

**ARTICLE****Short-Term Prediction of Photovoltaic Power Based on Fusion Device Feature-Transfer**Zhongyao Du<sup>1,\*</sup>, Xiaoying Chen<sup>1</sup>, Hao Wang<sup>2</sup>, Xuheng Wang<sup>1</sup>, Yu Deng<sup>1</sup> and Liying Sun<sup>1</sup><sup>1</sup>College of Electrical Engineering, Liaoning University of Technology, Jinzhou, 121001, China<sup>2</sup>School of Electrical Engineering, Shandong University, Jinan, 250012, China

\*Corresponding Author: Zhongyao Du. Email: druid\_y@live.com

Received: 15 November 2021 Accepted: 18 February 2022

**ABSTRACT**

To attain the goal of carbon peaking and carbon neutralization, the inevitable choice is the open sharing of power data and connection to the grid of high-permeability renewable energy. However, this approach is hindered by the lack of training data for predicting new grid-connected PV power stations. To overcome this problem, this work uses open and shared power data as input for a short-term PV-power-prediction model based on feature transfer learning to facilitate the generalization of the PV-power-prediction model to multiple PV-power stations. The proposed model integrates a structure model, heat-dissipation conditions, and the loss coefficients of PV modules. Clear-Sky entropy, characterizes seasonal and weather data features, describes the main meteorological characteristics at the PV power station. Taking gate recurrent unit neural networks as the framework, the open and shared PV-power data as the source-domain training label, and a small quantity of power data from a new grid-connected PV power station as the target-domain training label, the neural network hidden layer is shared between the target domain and the source domain. The fully connected layer is established in the target domain, and the regularization constraint is introduced to fine-tune and suppress the overfitting in feature transfer. The prediction of PV power is completed by using the actual power data of PV power stations. The average measures of the normalized root mean square error (NRMSE), the normalized mean absolute percentage error (NMAPE), and the normalized maximum absolute percentage error (NLAE) for the model decrease by 15%, 12%, and 35%, respectively, which reflects a much greater adaptability than is possible with other methods. These results show that the proposed method is highly generalizable to different types of PV devices and operating environments that offer insufficient training data.

**KEYWORDS**

Solar power generation; transfer learning; PV module; UMAP; GRU; overfitting

**Nomenclature**

$D_S$	Source domain dataset
$D_T$	Target domain dataset
$I_E$	Effective irradiance (W/m <sup>2</sup> )
$I_D$	Direct irradiance (W/m <sup>2</sup> )
$I_{DIF}$	Diffuse irradiance (W/m <sup>2</sup> )
$I_G$	Global irradiance (W/m <sup>2</sup> )



$\theta$	Solar declination angle on PV array (°)
$\beta$	PV array inclination (°)
$\gamma$	PV array azimuth (°)
$\mu$	Ground reflection coefficient
$\alpha$	Solar altitude angle (°)
$\gamma_s$	Solar altitude azimuth (°)
$\varepsilon$	Solar altitude declination angle (°)
$\omega$	Hour angle (°)
$\varphi$	Local latitude (°)
$t$	Time (minutes)
$T_m$	Module temperature (°C)
$T_{amb}$	Ambient temperature (°C)
$v$	Actual wind speed (m/s)
$v_w$	Wind speed in weather dataset (m/s)
$h$	Device height (m)
$a, b$	heat dissipation coefficient
$L_P$	System losses (%)
$L_{LID}$	Light-induced degradation losses (%)
$L_N$	Nominal losses (%)
$L_S$	Shadow Losses (%)
$I_C$	Clear-Sky irradiance (W/m <sup>2</sup> )
$d$	Euclidean distance
$M$	Fuzzy function
$\sigma$	Standard deviation
$E_C$	Clear-Sky entropy
$S_C$	Seasonal characteristic coefficient
$S_G$	Seasonal grade
$I_{dmax}$	Daily maximum effective irradiance (W/m <sup>2</sup> )
$I_{imax}$	Maximum effective irradiance of ideal value in current season (W/m <sup>2</sup> )
$G_{ij}$	Joint probability distribution of any two origin-data points
$Q_{ij}$	Joint probability distribution of any two features
$k, p, q$	UMAP hyperparameter
$C_E$	Cross entropy
$U$	Weather feature dataset
$Z_t$	Update gate
$R_t$	Reset gate
$x_t$	Input state
$h_t$	Hidden state
$W$	Weight matrix
$\delta$	Neural network parameter
$l$	Dataset of label
$F$	Dataset of feature
$\lambda$	L2 regularization penalty coefficient

## 1 Introduction

To attain the goal of carbon peaking and carbon neutrality, estimates indicate that the fraction of power generation from renewable energy in China must increase from the current 8% to over 60% by 2060 [1]. Unfortunately, power generation from renewable energy is intermittent, with random and volatile fluctuations in power output [2]. In particular, modern photovoltaic (PV) power stations lack power data labels, and their power output is difficult to predict with accuracy. Connecting such power stations to the large-scale power grid makes it challenging to maintain the stable operation of the power grid [3]. Accurate short-term forecasting of PV power generation is thus urgently needed for the day-ahead forecasting of PV power input into the power grid [4,5] and would significantly improve PV penetration into the power grid, thereby making a major step toward the goal of net-zero carbon dioxide emissions.

Short-term prediction of PV power generation essentially combines a physical model with a statistical model. Given the limited knowledge of the mechanism and the lack of corrective feedback, the physical model is rarely used alone [6]. Statistical models include the time series analysis method [7] and the machine learning method [8]. The literature widely discusses the machine learning method [9], which is highly generalizable for nonlinear mapping. For example, Chu et al. [10] developed the reforecasting method, which is based on artificial neural network optimization schemes and improves the performance of physical deterministic models based on cloud-tracking techniques. Nie et al. [11] proposed a framework that classifies input images of the sky into various sky conditions and then sends the classified images to specific convolutional neural network models to predict PV power output. These artificial neural network models can approximate the nonlinear relationship between the input features and the prediction target.

Moreover, the time series analysis and physical models in question are based on proven mathematical theories that can rapidly map features, which promotes the emergence of hybrid prediction models that offer multiple advantages [12]. Wang et al. [13] used time correlation modification in an integrated long-short-term memory (LSTM) recurrent neural network algorithm to calculate solar irradiance and PV power generation, which provided more accurate predictions than traditional machine learning models. In other work, Gao et al. [14] divided the daily weather conditions into ideal and non-ideal weather and developed a prediction model based on a LSTM neural network to efficiently map dynamic characteristics of both ideal and non-ideal weather. Adar et al. [15] used principal component analysis (PCA) to extract two principal components from 11 input variables for three technical types of PV modules to conveniently calculate the performance of each PV module.

Despite these advances, the application of machine learning algorithms in the power-generation industry remains challenging due to the lack of training data, especially for new PV power plants. To address this issue, Lee et al. [16] applied an online learning algorithm to gradually improve the adaptability of an algorithm for predicting PV power. Compared with the fixed model, the online algorithm predicts PV power with significantly greater accuracy. In addition, Wang et al. [17] studied the use of the generative adversarial network to expand an extreme weather training dataset.

Another solution to this problem originates from the open- and shared-power dataset, which is a new production factor [18]. Using this dataset, we develop in the present work a short-term PV-power-prediction model based on fusion device feature-transfer (FFT). Specifically, [Section 2](#) proposes applying a feature-fusion method to PV modules and adopts as parameters for the PV module its effective irradiance, temperature (considering heat dissipation), and system losses. The physical characteristics of the different technical types of PV modules are integrated. We then propose a local meteorological model for predicting PV power. The model extracts the Clear-Sky entropy, seasonal and

weather features to characterize the main meteorological (Section 3). Section 4 introduces a transfer learning (TL) neural network based on a gated recurrent unit (GRU) and a fully connected (FC) layer with fine-tuning and regularization constraints in feature-transfer to localize the prediction model. Section 5 evaluates the ability of the uniform manifold approximation and projection dimension reduction (UMAP) algorithm to extract a weather dataset given the global structure, redundancy, projection time, and feature-set fitting efficiency. Finally, we compare the PV-power predictions and errors produced by the local model, no-TL model, TL-PV model, and FFT-PV model and analyze the operation of the proposed model when localized to a specific environment.

