”, “NWP_temp”, “NWP_WS”, “rad”, and “NWP_press”.

Then, KPCA was used to obtain the final set of principal components (PCs) for the training set. As shown in Fig. 6, the contribution rate of the first PC of KPCA reached up to 62.67%, implying that 62.67% of the original data was covered. After this process, the final input indicators for the forecasting model were formed.

B. Forecast Process Using a GRU Model Based on Time-of-Day Clustering

In this paper, the preprocessed data were separated into three clusters based on time-of-day: 6 am–10 am, 11 am–3 pm, and 3 pm–6 pm. A model based on a GRU was built for each cluster. The process for forecasting solar power based on a GRU is shown in Fig. 7 [45]. The main steps are as follows:

- Step 1:* Use historical measured solar power time series $(t-1)$, $(t-2)$, $(t-3)$, ..., $(t-24)$ to forecast future solar power (t) . The matrix X contains the historical solar power data and the most important features obtained from the preprocessing process.
- Step 2:* Every row of matrix X is scaled and inserted into to the GRU block in the GRU layer. The two hidden layers of the GRU have 100 nodes for all networks. The batch size is 16 and there are 100 epochs for each network.

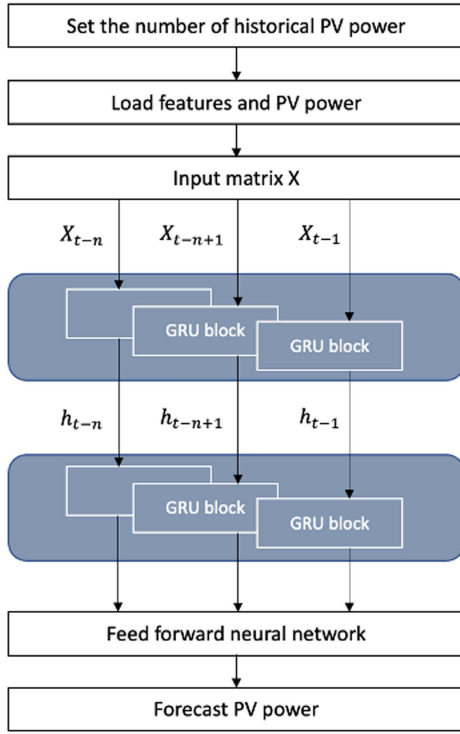


Fig. 7. The process forecasting solar power based on GRU.

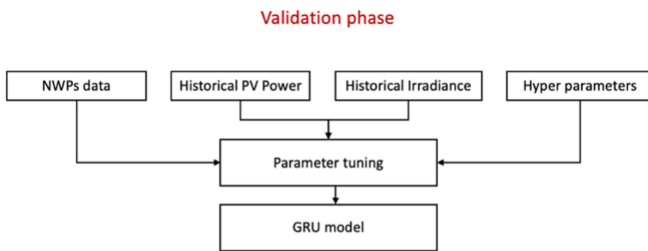


Fig. 8. Flowchart of the validation phase including hyper parameter tuning.

TABLE II
FINAL HYPERPARAMETER SETTINGS FOR GRU NETWORK

Parameters	Settings
Hidden layer nodes	64
Learning rate	0.002
Batch size	32
Epoch	50
Dropout	0.2

Step 3: The top GRU layer output is fed to a feed forward network that fits the output of GRU layer to solar power. The output size has one node for solar power forecasting.

Step 4: The hyperparameters to be optimized in the GRU are the number of hidden layer nodes and the learning rate, and the tuning process is shown in Fig. 8. The results after the tuning process are summarized in Table II.

