This article has been accepted for publication in a future issue of this journal, but has not been edited. Content will change prior to final publication.

Digital Twin Empowered PV Power Prediction

Xiaoyu Zhang, Yushuai Li, *Member, IEEE,* Tianyi Li, *Member, IEEE,* Yonghao Gui, *Member, IEEE,* Qiuye Sun, *Senior Member, IEEE,* David Wenzhong Gao, *Fellow, IEEE*

***Abstract*—The accurate prediction of photovoltaic (PV) power generation is significant to ensure the economic and safe operation of power systems. To this end, the paper proposes a new digital twin (DT) empowered PV power prediction framework that is capable of ensuring reliable data transmission and employing the DT to achieve high accuracy of power prediction. With this framework, considering potential data contamination in the collected PV data, a generative adversarial network is employed to restore the historical data set, which offers a prerequisite to ensure accurate mapping from the physical space to the digital space. Further, a new DT empowered PV power prediction method is proposed. Therein, we model a DT that encompasses a digital physical model for reflecting the physical operation mechanism and a neural network model (i.e., a parallel network of convolution and bidirectional long-short-term memory model) for capturing the hidden spatial-temporal features. The proposed method enables the use of the DT to take advantages of the digital physical model and the neural network model, resulting in enhanced prediction accuracy. Finally, a real data set is conducted to access the effectiveness of the proposed method.**

***Index Terms*—Photovoltaic Power Prediction; Digital Twin; Hybrid Prediction Method; Data Recovery.**

I. Introduction

With the increasing integration of PV power generation, its nonlinearity, periodicity, and volatility pose great challenges to the stable operation of power systems. The uncertain of the PV power generation and the randomness of the energy demand may lead to imbalance between the energy supply and demand. Accurate prediction models can mitigate the impacts of uncertainty of PV power generation, improve power system stability, and reduce the maintenance costs of additional equipments [1]–[3].

Currently, several studies on PV power prediction have been proposed, which can be roughly divided into three categories: 1) the physical methods; 2) the statistical methods; and 3) the artificial intelligence (AI)-based methods. The concept of physical methods is to use physical models to construct the relationship between PV power output and other factors, such as numerical weather prediction (NWP) data [4], sky images [5], and satellite images [6]. The concept of statistical methods is to apply statistical principles to extract correlations and variation patterns from historical data, such as Bayesian model averaging (BMA) [7], exponential smoothing [8], and autoregressive integrated moving average (ARIMA) [9]. Both physical and statistical methods have the advantage on obtaining stable PV power prediction. However, it is very difficult to establish a physical model that can obtain high accuracy prediction results for every prediction scenario, since there exist several hidden features that are hard to capture via mechanism analysis. Meanwhile, statistical methods mainly focus on using historical data of power generation, which ignores weather conditions. It results in limited prediction accuracy [10].

To cope with shortcomings of physical and statistical methods, the AI-based methods for PV power generation have been proposed and gained significant attentions. For instance, Convolutional Neural Networks (CNNs) were used for extracting spatial features [11]–[13], while Long Short-Term Memory (LSTM) networks were used for extracting temporal features [14]–[16]. CNNs do not fully consider the temporal characteristics of the input data, and LSTM networks have limited ability to capture causal relationships between input factors. To address this issue, hybrid models based on CNN and LSTM were proposed in [17], [18]. Additionally, Graph Neural Networks (GNNs) [19], [20], particularly Graph Convolutional Networks (GCNs) [21], [22], are often combined with graph modeling methods to explore causal relationships among input factors [23]. Nevertheless, GNNs and GCNs primarily focus on the neighboring information of nodes and have limited modeling capabilities for time-series data, which may pose challenges when dealing with graphs with complex topological structures. Recently, Generative Adversarial Networks (GANs) with capabilities in image restoration and data completion, have also been used to address PV power prediction problems. In [24], a generator based on Recurrent Neural Network (RNN) was employed to predict solar power, while a CNN discriminator was utilized to enhance the prediction accuracy of the generator. However, when training data is imbalanced or samples are scarce, it may lead to unreliable power prediction results generated by a GAN.