## 2 Characteristics of Photovoltaic Modules

The TL training set can be divided into source domain  $D_S$  and target domain  $D_T$ , where  $D_S$  is the historical power data of PV plants that have long operated while connected to the grid, and  $D_T$  is the power data from new PV plants. In TL, the effective irradiance, PV module temperature, and PV system loss are used to model the characteristics of PV modules so that the different types of PV modules can be mapped to a single feature space, thereby reducing differences between distributions to the different domains.

### 2.1 Effective Irradiance

To account for PV devices with different tracker systems, the effective solar irradiance on the PV array must be accurately modeled [19]. This section calculates the irradiance for a fixed PV array or a PV array with a single-axis tracker or a dual-axis tracker. To begin, the effective irradiance  $I_E$  of a PV array is calculated by using

$$I_E = I_D \cos \theta + I_{DIF} \left( \frac{1 + \cos \beta}{2} \right) + \mu I_G \left( \frac{1 - \cos \beta}{2} \right), \quad (1)$$

where  $I_D$  is the direct irradiance ( $\text{W}/\text{m}^2$ ),  $I_{DIF}$  is the diffuse irradiance ( $\text{W}/\text{m}^2$ ),  $I_G$  is the global irradiance ( $\text{W}/\text{m}^2$ ),  $\theta$  is the solar declination angle with respect to the PV array (degrees),  $\beta$  is the inclination of the PV array (degrees), and  $\mu$  is the ground reflection coefficient. When the system is fixed or has a single-axis tracker,  $\beta$  is the installed angle.

The solar declination angle  $\theta$  is

$$\theta = \arccos [\cos \beta \sin \alpha + \sin \beta \cos \alpha \cos (\gamma_s - \gamma)], \quad (2)$$

where  $\alpha$  is the solar altitude angle (degrees),  $\gamma$  is the PV array azimuth (degrees), and  $\gamma_s$  is the solar azimuth (degrees). When the PV array is fixed,  $\gamma$  is the installed value.

The solar azimuth  $\gamma_s$  and solar altitude angle  $\alpha$  are given by

$$\gamma_s = \arcsin \left( \frac{\cos \delta \sin \omega}{\cos \alpha} \right), \quad (3)$$

$$\alpha = \arcsin (\sin \varphi \sin \varepsilon + \cos \varphi \cos \varepsilon \cos \omega), \quad (4)$$

$$\sin \varepsilon = 0.39795 \cos [0.98563(n - 173)], \quad (5)$$

$$\omega = 15(t - 12), \quad (6)$$

where  $\varepsilon$  is the solar declination angle (degrees),  $\omega$  is the hour angle (degrees),  $\varphi$  is the local latitude (degrees),  $n$  is the number of days since January 01 [20], and  $t$  is the time of day in 24-hour format.

## 2.2 Module Temperature

The temperature of the PV module strongly affects the power generation, so accurately calculating the temperature is vital to improve the PV-power-prediction accuracy [21]. The heat dissipation of the PV module is affected by its mounting [22] and encapsulating materials [23]. This section quantifies how the installation orientation of PV modules and the module materials affect the heat dissipation. The temperature of the PV module is given by

$$T_m = T_{amb} + \left( \frac{\Delta T}{1000} + e^{a+bv} \right) G_c, \quad (7)$$

$$v = \begin{cases} 0.51v_w, & h \leq 5 \\ 0.61v_w, & h > 5, \end{cases} \quad (8)$$

where  $T_{amb}$  is the ambient air temperature ( $^{\circ}\text{C}$ ),  $a$  and  $b$  are the heat dissipation coefficients (given in Table 1 [24]),  $v$  is the actual wind speed (m/s),  $v_w$  is the wind speed from the weather dataset (m/s), and  $h$  is the height of the PV module (m).

**Table 1:** Heat dissipation coefficients for PV module

Mount	Encapsulate	$a$	$b$ (s/m)	$\Delta T$ ( $^{\circ}\text{C}$ )
Rack	Glass/cell/glass	-3.47	-0.0594	3
Rack	Glass/cell/polymer	-3.56	-0.0750	3
Floor	Glass/cell/glass	-2.98	-0.0471	1
Floor	Glass/cell/polymer	-2.81	-0.0455	0

## 2.3 System Losses

The PV power is affected by its local geographical environment, the degradation of facilities, and other factors [25]. The system losses  $L_p$  are expressed in the form of a percentage:

$$L_p = \left[ 1 - \prod_i \left( 1 - \frac{L_i}{100} \right) \right] \times 100\%, \quad (9)$$

where  $L_i$  is the set of  $L_{LID}$ ,  $L_N$ , and  $L_S$  [ $L_{LID}$  is the light-induced degradation (LID) losses (%),  $L_N$  is the nominal loss (%), and  $L_S$  is the shadow loss (%)]. Initially  $L_{LID} = 1.5\%$  for a PV array, and it increases by 0.5% per year [26]. The default shading loss for an “unshaded” PV module is 3% [27], which increases if the shadow area in the all-sky imager is calculated.

## 3 Main Meteorological Characteristics

### 3.1 Clear-Sky Entropy

To account for the influence of meteorological conditions, the solar irradiance is calculated based on the effective irradiance  $I_E$  and the ideal irradiance  $I_C$ , where  $I_C$  is the maximum irradiance not affected by atmospheric conditions (i.e., the Clear-Sky irradiance [28]). We use fuzzy entropy [29] to calculate the self-similarity of the difference  $x(t)$  between  $I_C$  and  $I_E$ , and the Clear-Sky entropy  $E_C$  serves to quantify the uncertainty of the weather conditions. The calculation is done in the following steps:

(1) Reconstruct the phase space of the time series  $x(t)$  and set the dimension to  $m$ :

$$X(i) = [x(i), x(i+1), \dots, x(i+m-1)] - x_0(i), \quad (10)$$

$$x_0(i) = \frac{1}{m} \sum_{j=0}^{m-1} x(i+j), \quad \{i | i \in Z\} = [1, n-m+1]. \quad (11)$$

(2) The distance  $d_{ij}$  separates the sequence  $x(i)$  from the sequence  $x(j)$ :

$$d_{ij}^m = \max_{k=0}^{m-1} |[x(i+k) - x_0(i)] - [x(j+k) - x_0(j)]|. \quad (12)$$

(3) Calculate the similarity  $M_{ij}^m$  between sequences by using the fuzzy function:

$$M_{ij}^m = \exp \left[ -\frac{(d_{ij}^m)^n}{\sigma} \right], \quad (13)$$

where  $\sigma$  is the standard deviation of the time series  $x(t)$ .

(4) Calculate the similarity of each  $j$  and take the average, then repeat Steps 1–3 to calculate the average of each  $i$ , which is given as

$$\Psi^m(\sigma) = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} \left( \frac{1}{n-m} \sum_{j=1, j \neq i}^{n-m+1} M_{ij}^m \right). \quad (14)$$

(5) The Clear-Sky entropy  $E_C$  of this time series is

$$E_C = \ln \Psi^m(\sigma) - \ln \Psi^{m+1}(\sigma). \quad (15)$$

During effective power generation, the short-term power prediction calculates the Clear-Sky entropy  $E_C$  in intervals of four hours.

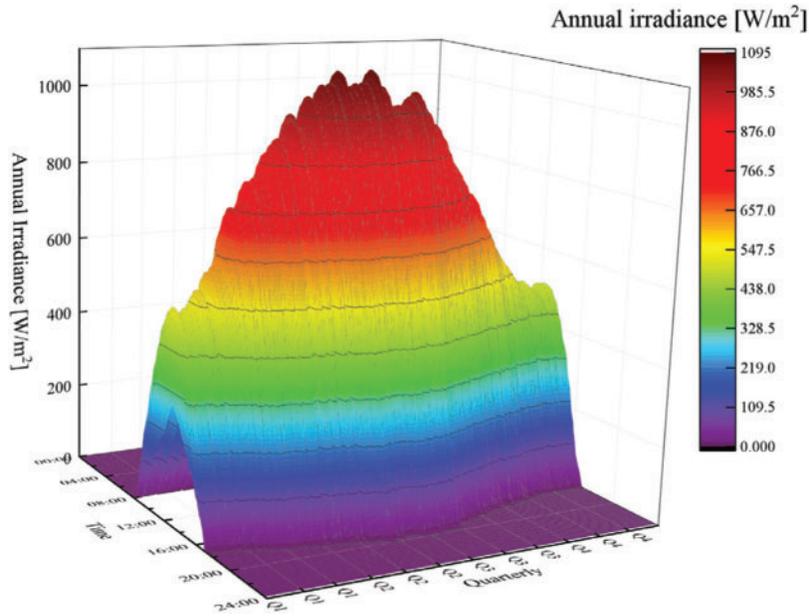
### 3.2 Seasonal Characteristics

The ideal solar irradiance varies seasonally with a strong time correlation. However, the effective irradiance is random, which is not conducive to extracting time-series features. Fig. 1 shows the annual irradiance.