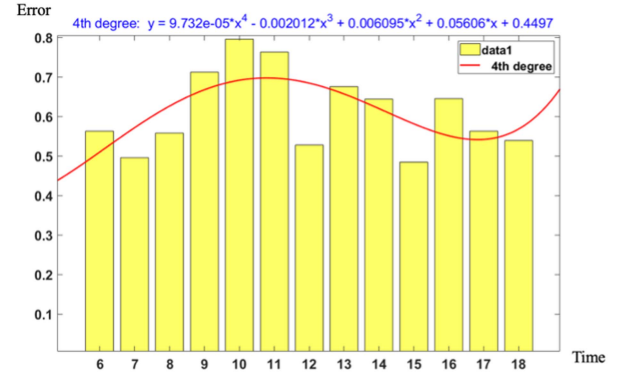


Fig. 9. Histogram and basic fitting curve of 6am-6pm (one day data).

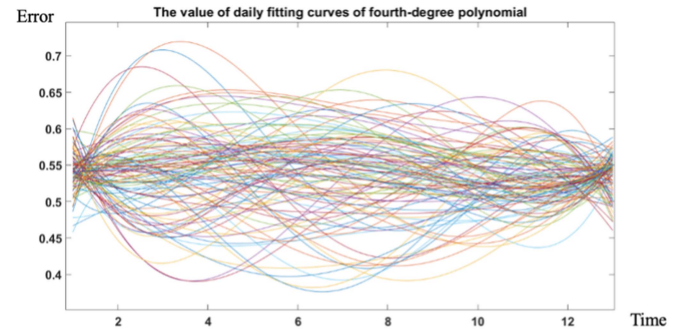


Fig. 10. The value of daily fitting curves of fourth-degree polynomial for forecasting error (4 months data).

C. Post Processing: Corrected Error Method

To enhance the accuracy of the proposed forecasting model, this paper proposed a postprocessing called error correction. The forecasting results that were obtained from the proposed model at this stage can be further analyzed to improve the forecasting performance. To obtain the maximum and minimum forecasting errors, the error between the actual values and the forecasted values is calculated as follows:

$$newerror = \frac{error - min_error}{max_error - min_error} \quad (19)$$

where min_error and max_error are the minimum and maximum of errors, respectively. Fig. 2 shows the proposed framework where the data in the training set were used to correct the error. More specifically, four months of training data (July to October) were used to obtain the fitting error curve, and then the outliers (beyond one standard deviation) were removed. Finally, an average error curve was obtained and added to it as a bias to correct our forecasting results. The postprocessing process significantly enhanced the forecasting results.

The following steps describe the process of correcting the forecasting error during the training process, and the corresponding results are shown in Figs. 9–12.

Step 1: Use training data to predict and obtain the error curve.

Based on our analyses, the error curve can be represented by a fourth-degree polynomial formula. Notably, the forecasting

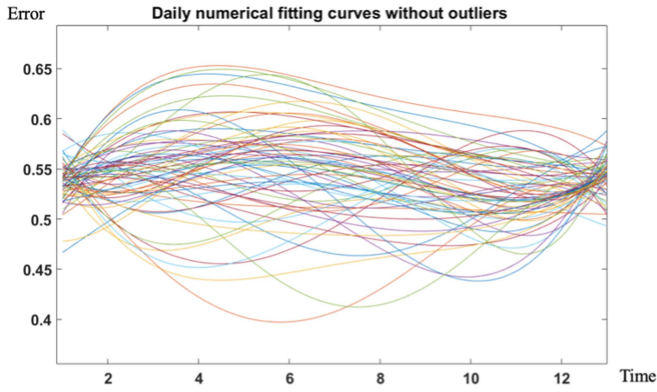


Fig. 11. Daily numerical fitting curves without outliers (4 months data).

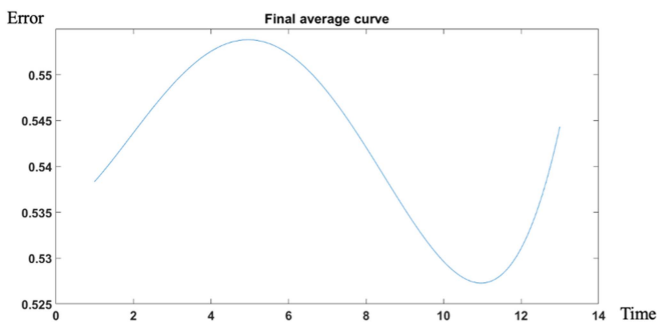


Fig. 12. Final average curve.

error at each curve is calculated by a p.u. value within the range [0,1]

Step 2: Remove the outliers (beyond one standard deviation) of these forecasting errors.