The aforementioned PV prediction models are built up on the assumption that the dataset is complete [11]–[24]. In fact, varying degrees of pollution are usually observed in the collected measurement data, which may be caused by data logger failures, communication network failures, and inaccurate instruments, etc. Learning samples with these unexpected pollutants may lead to bias in prediction results. To address this issue, the PV power generation is predicted based on a recursive long and short-term memory network in [25], which considers the possible quality problems of the dataset. The

Manuscript received: May 24, 2023; revised: August 19, 2023; accepted: October 15, 2023. Date of CrossCheck: October 15, 2023. Date of online publication: XX XX, XXXX.

This work is supported by European Horizon 2020 Marie Sklodowska-Curie Actions under Grant 101023244.

X. Zhang and Q. Sun, are with the School of Information Science and Engineering, Northeastern University, Shenyang, Liaoning, 110004, China.(e-mail:2100689@stu.neu.edu.cn).

Y. Li (corresponding author) is with the Department of Informatics, University of Oslo, Oslo 0316, Norway (e-mail: yushuaili@ieee.org).

T. Li is with the Department of Computer Science, Aalborg University, 9220 Aalborg, Denmark (e-mail: tianyi@cs.aau.dk).

Y. Gui is with the Electrification and Energy Infrastructures Division, Oak Ridge National Laboratory, USA (e-mail: guiy@ornl.gov).

D. W. Gao is with the Department of Electrical and Computer Engineering, University of Denver, Denver, CO 80208 USA (e-mail: Wenzhong.Gao@du.edu).

DOI: 10.35833/MPCE.2023.000351

This article has been accepted for publication in a future issue of this journal, but has not been edited. Content will change prior to final publication.

missing data is estimated using a recursive process. However, the robustness of the method is reduced when the testing data loss rate significantly differs from the training data loss rate. Moreover, this approach does not consider the continuous missing data patterns in the data-set. In [26], a super-resolution perception convolutional neural network was employed to recover missing data, and a stochastic configuration network (SCN) was utilized for PV power generation prediction. However, the quality of data recovery needs improvement, which subsequently affects the accuracy of PV power prediction. In addition, even if the data set is complete, there may exist data imbalance. In [27], data augmentation methods, e.g. noise injection, color space transformations, and mixing of images, were used to expand a small amount of sky image data under cloudy conditions. Meanwhile, the CNN was used to predict short-term PV output. In [28], the dataset was augmented with complementary exogenous features, including the periodic properties of the production, altitude, azimuth, and irradiance of solar, and clear and overcast days, etc. Then, a hybrid neural network model was proposed to predict PV power generation.

The aforementioned AI-based methods have yielded remarkable outcomes. However, there exist two challenges. On the one hand, although the consideration of data recovery is presented in [25] and [26], the implementation of these methods is challenging; meanwhile, the quality of the recovered data is insufficient. To address this issue, a potential approach is the utilization of the GAN. It is a framework for training parameter generation models, which is capable of learning arbitrarily complex probability distributions. The success of GANs in image restoration [29] and traffic data completion [30] serves as inspiration for applying GANs to learn the distribution of PV data, thus tackling the challenging task of recovering large-scale historical data. On the other hand, these AI-based methods [11]–[26] are predominantly developed using historical data, such as power generation and meteorology data, without taking into account the specific physical characteristics of the PV system itself. It should be noted that the actual state of the PV power station, particularly the physical condition of the PV panels, significantly impacts the power generation process. To address this issue, the DT technology provides an alternative solution. The DT refers to the construction of a virtual system in a virtual space that utilizes physical models and operational historical data to accurately represent and map the physical entity or process [31]. The advantages of DTs can be divided into the following three points: 1) The DT is an accurate virtual simulation of a real world entity, process, or system, which allows us to perform various tests, predictions, and optimizations in a virtual environment without actually manipulating real world objects. It results in saved time and money [32]; 2) The DT is capable of sharing information with the physical entities in real time, resulting in the synchronization of information. This is helpful to make fast and accurate decision making [33]; 3) The DT can leverage the digital physical models to describe the behavior of real systems, and combine with data-driven machine learning methods to achieve accurate modeling and prediction of real systems [34]. These advantages make DT be an innovative method and tool that can be applied to multiple fields. For instance, a two-level hierarchical learning process using the real-time model state stored on the DT server was proposed in [35], which aims to enhance the ML-based product design on a DT-aided IoT platform. An intelligent context-aware medical system was implemented in [36] by using a DT-based framework. Meanwhile, an electrocardiogram (ECG) heart rhythms classifier model was built by using machine learning to diagnose heart disease and detect heart problems. In addition, the DT was also used for product quality prediction [37], intelligent transportation [38], and smart home [39]. Although the DT has gained broad applications, it has not be applied to PV power prediction. Based on the advantages of the DT, our aim is to jointly create the digital physical model to reflect the inherent mechanism of PVs and use the neural networks to capture the hidden features that are hard to be modeled by physical model. In the sense, we can create a high-fidelity DT to reflect the reality well by taking advantages of physical knowledge and learned data knowledge, resulting in enhanced prediction accuracy. However, no attention has so far been paid to this aspect.