To improve the efficiency with which the irradiance time series characteristics are recognized, we divide the irradiance into four levels and construct the seasonal characteristic coefficient  $S_C$  to quantify the irradiance time series, as follows:

$$S_C = S_G + \frac{I_{d\max}}{I_{i\max}}, \quad (16)$$

where  $S_G$  is the seasonal grade (i.e., the average irradiance of each of the four seasons from small to large are denoted grades 1 to 4, respectively),  $I_{d\max}/I_{i\max}$  is the uncertainty coefficient (the closer it approaches to unity, the closer the irradiance approaches that of the next season),  $I_{d\max}$  is the maximum effective daily irradiance ( $\text{W}/\text{m}^2$ ), and  $I_{i\max}$  is the maximum ideal irradiance in the current season ( $\text{W}/\text{m}^2$ ).



**Figure 1:** Annual irradiance as a function of time of day and month of year

### 3.3 Weather Feature Based on Uniform Manifold Approximation

The weather dataset is characterized by numerous features, which ensure a large amount of data and little distinction between features. The feature extraction based on t-distributed stochastic neighbor embedding (t-SNE) affects how the relationship between the original data and the prediction algorithm are recognized due to the loss of global structure. This section uses the UMAP algorithm [30] to construct as follows the features of the weather dataset:

- (1) Consider a dataset  $D = \{x_1, x_2, \dots, x_n\}$ , where the data points are unnormalized Euclidean distance to increase the reliability and speed of calculating high-dimensional data compared with t-SNE. The conditional probability  $G_{ij}$  is

$$G_{ij} = \exp\left(-\frac{d(x_i, x_j) - \rho_i}{\sigma_i}\right), \tag{17}$$

$$\rho_i = \min\{d(x_i, x_j) \mid 1 \leq j \leq k, d(x_i, x_j) > 0\}, \tag{18}$$

where  $d$  is the Euclidean distance between points  $x_i$  and  $x_j$ ,  $\rho_i$  is the distance from data point  $i$  to the nearest data point  $j \neq i$ , and  $\sigma_i$  is the variance of distance.

- (2) The parameter  $k$  controls the number of neighbors in the cluster and satisfies

$$\log_2 k = \sum_j G_{ij}, \tag{19}$$

where  $G_{ij}$  is the joint probability distribution between points  $x_i$  and  $x_j$ . The sum of the Euclidean distance between point  $i$  and all other points replaces the information entropy in t-SNE.

- (3) Using  $\rho_i$ , a locally similar data cluster is formed, and the joint probability distribution  $G_{ij}$  of any two origin-data points in the space is symmetrically calculated as

$$G_{ij} = G_{ij} + G_{ji} - G_{ij}G_{ji}. \tag{20}$$

- (4) Let the feature set  $U$  be the probability map modeled by  $D$  in low-dimensional space using  $1/(1+py^{2q})$  rather than the  $t$  distribution in t-SNE. The joint probability distribution  $Q_{ij}$  of any two features of  $N$  is then given as

$$Q_{ij} = \left[ 1 + p(y_i - y_j)^{2q} \right]^{-1}, \quad (21)$$

where  $p$  and  $q$  are the UMAP hyperparameters, whose default values are 1.93 and 0.79, respectively.

- (5) The cross-entropy  $C_E$  is the loss function and balances the local structure and the total structure. It is calculated as follows:

$$C_E = \sum_i \sum_j \left[ G_{ij} \lg \left( \frac{G_{ij}}{Q_{ij}} \right) + (1 - G_{ij}) \lg \left( \frac{1 - G_{ij}}{1 - Q_{ij}} \right) \right]. \quad (22)$$

- (6) The graph Laplace initializes the low-dimensional coordinates and uses a spectral embedding projection to a low-dimensional space. Stochastic gradient descent [31] is used to optimize, to calculate the gradient with a random subset, to improve the iterative minimum cross-entropy speed, and to calculate the optimal feature set  $U$ . The cross-entropy of the stochastic gradient descent optimization is

$$C_E = \sum_j \left\{ -G \lg Q(d_{ij}) + (1 - G) \lg [1 - Q(u_{ij})] \right\}, \quad (23)$$

where  $u_{ij}$  is the distance between characteristic data points.

## 4 Prediction Model of Photovoltaic Power

### 4.1 Gated Recurrent Unit Framework for Feature Transfer

Compared with the commonly used LSTM, the GRU framework offers the advantages of simple structure and fast calculations [32]. Nevertheless, the prediction model constructed from a traditional GRU neural network cannot predict the output of new PV power plants with sufficient accuracy. Therefore, this section constructs a GRU neural network based on TL (TL-GRU) with one input layer, one output layer, two hidden layers, and one FC layer. The hidden layer is initialized by using the source domain dataset and is shared with the target domain. Fig. 2 shows the structure.

The information transmission through a GRU cell is described by Eqs. (24)–(27):

*Gate control.* Let  $Z_t$  be the update gate of the GRU at time  $t$ , and let  $R_t$  be the reset gate:

$$Z_t = S(W_{zx}x_t + W_{zh}h_{t-1}), \quad (24)$$

$$R_t = S(W_{rx}x_t + W_{rh}h_{t-1}), \quad (25)$$

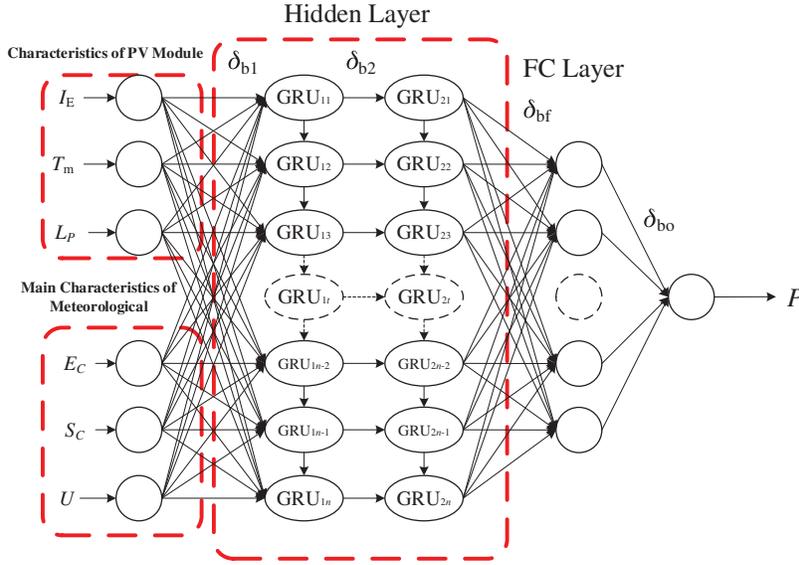
where  $S$  is the stimulus function sigmoid,  $x_t$  and  $h_{t-1}$  are the input at the current moment and the hidden state at the previous moment, respectively, and  $W_x$  and  $W_h$  are the weighting matrix of the input and loop connections, respectively.

*State retention.* Let  $h_t$  be the current hidden state of the GRU at time  $t$ , and let  $h'_t$  be the hidden state at the previous time:

$$h_t = (1 - Z_t) \odot h_{t-1} + Z_t \odot h'_t, \quad (26)$$

$$h'_t = \tanh(W_{h'x}x_t + R_t \odot U_{h'}h_{t-1}), \quad (27)$$

where  $\odot$  is the Hadamard product.