Step 3: Obtain an average curve of these forecasting errors.

Step 4: Add this average curve to the forecast results in the testing process.

IV. FORECASTING RESULTS

For practical applications, the dataset from ten solar sites and NWP data from the CWB in Taiwan were used to verify the performance of the proposed model. The installed capacities of the ten PV sites were 177.735 kW, 580.0 kW, 493.92 kW, 387.0 kW, 182.16 kW, 203.42 kW, 368.42 kW, 251.685 kW, 248.0 kW, and 907.68 kW.

The inputs to the proposed forecasting model included historical solar power, historical solar irradiance, and NWP data, including solar irradiance, temperature, rainfall, air pressure, wind speed, and humidity. In the first nine PV sites, this work collected the data from June to November 2019, whereas PV site 10 had a two-year data range (2019–2021). The data collection for both sources was difficult and so the data collection at the first nine PV sites was of short duration. The datasets from PV sites 1–9 were split as follows:

- Data from July 3rd to October 13th (2280 data points) were used as the training set.

TABLE III
TRAINING AND TESTING OF THE PROPOSED MODEL AT 10 SITES

PV Site	Training		Testing	
	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)
1	2.81	5.52	1.52	2.92
2	2.45	4.28	1.64	3.08
3	3.56	5.32	2.18	4.43
4	4.19	7.14	2.67	5.29
5	5.23	7.55	2.84	5.67
6	5.37	7.03	2.89	5.27
7	3.20	5.00	1.77	3.45
8	3.56	5.82	1.99	3.74
9	4.81	9.79	3.69	7.51
10	6.89	7.76	7.72	8.11

- Data from October 14th to October 31st (408 data points) were used as the validation set.
- Data from November 1st to November 18th (432 data points) were used as the testing set.

Therefore, selecting 18 days for the prediction period (70% training, 15% validation, and 15% testing) is reasonable because the periods of the collected data are of short duration.

The dataset for PV site 10 was split as follows:

- Data from July 2019 to July 2021 were used for training (80%).
- Data from August 2021 were used as the validation set (10%).
- Data from September 2021 were used as the testing set (10%).

Forecasting results using the proposed model for all ten PV sites are summarized in Table III. Although the testing days were relatively few, a total of ten sites were considered. Therefore, the forecasting results for the ten PV sites can be used to justify the model performance.

Based on Table III, owing to the limited data at the first nine sites, the training errors exceeded the testing errors. For a PV power forecasting task, it is possible to have testing error lower than training error, which means the forecasting method generalizes well when testing set is smaller than training set, and the testing set has less noise from moving clouds or other weather factors. To reduce the effect of data size, this study considered ten PV sites at different areas, which is enough to confirm the superiority of the proposed method. However, because the training period for PV site 10 was of a longer duration, the prevalence of training errors was close to that of testing errors. In addition, for all ten sites, the training errors decreased inversely with the number of epochs of the GRU model, revealing that the proposed model fits well.

To demonstrate the effectiveness of the proposed approaches, the results of solar power forecasting using the developed model were compared with other benchmark methods, including ANN, LSTM, XGBoost, and GRU. To verify the importance of data

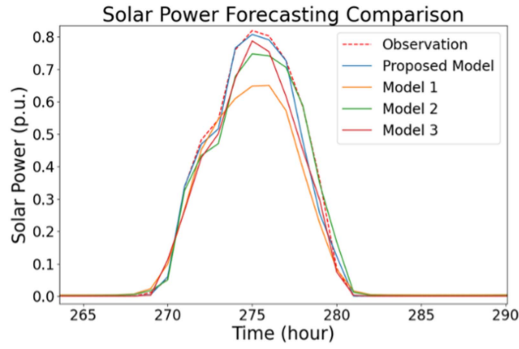


Fig. 13. Effectiveness of the proposed model during a sunny day.