To tackle those challenges, the paper proposes the DT empowered framework, model, and method for PV power generation prediction. The main contributions are as follows:

1) We propose a novel DT empowered PV power prediction framework, composed of a physical layer, a data transmission layer, a DT layer, and a service layer, while defining the detailed functionality of each layer. This is a universal reference framework that enables the integration of the DT to empower the PV power prediction.

2) To ensure accurate mapping from the physical to the digital space, a GAN is employed to restore the historical dataset, considering potential data contamination in the collected PV data. This restoration process serves as a prerequisite for reliable data analysis and prediction within the DT framework.

3) A DT empowered PV power prediction method is proposed, where the DT is constructed with a digital physical model and a parallel network of convolution and bidirectional long-short-term memory (CNN-BiLSTM) model. This proposed method captures both the physical operation mechanism and hidden spatial-temporal features, leveraging the strengths of both models to increase prediction accuracy.

The remainder is summarized as follows. Section II presents the DT empowered PV power prediction framework. Section III provides the DT empowered prediction method within the proposed framework. Section IV presents the simulations to evaluate the performance of the proposed method. Finally, Section V concludes the paper.

II. DT Empowered PV Power Prediction Framework

Fig. 1 shows that the proposed DT empowered PV power prediction framework, which is composed of a physical layer, a data transmission layer, a DT layer, and a service layer.

1) Physical layer: It refers to physical objects in the real world, such as PV panels and sensors. The layer will collect and store device parameters, the PV power generation data and the meteorological data. Device parameters include short-circuit current I_{SC}, open-circuit voltage U_{OC}, data at the

This article has been accepted for publication in a future issue of this journal, but has not been edited. Content will change prior to final publication.