**Figure 2:** TL-GRU neural network

If the input layer of the neural network is  $F$ , then the hidden layers and the output  $P$  of the FC layer are

$$H_1 = S \left( \sum_j^m W_{1j} F_j + \delta_{b1} \right), \quad (28)$$

$$H_2 = S \left( \sum_j^m W_{2j} H_{1j} + \delta_{b2} \right), \quad (29)$$

$$P = W_o \left[ S \left( \sum_j^m W_{fj} H_{2j} + \delta_{bf} \right) \right] + \delta_{bo}, \quad (30)$$

where  $P$  is the normalized PV-power prediction (dimensionless),  $W_1$ ,  $W_2$ ,  $W_f$ , and  $W_o$  are the weighting matrices to transform the input layer to the first hidden layer, the first hidden layer to the second layer, the second hidden layer to the FC layer, and the FC layer to the output layer, respectively, and  $\delta_{b1}$ ,  $\delta_{b2}$ ,  $\delta_{bf}$ , and  $\delta_{bo}$  are the biases of the two hidden layers, the FC layer, and the output layer, respectively.

#### 4.2 Regularized Transfer Learning

When the prediction model learns the target domain, ‘‘catastrophic forgetting’’ is triggered, which leads to overfitting [33]. The FC layer serves as a ‘‘firewall’’ for TL in this section and is fine-tuned according to the target domain [34].

Let the source domain dataset  $D_S = [I_S, F_S]$ , the target domain  $D_T = [I_T, F_T]$ , and the prediction dataset  $D_P \subset D_T$ , where  $I$  and  $F$  are the label and feature sets, respectively. The minimum empirical loss function of the source domain is

$$L' = \operatorname{argmin}_f \frac{1}{n} \sum_{i=1}^n L(l, F, \delta, f), \quad (31)$$

$$\delta = [\delta_H, \delta_{SF}], \quad (32)$$

where  $\delta_H$  is the shared hidden layer parameter, and  $\delta_{SF}$  is the FC layer parameter of the source domain.

Fine-tuning the FC layer over the target domain is problematic because the model becomes excessively complex, and the training process leads to overfitting. We thus introduce a structural penalty on  $L'$  to constrain the model complexity:

$$L' = \operatorname{argmin}_f \left[ \frac{1}{n_s} \sum_{i=1}^{n_s} L(l, F, \delta, f) + \frac{\lambda}{2} \|W\|_2^2 \right], \quad (33)$$

where  $n_s$  is the number of source domain labels,  $\lambda$  is the L2 regularization-penalty coefficient, and  $W$  is the regular structural term of the weighting matrix.

### 4.3 Feature-Transfer Process

Fig. 3 summarizes the steps of the model for short-term prediction of PV power based on the fusion device feature-transfer (FFT-PV) model.

- (1) *Normalization.* To ensure that different features have the same gain as the mapping in the domain, we normalize the data to (0, 1):

$$X_N = \frac{X_A - X_{\min}}{X_{\max} - X_{\min}}, \quad (34)$$

where  $X_N$ ,  $X_A$ ,  $X_{\min}$ ,  $X_{\max}$  are the normalized variable, the actual value of the variable, the minimum value of the variable, and the maximum value of the variable, respectively. When the variable is the generated PV power, the characteristics of the PV power plant are considered, and the maximum power is the operating capacity of the PV power plant.

- (2) *Construction of transfer characteristics.* We use Eqs. (1)–(9) to calculate the characteristics of the PV modules, then calculate the Clear-Sky entropy  $E_C$  for 4-hour intervals, and finally construct the main meteorological characteristics together with the Clear-Sky entropy  $E_C$ , the seasonal characteristic coefficient  $S_C$ , and the weather feature set  $U$ . To summarize, we construct the following transfer characteristics  $F$  of the FFT-PV model:

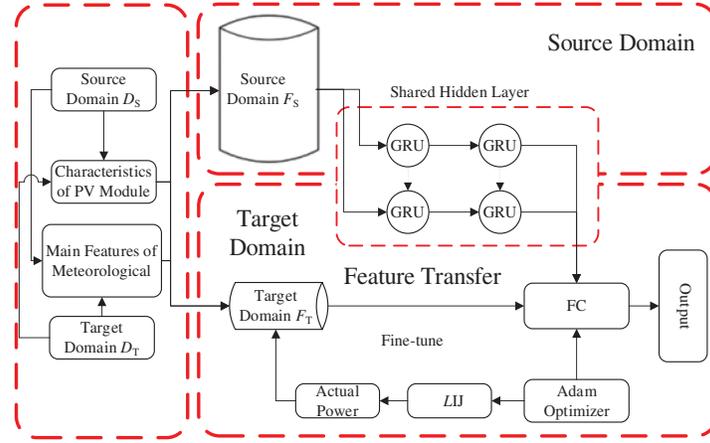
$$F = [I_E, T_m, L_P, E_C, S_C, U] \quad (35)$$

- (3) *Model pre-training.* To establish the TL-GR U framework, we use the source domain  $D_S = [I_S, F_S]$  to pre-train, randomly initialize the weighting matrix and bias, set the learning rate to 0.9, and apply the Adam optimization algorithm [35] to iteratively minimize the mean square error of the loss function.
- (4) *Model localization.* We use the source domain  $D_T = [I_T, F_T]$  to localize the PV module, the hidden layer is shared between the source domain and the target domain, the FC layer is fine-tuned by applying Eq. (30), regularization iteration is done according to Eq. (33), and the learning rate is 0.01.

## 5 Short-Term Prediction of Photovoltaic Power

To verify the performance of the FFT-PV model, we use the PV-power data from Pianguan, China [2014/07/26 to 2019/07/26] as the source domain to label  $l_s$  for 1826 days. The target domain dataset comes from the Desert Knowledge Australia Solar Centre, the power data of [2021/03/09 to 2021/04/09] is target domain A (labeled  $l_{TA}$ ), and the power data of [2021/02/26 to 2021/03/26] is target domain B (labeled  $l_{TB}$ ), and both are labeled  $l_T$  for 30 days. The weather dataset contains ten independent variables: GHI, POA irradiance, DNI, surface pressure, wind direction, wind speed, vertical wind speed, temperature, relative humidity, and rainfall. The time resolution of the power prediction is 15

min, which satisfies the requirements of the grid dispatching agency for short-term renewable-energy forecasting [36].



**Figure 3:** Model for predicting power of PV power plants

### 5.1 Evaluation of Model Performance

The Spearman correlation coefficient  $C_s$  represents the correlation between variables, the adjusted correlation coefficient  $R_A^2$  represents how input affects the model and offsets the contribution to the model of additional variables. The normalized root mean square error (NRMSE) gives the degree of dispersion in the normalized error, the normalized mean absolute percentage error (NMAPE) gives the average normalized percentage error, the normalized maximum absolute percentage error (NLAE) gives the maximum value of the normalized percentage error, and the qualified rate  $Q_R$  is the percent of predicted qualified points with respect to the total number of points in the evaluation period. These quantities are expressed as follows:

$$C_s = 1 - \frac{6 \sum_{i=1}^n \Delta_i^2}{n(n^2 - 1)}, \quad (36)$$

$$R_A^2 = 1 - \frac{(n-1) \sum_{i=1}^N (y_i - p_i)^2}{(n-k-1) \sum_{i=1}^N (y_i - \bar{y})^2} \times 100\%, \quad (37)$$

$$NRMSE = \frac{\sqrt{\sum_{i=1}^n (y_i - p_i)^2}}{\sqrt{\sum_{i=1}^n y_i^2}} \times 100\%, \quad (38)$$

$$NMAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - p_i|}{\frac{1}{n} \sum_{i=1}^n y_i} \times 100\%, \quad (39)$$

$$NLAE = \frac{1}{P_{\max}} |y_i - p_i|_{\max} \times 100\%, \quad (40)$$

$$Q_R = \frac{1}{n} \sum_{i=1}^n B_i \times 100\%, \quad B_i = \begin{cases} 1, & \frac{|y_i - p_i|}{P_N} \leq B \\ 0, & \frac{|y_i - p_i|}{P_N} > B, \end{cases} \quad (41)$$

where  $\Delta_i$  is the difference in grade of a variable (dimensionless),  $y_i$ ,  $p_i$ ,  $P_{\max}$ , and  $P_N$  are the actual value, predicted value, actual maximum power in the interval, and operating capacity (kW), respectively, and  $B$  is the qualified threshold, which is 6.5% in this section.