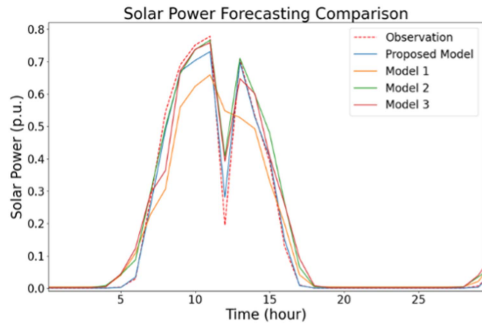


Fig. 14. Effectiveness of the proposed model during a non-sunny day.

TABLE IV
PERFORMANCE OF DIFFERENT ALGORITHMS

Model	Training		Testing	
	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)
Proposed model	2.81	5.52	1.52	2.92
ANN	6.37	9.67	3.28	5.26
LSTM	4.03	7.81	2.33	3.10
XGBoost	1.17	2.66	1.95	3.05
GRU	3.22	6.40	2.74	3.34

preprocessing, feature engineering and postprocessing of error correction, this work compared the proposed model with the following models:

- *Model 1*: a typical ANN model with preprocessing and postprocessing.
- *Model 2*: a traditional single layer GRU with the same parameters as our proposed model but without preprocessing, feature engineering and postprocessing.
- *Model 3*: a GRU model with preprocessing but without postprocessing.

Figs. 13 and 14 display the real measurements and forecasts of a sunny day and non-sunny day, respectively, using the proposed model, Model 1, Model 2, and Model 3 for solar site 10. To investigate the performance of the proposed model, the evaluation indexes NRMSE and NMAPE were used. Table IV shows the

forecasting results of the models for solar site 1. Note that all models involved data preprocessing, feature engineering, and postprocessing processes. Table V shows a comparison between the proposed model and Model 1, Model 2, and Model 3 for the ten solar sites. The following three perspectives can be taken:

Perspective 1. Performance of the proposed model: The training and testing results for all algorithms are shown in Table IV. The proposed model has a testing NMAPE of 1.52% and NRMSE of 2.92%, indicating superior performance compared to the other algorithms. This is because the GRU considers the effect of features on solar power output at the next time step and improves the forecasting accuracy by using historical time series data. The training results of XGBoost were the smallest and the testing results were close to the proposed model. The ANN model had the worst prediction performance. However, the proposed model had a longer training time because the RNNs needed to be processed sequentially. In addition, every GRU cell had its own update gate and reset gate, creating more operations in each of them.

Perspective 2. The impact of preprocessing, feature engineering and post processing on solar power forecasting: As can be seen from Table V, by using the proposed model, one can obtain an average error of 2.90% (NMAPE) or 4.95% (NRMSE) for all ten PV sites, indicating that the proposed model outperforms the other three models. In Model 1, the use of the ANN model obtained the worst forecasting result even though it included preprocessing and postprocessing, demonstrating the superiority of deep learning methods. In Model 2, the use of a GRU without pre- and post-processing generated a low degree of forecasting accuracy, demonstrating the importance of pre- and postprocessing. Model 3, which consisted of the same GRU forecasting model and preprocessing methods including feature selection and KPCA but without an error correction method (postprocessing technique), performed well compared to the proposed model at four of the PV sites. However, the average error for the 10 sites exceeded that of the proposed model. It can be concluded that the proposed model tracks the observed values better than Model 1, Model 2, and Model 3 in both examined cases: a sunny day and a non-sunny day.

Perspective 3. Contribution of preprocessing, feature engineering and post processing: The main contribution of this paper is to combine preprocessing, feature engineering, and postprocessing. Notably, all models were designed to use the same forecasting engine, i.e., a GRU, so as to demonstrate the independent contributions of preprocessing and postprocessing. The comparisons between the proposed model and the reference models were made as follows:

- 1) Preprocessing only: Models A, B, and C:
 - a) *Model A*: Include feature selection using the Pearson Correlation Coefficient (PCC) and XGBoost.
 - b) *Model B*: Used KPCA to reduce the dimension of the dataset.
 - c) *Model C*: Used both Feature selection and KPCA (a complete preprocessing method combining both A and B).
- 2) Postprocessing only: Model D.
- 3) No preprocessing or postprocessing: Model E.

TABLE V
COMPARISON OF FORECASTING RESULTS OF 10 PV SITES

	Proposed model		Model 1 Traditional ANN (with pre- and post- processing)		Model 2 Deep learning - GRU (without pre- and post- processing)		Model 3 Deep learning - GRU (without post-processing)	
	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)
Site 1	1.52	2.92	3.28	5.74	2.80	4.96	2.63	4.78
Site 2	1.64	3.08	2.90	5.57	2.52	4.59	2.18	4.24
Site 3	2.18	4.43	3.45	6.81	2.65	5.27	2.36	4.90
Site 4	2.67	5.29	3.56	5.95	2.81	5.35	2.71	5.30
Site 5	2.84	5.67	3.22	6.82	2.60	4.80	2.57	4.78
Site 6	2.89	5.27	5.07	8.12	2.86	5.07	2.85	5.05
Site 7	1.77	3.45	3.42	5.23	1.70	3.29	1.69	3.24
Site 8	1.99	3.74	3.70	6.85	3.73	6.46	2.87	5.92
Site 9	3.69	7.51	6.56	9.20	3.64	7.71	3.58	7.46
Site 10	7.72	8.11	9.80	10.56	8.43	9.03	8.05	8.92
Average	2.90	4.95	4.50	7.00	3.40	5.65	3.15	5.46

TABLE VI
CONTRIBUTION OF EACH COMPONENT IN PROPOSED METHOD

	With pre- and post-processing		With only preprocessing						With only postprocessing		Without pre- and post-processing	
	Proposed model		Model A With feature selection (PCC + XGBoost)		Model B With KPCA (No feature selection)		Model C With feature selection and KPCA (PCC + XGBoost + KPCA)		Model D Error correction method		Model E Only GRU	
	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)	NMAPE (%)	NRMSE (%)
Site 1	1.52	2.92	2.65	4.79	2.66	4.80	2.63	4.78	2.72	4.83	2.80	4.96
Site 2	1.64	3.08	2.19	4.26	2.21	4.29	2.18	4.24	2.35	4.56	2.52	4.59
Site 3	2.18	4.43	2.48	5.00	2.52	5.04	2.36	4.90	2.57	5.11	2.65	5.27
Site 4	2.67	5.29	2.89	5.33	2.97	5.35	2.71	5.30	2.78	5.32	2.81	5.35
Site 5	2.84	5.67	2.58	4.78	2.59	4.79	2.57	4.78	2.58	4.79	2.60	4.80
Site 6	2.89	5.27	2.87	5.11	2.92	5.19	2.85	5.05	2.86	5.06	2.86	5.07
Site 7	1.77	3.45	1.69	3.26	1.70	3.28	1.69	3.24	1.70	3.26	1.70	3.29
Site 8	1.99	3.74	3.03	5.98	3.34	6.21	2.87	5.92	3.49	6.15	3.73	6.46
Site 9	3.69	7.51	3.60	7.65	3.63	7.59	3.58	7.46	3.62	7.68	3.64	7.71
Site 10	7.72	8.11	8.27	8.96	8.38	8.99	8.05	8.92	8.33	9.00	8.43	9.03
Average	2.90	4.95	3.23	5.51	3.29	5.55	3.15	5.46	3.30	5.58	3.40	5.65

Based on Table VI, it can be observed that the average error of the proposed model was the lowest. Model C had the best performance of the preprocessing only group, and generated good forecasting results at PV sites 5, 6, 7, and 9. The average error using Model A, B, or C was less than that using Model D,

revealing that preprocessing plays a larger role than postprocessing. Model E (no preprocessing and no postprocessing) had the worst performance. Therefore, it can be concluded that removing either preprocessing or postprocessing process does not lead to better performance.