### 5.2 Evaluation of Model Feature

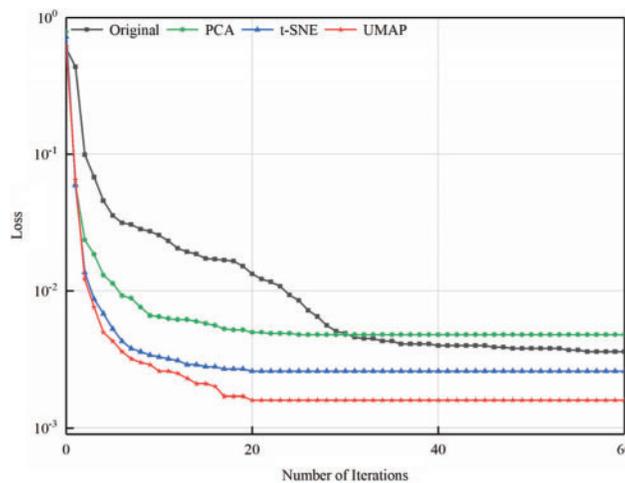
The ability of UMAP to extract features from the weather dataset is accurately evaluated and compared with the original dataset  $D$  and with the PCA and the t-SNE algorithms.

The global structure is evaluated by the coefficient  $C_S$  of the distance between the  $D$  data points and the distance between the feature set data points ( $C_{SD}$ ). The average value of the coefficient  $C_S$  between variables ( $C_{SV}$ ) evaluates the redundancy between features (see Table 2). The feature set  $PC$  exhibits a high redundancy between features, so the global structure is difficult to maintain.  $C_{SD} = 0.86$  for feature set  $U$  and  $C_{SV} = 0.36$ , which is significantly better than the feature set  $T$ . Therefore, UMAP offers the most balance in preserving the global structure.

**Table 2:** Feature extraction performance

Feature set	Algorithm	Dimension	Time/s	$C_{SD}$	$C_{SV}$	$R_A^2/\%$
$D$		10		1.00	0.12	72.43
$PC$	PCA	3	0.03	0.52	0.63	68.62
$T$	t-SNE	3	0.49	0.65	0.51	76.51
$U$	UMAP	3	0.27	0.86	0.36	85.26

The feature set iteratively trains the neural network, as shown in Fig. 4. The feature sets  $PC$ ,  $T$ , and  $U$  can accelerate the iteration speed, but the loss of  $PC$  information reduces the accuracy. Comparing the fitting degree of each feature set to the PV power by calculating  $R_A^2$  shows that feature sets  $T$  and  $U$  improve the fitting accuracy of features. For  $U$ ,  $R_A^2 = 85.26\%$ , and the calculation time of UMAP decreases by 8.75% with respect to t-SNE.



**Figure 4:** Iteration process for calculating loss

To summarize, the UMAP algorithm rapidly and accurately constructs the weather feature set to replace the original meteorological data.

### 5.3 Error Analysis

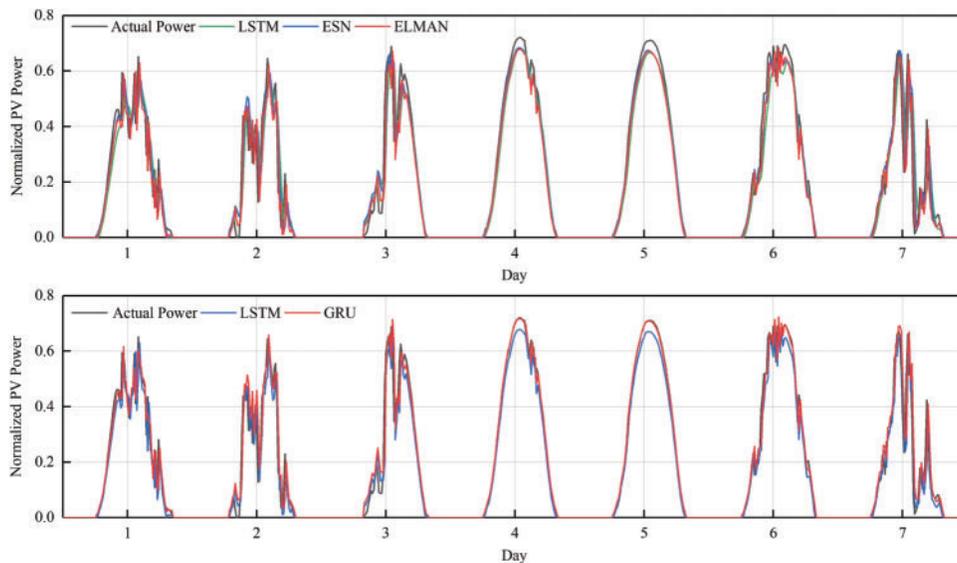
To test that the FFT-PV model can be easily generalized for use with different PV power stations, Section 5.3.1 evaluates the data mining ability of the prediction model trained by the source domain feature set  $F_s$ , and Section 5.3.2 transfers learning from the source domain model to target domains A and B and tests the adaptability of the FTL-PV model to different domains. Table 3 lists part of the metadata of each domain.

**Table 3:** Metadata of domains

Domain	Structure	Encapsulation	Tracker	On-grid time/year	Location
Source	Roof	Glass/cell/glass	Fixed	2016	39.33°N, 111.83°E
Target A	Rack	Glass/cell/polymer	Single	2008	23.42°S, 133.52°E
Target B	Rack	Glass/cell/polymer	Dual	2010	25.24°S, 130.99°E

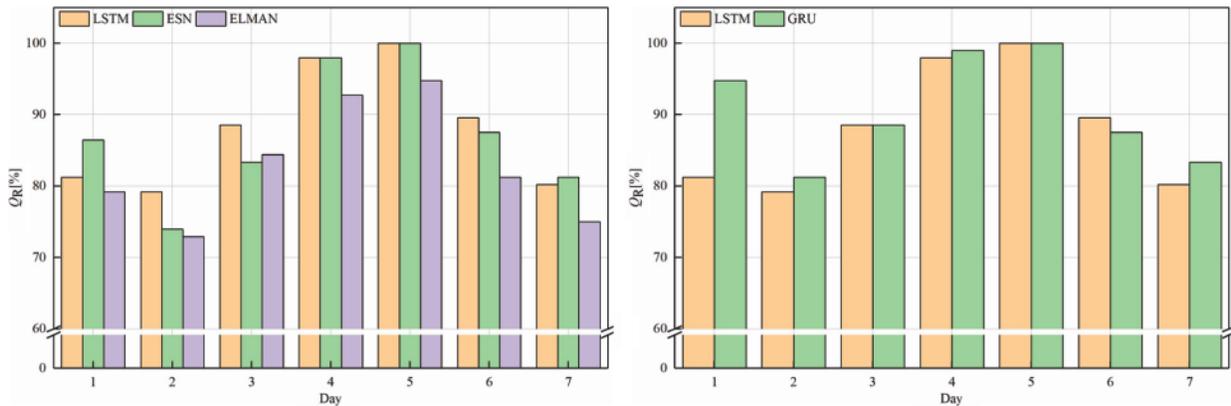
#### 5.3.1 Source Domain Error

The source domain model is trained with  $F_s$  and is based on a GRU neural network. Fig. 5 compares the PV power predicted by the source domain feature set  $F_s$  with that predicted by the Elman network, the echo state network (ESN), and the LSTM network. Fig. 6 shows  $Q_R$  for the source domain.



**Figure 5:** Comparison of source domain predictions

Fig. 6 shows that the time series-based LSTM and GRU framework produce a greater  $Q_R$ , with the simpler structure of the GRU producing the highest  $Q_R$ . Table 4 lists the prediction indicators of the source domain.