This study aims to use either short-term or long-term training data for model learning. PV sites 1–9 provide a short-term training data (around 3–4 months), while PV site 10 provide a long-term training data (2 years). Both forecasting cases indicate that the proposed forecasting model surpasses other state-of-the-art models no matter the length of training data, demonstrating that the proposed method has good reliability and stability. Indeed, a long-term data would be more useful for training deep learning models, and that is the reason why PV site 10 was selected in this study. Two-year training data at PV site 10 already cover all weather scenarios in a year, and the forecasting result at PV site 10 also has the best performance.

In this study, the testing data for all PV sites are in Autumn (from September to November), so weather conditions at different sites are similar. As for the PV forecasting in different seasons, predictors need to consider more techniques, such as weather classification, data decomposition or denoising, to predict PV power generation at different seasons.

V. CONCLUSION

This study proposed a novel model for forecasting solar power generation, which involves data preprocessing, feature engineering based on the PCC and XGBoost, KPCA calculation, a GRU model with time-of-day clustering as well as postprocessing based on error correction. The experimental results demonstrated that the proposed model has better forecasting accuracy than other typical models and can achieve good prediction results. In addition, a GRU can deal with a larger dataset because it has fewer parameters and shorter training time compared with LSTM. The process of feature engineering was essential because it captured the most important features of the dataset and improved the quality of the dataset for training the model. From the correlation analysis results, it was found that the PCC could extract the main features along with analyzing the relationship between historical and future solar power generation. KPCA had the ability to reduce noise and improve the validity of the input information. Although several models with feature extraction have been mentioned in the literature—which is also an important process for stabilizing a model—it has rarely been of serious interest.

In conclusion, this study developed a complete framework for short-term solar power forecasting. Achieving more accurate solar power forecasting can effectively help system operators make reasonable dispatching decisions and engage in energy planning.

REFERENCES

- [1] Z. Xin-gang and Z. You, "Technological progress and industrial performance: A case study of solar photovoltaic industry," *Renewable Sustain. Energy Rev.*, vol. 81, pp. 929–936, 2018.
- [2] M. J. Hossain and M. A. Mahmud, *Renewable Energy Integration Challenges and Solutions*. Singapore: Springer, 2014.
- [3] A. Taşçıkaraoğlu, "Chapter 5 - Impacts of accurate renewable power forecasting on optimum operation of power system," in *Optimization in Renewable Energy Systems*, O. Erdinc, Ed. Boston, MA, USA: Butterworth-Heinemann, 2017, pp. 159–175.
- [4] S. J. Steffel, P. R. Caroselli, A. M. Dinkel, J. Q. Liu, R. N. Sackey, and N. R. Vadhar, "Integrating solar generation on the electric distribution grid," *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 878–886, Jun. 2012.
- [5] R. Ahmed, V. Sreeram, Y. Mishra, and M. D. Arif, "A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization," *Renew. Sustain. Energy Rev.*, vol. 124, 2020, Art. no. 109792.
- [6] S. I. Sulaiman, T. K. Abdul Rahman, I. Musirin, and S. Shaari, "Artificial neural network versus linear regression for predicting grid-connected photovoltaic system output," in *Proc. IEEE Int. Conf. Cyber Technol. Automat. Control Intell. Syst.*, 2012, pp. 170–174.
- [7] Y. Hu, W. Lian, Y. Han, S. Dai, and H. Zhu, "A seasonal model using optimized multi-layer neural networks to forecast power output of PV plants," *Energies*, vol. 11, no. 2, 2018, Art. no. 326.
- [8] H. S. Jang, K. Y. Bae, H.-S. Park, and D. K. Sung, "Solar power prediction based on satellite images and support vector machine," *IEEE Trans. Sustain. Energy*, vol. 7, no. 3, pp. 1255–1263, Jul. 2016.
- [9] V. Vapnik, *The Nature of Statistical Learning Theory*. Berlin, Germany: Springer, pp. 289–290, 1999.
- [10] F. Bizzarri, M. Bongiorno, A. Brambilla, G. Grusso, and G. S. Gajani, "Model of photovoltaic power plants for performance analysis and production forecast," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 278–285, Apr. 2013.
- [11] E. İzgi, A. Öztöpal, B. Yerli, M. K. Kaymak, and A. D. Şahin, "Short-term solar power prediction by using artificial neural networks," *Sol. Energy*, vol. 86, no. 2, pp. 725–733, 2012.
- [12] H. Zhu, W. Lian, L. Lu, S. Dai, and Y. Hu, "An improved forecasting method for photovoltaic power based on adaptive BP neural network with a scrolling time window," *Energies*, vol. 10, no. 10, 2017, Art. no. 1542.
- [13] J. Song, V. Krishnamurthy, A. Kwasinski, and R. Sharma, "Development of a Markov-chain-based energy storage model for power supply availability assessment of photovoltaic generation plants," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 491–500, Apr. 2013.
- [14] H. Long, Z. Zhang, and Y. Su, "Analysis of daily solar power prediction with data-driven approaches," *Appl. Energy*, vol. 126, pp. 29–37, 2014.
- [15] A. U. Haque, M. H. Nehrir, and P. Mandal, "Solar PV power generation forecast using a hybrid intelligent approach," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2013, pp. 1–5.
- [16] M. Abuellla and B. Chowdhury, "Random forest ensemble of support vector regression models for solar power forecasting," in *Proc. IEEE Power Energy Soc. Innov. Smart Grid Technol. Conf.*, 2017, pp. 1–5.
- [17] G. Cervone, L. Clemente-Harding, S. Alessandrini, and L. Delle Monache, "Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble," *Renew. Energy*, vol. 108, pp. 274–286, 2017.
- [18] M. A. Mat Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, and F. Hussin, "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review," *Renew. Sustain. Energy Rev.*, vol. 70, pp. 1108–1118, 2017.
- [19] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, 2018, Art. no. 1255.
- [20] A. Yona, T. Senjyu, and T. Funabashi, "Application of recurrent neural network to short-term-ahead generating power forecasting for photovoltaic system," in *Proc. IEEE Power Eng. Soc. Gen. Meeting*, 2007, pp. 1–6.
- [21] A. Alzahrani, P. Shamsi, M. Ferdowsi, and C. Dagli, "Solar irradiance forecasting using deep recurrent neural networks," in *Proc. IEEE 6th Int. Conf. Renewable Energy Res. Appl.*, 2017, pp. 988–994.
- [22] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, 2018.
- [23] S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 6, pp. 107–116, 1998.
- [24] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, pp. 2727–2740, 2019.
- [25] Y. Wang, W. Liao, and Y. Chang, "Gated recurrent unit network-based short-term photovoltaic forecasting," *Energies*, vol. 11, no. 8, 2018, Art. no. 2163.
- [26] B. Liu, C. Fu, A. Bielefeld, and Y. Q. Liu, "Forecasting of Chinese primary energy consumption in 2021 with GRU artificial neural network," *Energies*, vol. 10, no. 10, 2017, Art. no. 1453.
- [27] N. Sodsong, K. M. Yu, and W. Ouyang, "Short-Term solar PV forecasting using gated recurrent unit with a cascade model," in *Proc. IEEE Int. Conf. Artif. Intell. Inf. Commun.*, 2019, pp. 292–297.
- [28] H. Nielsen, T. Nielsen, H. Madsen, M. Pindado, and I. Marti, "Optimal combination of wind power forecasts," *Wind Energy*, vol. 10, pp. 471–482, Nov. 2007.