**Figure 6:**  $Q_R$  of source domain

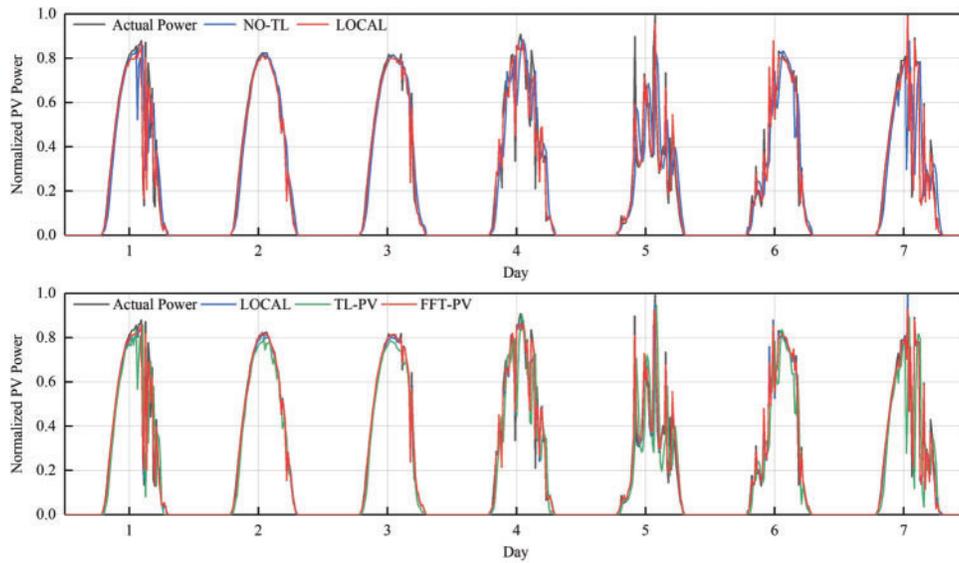
**Table 4:** Predicted indicators of source domain

Model	Indicator (%)	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Total
Elman	NRMSE	21.13	32.15	25.07	8.15	8.20	15.75	38.43	20.29
	NMAPE	19.39	31.06	20.23	8.21	8.05	14.05	33.88	17.16
	NLAE	19.51	26.66	27.86	7.99	9.05	18.59	37.83	37.83
ESN	NRMSE	18.75	28.15	22.41	6.01	5.60	11.75	31.85	17.09
	NMAPE	15.88	26.68	18.75	5.17	5.48	10.16	27.25	13.67
	NLAE	18.22	23.74	28.03	10.69	4.64	14.05	30.56	30.56
LSTM	NRMSE	20.64	23.92	19.98	7.47	6.67	12.57	25.93	15.84
	NMAPE	16.96	23.50	16.96	6.99	6.93	10.61	23.02	13.51
	NLAE	19.63	17.34	27.45	12.41	4.99	16.08	25.92	27.45
GRU	NRMSE	12.50	25.25	19.29	3.96	1.79	10.62	27.80	14.35
	NMAPE	9.79	23.35	15.55	2.78	1.66	8.29	24.95	10.42
	NLAE	15.56	20.47	26.93	8.71	2.72	13.14	24.29	26.93

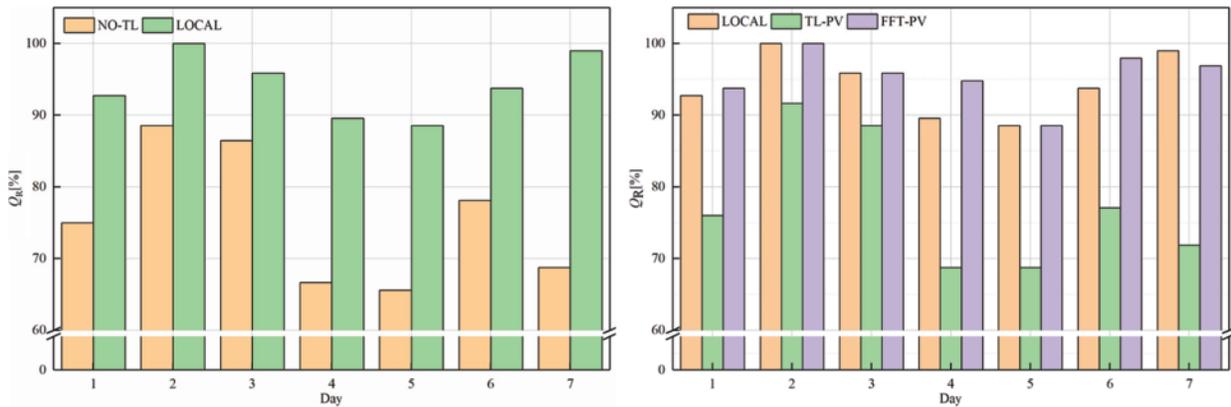
Overall, the indicators of the GRU framework in the source domain are significantly better than those of the other frameworks. On average, the indicators of the GRU framework are reduced by 3% for NRMSE, 4% for NMAPE, and 5% for NLAE. Therefore, the GRU network based on time series produces the most accurate PV-power forecasts of all frameworks studies herein.

### 5.3.2 Target Domain of Error

To verify the generalizability of feature transfer, TL is applied to the target domains A and B, and the FFT-PV model is compared with the no-TL model, the prediction model based on the local dataset (LOCAL), and the TL model without constructing the characteristics of the PV module (TL-PV) in Figs. 7–10 and Tables 5, 6 present the results for target domain A (B).

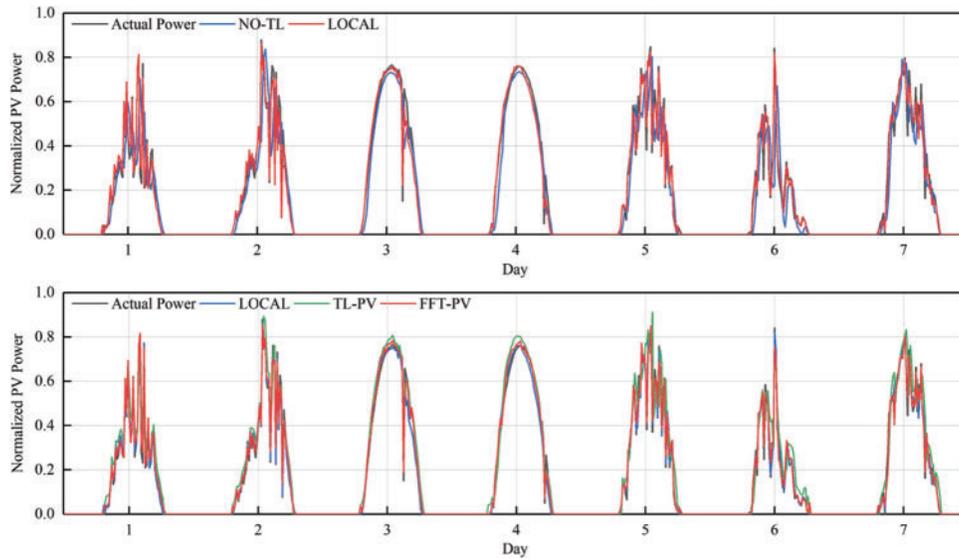


**Figure 7:** Comparison of PV-power predictions of the various models in target domain A

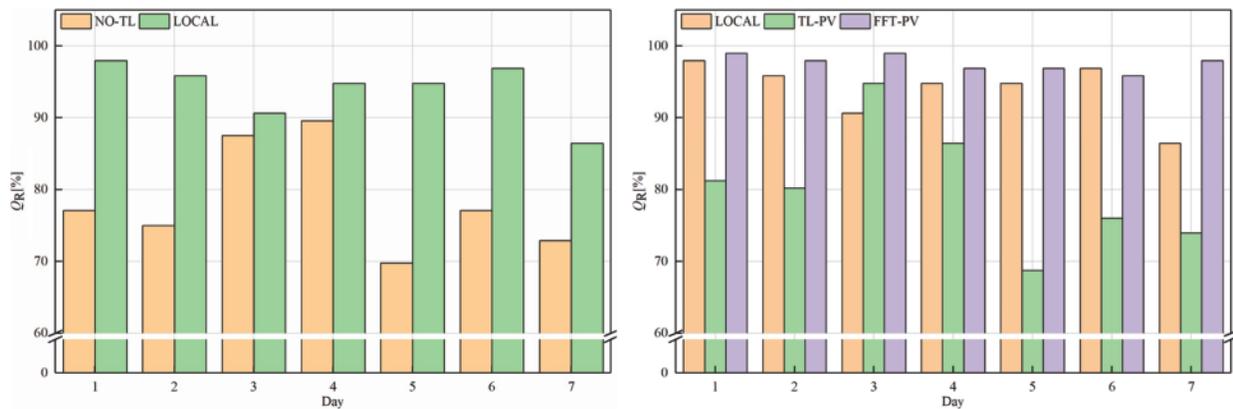


**Figure 8:**  $Q_R$  of target domain A

Comparing  $Q_R$  of the LOCAL model and no-TL model in target domain A shows that the qualification rate for the prediction model without TL decreases significantly, which prevents it from meeting the accuracy requirements. Moreover, the prediction model based on FFT-PV produces significantly more accurate predictions of PV power than does the traditional TL-PV model, and the prediction accuracy of the local model is similar to FFT-PV when given sufficient data.