- [29] Q.-T. Phan, Y.-K. Wu, and Q.-D. Phan, "An overview of data preprocessing for short-term wind power forecasting," in *Proc. IEEE 7th Int. Conf. Appl. System Innov.*, 2021, pp. 121–125.
- [30] Z. Zhang et al., "Solar radiation intensity probabilistic forecasting based on K-Means time series clustering and gaussian process regression," *IEEE Access*, vol. 9, pp. 89079–89092, 2021.
- [31] C.-M. Huang, C.-J. Kuo, S.-J. Chen, and S.-P. Yang, "One-day-ahead hourly forecasting for photovoltaic power generation using an intelligent method with weather-based forecasting models," *IET Gener., Transmiss. Distrib.*, vol. 9, pp. 1874–1882, 2015.
- [32] H. Zhu, Y. Shi, H. Wang, and L. Lu, "New feature extraction method for photovoltaic array output time series and its application in fault diagnosis," *IEEE J. Photovolt.*, vol. 10, no. 4, pp. 1133–1141, Jul. 2020.
- [33] Y. N. Huizhi Gou, "Forecasting model of photovoltaic power based on KPCA-MCS-DCNN," *Comput. Model. Eng. Sci.*, vol. 128, no. 2, pp. 803–822, 2021.
- [34] R. Hossain, A. Than Oo, and A. Ali, "The effectiveness of feature selection method in solar power prediction," *J. Renew. Energy*, vol. 2013, 2013, Art. no. 952613.
- [35] E. Lorenz, J. Hurka, D. Heinemann, and H. G. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 2, no. 1, pp. 2–10, Mar. 2009.
- [36] S. Pelland, G. N. Galanis, and G. Kallos, "Solar and photovoltaic forecasting through post-processing of the global environmental multiscale numerical weather prediction model," *Prog. Photovolt. Res. Appl.*, vol. 21, pp. 284–296, 2013.
- [37] H. Verbois, Y.-M. Saint-Drenan, A. Thiery, and P. Blanc, "Statistical learning for NWP post-processing: A benchmark for solar irradiance forecasting," *Sol. Energy*, vol. 238, pp. 132–149, 2022.
- [38] C. Voyant et al., "Machine learning methods for solar radiation forecasting: A review," *Renew. Energy*, vol. 105, pp. 569–582, 2017.
- [39] D. W. van der Meer, J. Widén, and J. Munkhammar, "Review on probabilistic forecasting of photovoltaic power production and electricity consumption," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 1484–1512, 2018.
- [40] K. Cho, B. Merriënboer, C. Gulcehre, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2014, pp. 1724–1734.
- [41] R. K. Agrawal, F. Muchahary, and M. M. Tripathi, "Long term load forecasting with hourly predictions based on long-short-term-memory networks," in *Proc. IEEE Texas Power Energy Conf.*, 2018, pp. 1–6.
- [42] O. De Jesus and M. T. Hagan, "Backpropagation algorithms for a broad class of dynamic networks," *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp. 14–27, Jan. 2007.
- [43] E. Yazan and M. F. Talu, "Comparison of the stochastic gradient descent based optimization techniques," in *Proc. IEEE Int. Artif. Intell. Data Process. Symp.*, 2017, pp. 1–5.
- [44] Q. Li, X. Zhang, T. Ma, C. Jiao, H. Wang, and W. Hu, "A multi-step ahead photovoltaic power prediction model based on similar day, enhanced colliding bodies optimization, variational mode decomposition, and deep extreme learning machine," *Energy*, vol. 224, 2021, Art. no. 120094.
- [45] Q.-T. Phan, Y.-K. Wu, Q.-D. Phan, and H.-Y. Lo, "A novel forecasting model for solar power generation by a deep learning framework with data preprocessing and postprocessing," in *Proc. IEEE/IAS 58th Ind. Commercial Power Syst. Tech. Conf.*, 2022, pp. 1–11.