**Figure 9:** Comparison of PV-power predictions of the various models in target domain B



**Figure 10:**  $Q_R$  of target domain B

As shown in Table 5, the indicators of FFT-PV are significantly better than those of the no-TL and TL-PV models, and the accuracy is equivalent to that of the LOCAL model. Compared with the no-TL model, the NRMSE, NMAPE, and NLAE of the FFT-PV model decrease by 23.11%, 18.17%, and 39.26%, respectively. Compared with the TL-PV model, the NRMSE, NMAPE, and NLAE of the FFT-PV model decrease by 24.29%, 18.35%, and 55.53%, respectively.

For target domain B, compared with the no-TL model, the NRMSE, NMAPE, and NLAE of the FFT-PV model decrease by 24.43%, 19.28%, and 53.06%, respectively. Compared with the TL-PV model, the NRMSE, NMAPE, and NLAE of the FFT-PV model decrease by 14.78%, 12.65%, and 44.87%, respectively. Fig. 10 shows that unstable weather conditions amplify the advantages of the FFT-PV model.

Like target domain A, the FFT-PV model of target domain B increases the prediction accuracy with respect to the no-TL and TL-PV models and suppresses overfitting.

**Table 5:** Prediction indicators of target domain A

Model	Indicator (%)	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Total
LOCAL	NRMSE	7.80	3.53	6.01	9.30	20.31	7.80	6.49	8.95
	NMAPE	7.03	3.36	4.40	7.77	14.60	6.08	5.12	6.55
	NLAE	14.74	4.86	12.33	15.06	31.86	14.76	16.23	31.86
NO-TL	NRMSE	33.15	9.73	12.10	29.33	49.76	25.60	43.56	29.94
	NMAPE	24.56	7.79	10.00	25.55	45.15	21.83	35.56	22.98
	NLAE	60.92	21.99	24.16	53.04	61.76	41.19	62.91	62.91
TL-PV	NRMSE	35.43	8.85	11.17	31.03	52.05	25.99	44.79	31.12
	NMAPE	26.20	8.09	9.72	26.27	46.11	21.37	33.89	23.16
	NLAE	79.17	15.68	25.26	51.45	59.11	36.72	77.36	79.18
FFT-PV	NRMSE	6.35	2.81	5.38	6.90	14.10	5.92	6.17	6.83
	NMAPE	4.87	2.26	3.21	5.24	12.16	3.75	4.44	4.81
	NLAE	12.86	5.22	12.22	13.41	23.65	14.64	16.00	23.65

**Table 6:** Prediction indicators of target domain B

Model	Indicator (%)	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Total
LOCAL	NRMSE	8.72	7.82	8.65	7.31	7.94	11.43	12.13	9.04
	NMAPE	9.17	7.26	7.52	6.75	6.99	10.63	9.91	8.08
	NLAE	6.49	9.47	13.25	10.72	10.35	14.18	15.41	15.41
NO-TL	NRMSE	48.28	40.60	18.18	13.71	32.81	49.44	24.13	30.46
	NMAPE	40.33	33.10	11.87	10.80	29.78	42.80	22.68	24.63
	NLAE	42.97	45.85	48.33	23.29	43.53	62.30	26.09	62.30
TL-PV	NRMSE	22.17	17.40	9.29	8.96	33.22	30.13	24.22	20.81
	NMAPE	22.06	17.24	8.22	8.38	27.67	30.69	22.41	18.00
	NLAE	23.86	19.92	19.70	10.08	54.11	27.09	28.19	54.11
FFT-PV	NRMSE	6.56	6.91	4.67	5.23	6.36	9.00	5.69	6.03
	NMAPE	6.04	6.58	4.15	4.35	5.67	7.88	4.67	5.35
	NLAE	6.29	7.83	6.73	7.76	6.86	8.58	9.24	9.24

Synthesizing target domains A and B shows that the LOCAL model uses sufficient local data for training, so the prediction accuracy is high. The no-TL model uses the target domain data to directly train the source domain model and overfits the learning of small samples, so the prediction accuracy is the worst of all models studied. The TL-PV model is not a physically accurate model, so its results are not much better than those of the no-TL model. The fluctuation range of the FFT-PV model is

closest to the true range of fluctuations, and the recognition accuracy and real-time performance of the features are also better than those of the other solutions.

To summarize, the FFT-PV model based on the GRU framework proposed herein not only can be applied to different data domains to overcome the problem of insufficient local training data for newly built PV power plants but also can be widely generalized.

## 6 Conclusions

Based on open and shared power data, we develop herein a model to make short-term predictions of PV power based on fusion device feature transfer. The results lead to the following main conclusions:

- The feature extraction based on UMAP is clearly superior to that based on the PCA and t-SNE algorithms for processing weather datasets and accelerates the projection by balancing the global structure with the local structure. The weather feature constructed by the UMAP algorithm increases the iteration speed of the neural network.
- Taking the PV power data of Pianguan, China as the source domain label, the GRU framework with the simple structure prediction model produces an average reduction of 3%, 4%, and 5% for the NRMSE, NMAPE, and NLAE with respect to the other frameworks.
- Taking the PV power data of the Desert Knowledge Australia Solar Centre as the target domain label, we establish a fine-tuned FC layer by sharing the hidden layer and introducing regularization constraints in the feature transfer. The FFT-PV produces an average reduction of 15%, 12%, and 35% for the NRMSE, NMAPE, and NLAE with respect to the other methods, and overfitting is suppressed for target-domain training.
- The innovation introduced by this study is the integration of transfer learning with the physical characteristics of PV-power plants. By establishing high-precision models for different PV devices and local meteorological models, and fine-tuning network parameters according to structural constraints and PV-plant features, we resolve the problems whereby new PV-power stations have insufficient data and trained prediction models are difficult to generalize for use on different PV-power stations. The proposed model enables new grid-connected PV-power plants to benefit from high-precision PV-power prediction, which increases the use of PV power and thereby contributes significantly to the goal of carbon peaking and carbon neutralization.

**Acknowledgement:** This research was supported by the National Natural Science Foundation of China (No. 6180802161), the Educational Commission of Liaoning Province of China (No. JZL201915401). We thank TopEdit ([www.topedit.com](http://www.topedit.com)) for its linguistic assistance during the preparation of this manuscript.

**Funding Statement:** The authors received no specific funding for this study.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

1. Li, Z., Chen, S., Dong, W., Liu, P., Du, E. et al. (2021). Low carbon transition pathway of power sector under carbon emission constraints. *Proceedings of the CSEE*, 41(12), 3987–4001. DOI 10.13334/j.0258-8013.pcsee.210671.

2. Faraji, J., Ketabi, A., Hashemi-Dezaki, H., Shafie-Khah, M., Catalao, J. P. (2020). Optimal day-ahead scheduling and operation of the prosumer by considering corrective actions based on very short-term load forecasting. *IEEE Access*, 8, 83561–83582. DOI 10.1109/ACCESS.2020.2991482.
3. Das, U. K., Tey, K. S., Seyedmahmoudian, M., Mekhilef, S., Idris, M. Y. I. et al. (2018). Forecasting of photovoltaic power generation and model optimization: A review. *Renewable & Sustainable Energy Reviews*, 81, 912–928. DOI 10.1016/j.rser.2017.08.017.
4. Ahmed, R., Sreeram, V., Mishra, Y., Arif, M. D. (2020). A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. *Renewable and Sustainable Energy Reviews*, 124, 109792. DOI 10.1016/j.rser.2020.109792.
5. Faraji, J., Abazari, A., Babaei, M., Muyeen, S. M., Benbouzid, M. (2020). Day-ahead optimization of prosumer considering battery depreciation and weather prediction for renewable energy sources. *Applied Sciences*, 10(8), 2774. DOI 10.3390/app10082774.
6. Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-de-Pison, F. J. et al. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136, 78–111. DOI 10.1016/j.solener.2016.06.069.
7. Niu, D., Wang, K., Sun, L., Wu, J., Xu, X. (2020). Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study. *Applied Soft Computing*, 93, 106389. DOI 10.1016/j.asoc.2020.106389.
8. Liu, L., Zhao, Y., Chang, D., Xie, J., Ma, Z. et al. (2018). Prediction of short-term PV power output and uncertainty analysis. *Applied Energy*, 228, 700–711. DOI 10.1016/j.apenergy.2018.06.112.
9. Massaoudi, M., Chihi, I., Abu-Rub, H., Refaat, S. S., Oueslati, F. S. (2021). Convergence of photovoltaic power forecasting and deep learning: State-of-art review. *IEEE Access*, 9, 136593–136615. DOI 10.1109/ACCESS.2021.3117004.
10. Chu, Y., Urquhart, B., Gohari, S. M., Pedro, H. T., Kleissl, J. et al. (2015). Short-term reforecasting of power output from a 48 MWe solar PV plant. *Solar Energy*, 112, 68–77. DOI 10.1016/j.solener.2014.11.017.
11. Nie, Y., Sun, Y., Chen, Y., Orsini, R., Brandt, A. (2020). PV power output prediction from sky images using convolutional neural network: The comparison of sky-condition-specific sub-models and an end-to-end model. *Journal of Renewable and Sustainable Energy*, 12(4), 46101. DOI 10.1063/5.0014016.
12. Faraji, J., Ketabi, A., Hashemi-Dezaki, H., Shafie-Khah, M., Catalão, J. P. (2020). Optimal day-ahead self-scheduling and operation of prosumer microgrids using hybrid machine learning-based weather and load forecasting. *IEEE Access*, 8, 157284–157305. DOI 10.1109/ACCESS.2020.3019562.
13. Wang, F., Xuan, Z., Zhen, Z., Li, K., Wang, T. et al. (2020). A Day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. *Energy Conversion and Management*, 212, 112766. DOI 10.1016/j.enconman.2020.112766.
14. Gao, M., Li, J., Hong, F., Long, D. (2019). Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. *Energy*, 187, 115838. DOI 10.1016/j.energy.2019.07.168.
15. Adar, M., Najih, Y., Gouskir, M., Chebak, A., Mabrouki, M. et al. (2020). Three PV plants performance analysis using the principal component analysis method. *Energy*, 207, 118315. DOI 10.1016/j.energy.2020.118315.
16. Lee, H., Lee, J. G., Kim, N. W., Lee, B. T. (2021). Model-agnostic online forecasting for PV power output. *IET Renewable Power Generation*, 15(15), 3539–3551. DOI 10.1049/rpg2.12243.
17. Wang, F., Zhang, Z., Liu, C., Yu, Y., Pang, S. et al. (2019). Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting. *Energy Conversion and Management*, 181, 443–462. DOI 10.1016/j.enconman.2018.11.074.
18. Zhao, Y., Xia, S., Zhang, J., Hu, Y., Wu, M. (2021). Effect of the digital transformation of power system on renewable energy utilization in China. *IEEE Access*, 9, 96201–96209. DOI 10.1109/ACCESS.2021.3094317.
19. Dolara, A., Leva, S., Manzolini, G. (2015). Comparison of different physical models for PV power output prediction. *Solar Energy*, 119, 83–99. DOI 10.1016/j.solener.2015.06.017.

20. Zhang, L., Chen, Z., Zhang, H., Ma, Z., Cao, B. et al. (2020). Accurate study and evaluation of small PV power generation system based on specific geographical location. *Energy Engineering*, 117(6), 453–470. DOI 10.32604/EE.2020.013276.
21. Yuan, W., Ji, J., Li, Z., Zhou, F., Ren, X. et al. (2018). Comparison study of the performance of two kinds of photovoltaic/thermal(PV/T) systems and a PV module at high ambient temperature. *Energy*, 148, 1153–1161. DOI 10.1016/j.energy.2018.01.121.
22. Kurnik, J., Jankovec, M., Brecl, K., Topic, M. (2011). Outdoor testing of PV module temperature and performance under different mounting and operational conditions. *Solar Energy Materials and Solar Cells*, 95(1), 373–376. DOI 10.1016/j.solmat.2010.04.022.
23. Walwil, H. M., Mukhaimer, A., Al-Sulaiman, F. A., Said, S. A. (2017). Comparative studies of encapsulation and glass surface modification impacts on PV performance in a desert climate. *Solar Energy*, 142, 288–298. DOI 10.1016/j.solener.2016.12.020.
24. Holmgren, W. F., Lorenzo, A. T., Hansen, C. (2017). A comparison of PV power forecasts using PVLlib-Python. *IEEE 44th Photovoltaic Specialist Conference (PVSC)*, pp. 1127–1131. Washington DC, USA. DOI 10.1109/PVSC.2017.8366724.
25. Olalla, C., Hasan, M., Deline, C., Maksimović, D. (2018). Mitigation of hot-spots in photovoltaic systems using distributed power electronics. *Energies*, 11(4), 726. DOI 10.3390/en11040726.
26. Jordan, D. C., Silverman, T. J., Sekulic, B., Kurtz, S. R. (2017). PV degradation curves: Non-linearities and failure modes. *Progress in Photovoltaics*, 25(7), 583–591. DOI 10.1002/pip.2835.
27. Bhallamudi, R., Kumarasamy, S., Sundarabalan, C. K. (2021). Effect of dust and shadow on performance of solar photovoltaic modules: Experimental analysis. *Energy Engineering*, 118(6), 1827–1838. DOI 10.32604/EE.2021.016798.
28. Maitanova, N., Telle, J. S., Hanke, B., Grottke, M., Schmidt, T. et al. (2020). A machine learning approach to Low-cost photovoltaic power prediction based on publicly available weather reports. *Energies*, 13(3), 735. DOI 10.3390/en13030735.
29. Singh, P., Dhiman, G. (2018). Uncertainty representation using fuzzy-entropy approach: Special application in remotely sensed high-resolution satellite images (RSHRSIs). *Applied Soft Computing*, 72, 121–139. DOI 10.1016/j.asoc.2018.07.038.
30. McInnes, L., Healy, J., Melville, J. (2020). UMAP: Uniform manifold approximation and projection for dimension reduction. <https://arxiv.org/abs/1802.03426>.
31. Keskar, N. S., Socher, R. (2017). Improving generalization performance by switching from adam to sgd. <https://arxiv.org/abs/1712.07628>.
32. Gao, S., Huang, Y., Zhang, S., Han, J., Wang, G. et al. (2020). Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation. *Journal of Hydrology*, 589, 125188. DOI 10.1016/j.jhydrol.2020.125188.
33. Ghazi, M. M., Yanikoglu, B., Aptoula, E. (2017). Plant identification using deep neural networks via optimization of transfer learning parameters. *Neurocomputing*, 235, 228–235. DOI 10.1016/j.neucom.2017.01.018.
34. Tan, Y., Zhao, G. (2020). Transfer learning with long short-term memory network for state-of-health prediction of lithium-ion batteries. *IEEE Transactions on Industrial Electronics*, 67(10), 8723–8731. DOI 10.1109/TIE.2019.2946551.
35. Iduka, H., Kobayashi, Y. (2020). Training deep neural networks using conjugate gradient-like methods. *Electronics*, 9(11), 1809. DOI 10.3390/electronics9111809.
36. Raza, M. Q., Nadarajah, M., Ekanayake, C. (2016). On recent advances in PV output power forecast. *Solar Energy*, 136, 125–144. DOI 10.1016/j.solener.2016.06.073.