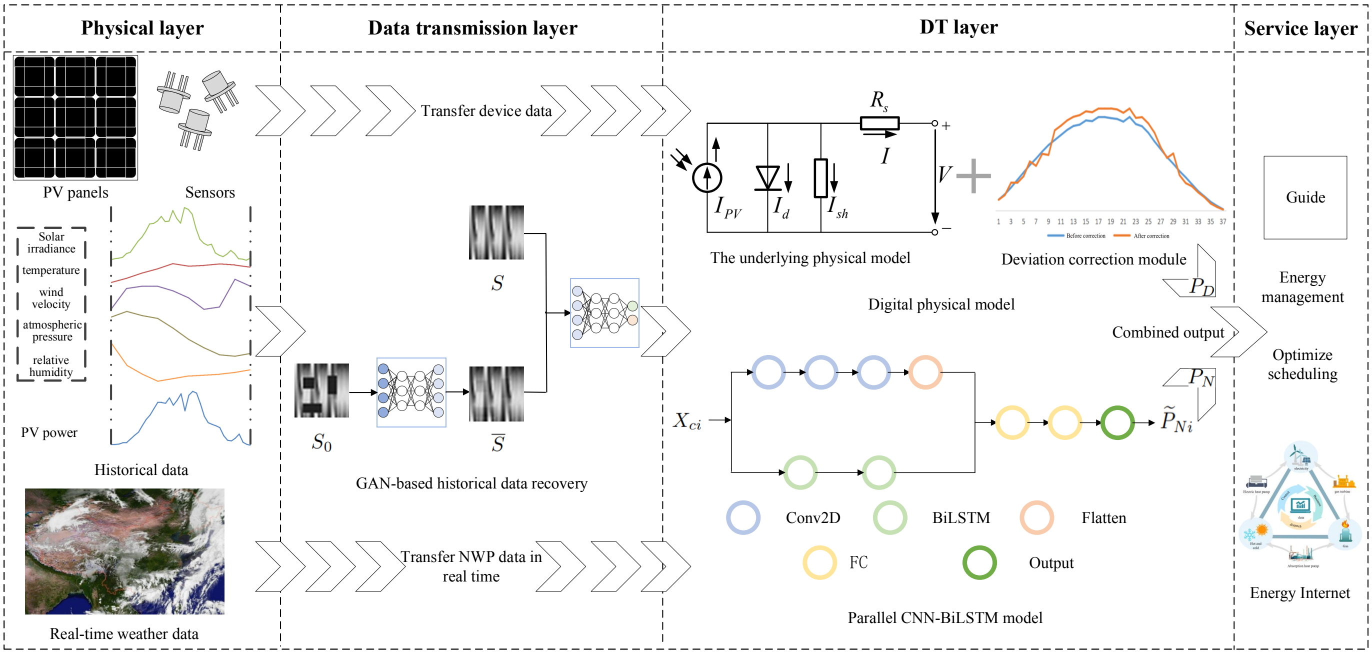

Fig. 1. PV power prediction framework.

maximum power point (current I_m, voltage U_m, and maximum power P_{0m}), and volt-ampere characteristic curve of the PV cell. According to different sampling time points, historical datasets can be expressed as:

$$D_{xp} = [D_{xp}(1), D_{xp}(2), \cdots, D_{xp}(\mathrm{n}_T)]^{\mathrm{T}}, \tag{1}$$

$$X = [X(1), X(2), \cdots, X(n_T)]^{\mathrm{T}}, \tag{2}$$

$$P = [P(1), \mathrm{P}(2), \cdots, \mathrm{P}(n_T)]^{\mathrm{T}}, \tag{3}$$

where $D_{xp}(j)$ represents historical data, including temperature, wind speed, solar radiation, relative humidity, and PV power generation, etc., collected at the jth sampling; $D_{xp}(j) = \{X(j), P(j)\}$; n_T represents the time dimension of the data; $X(j)$ and $P(j)$ represent the historical meteorological data and historical PV power generation data collected at the jth sampling time point, respectively.

2) Data transmission layer: This layer serves as the connection channel between the physical and virtual spaces, enabling the collection and transmission of relevant data information from the PV power station. During data collection, the loss of data packets is possible, leading to incomplete time series data in the analysis of historical PV power generation data. To address this issue, we propose the utilization of a GAN for data recovery, which will be discussed in Section III-A. The historical data restored by GAN and the parameter data of PV panels, sensors, and other devices are transmitted from the physical space to the virtual space at one time, participating in the construction of the DT model of the PV power station. The real-time weather data is transmitted in real-time from the physical space to the virtual space, which enables participating in the power prediction of the DT layer.

3) DT layer: As the main part of this paper, this layer focuses on creating the DT model and using it to achieve PV power generation. In order to accurately reflect the real world and create a high-fidelity DT model, it is necessary to consider the physical characteristics of the PV system and extract the inherent relationships within the historical data simultaneously. In the virtual space, we set up a digital physical model that can reflect the physical operation mechanism and a parallel CNN-BiLSTM model to capture hidden temporal and spatial features. These components are combined using a fusion formula to accomplish the prediction of PV power. The detailed DT modeling process and the prediction procedure will be discussed in Section III-B and Section III-C, respectively.

4) Service layer: This layer receives the prediction results from the DT layer to meet diverse services, such as: 1) providing reference for energy dispatch and optimization; 2) optimizing the charging/discharging control for battery; and 3) facilitating demand response programs.

III. DT Empowered PV Power Prediction Method

Within the proposed framework, we propose the DT empowered prediction method that contains three phases: 1) data preparation phase, 2) DT modeling phase; and 3) power prediction phase. Fig. 2 illustrates the overall flowchart. Next, we proceed to elaborate the design of each phase.

A. Data Preparation Phase

The data preparation phase is performed within the data transmission layer. In this phase, the data transmission layer retrieves pertinent data from the PV power station. The historical meteorological and power data are fed into the GAN. Subsequently, the recovered historical data and device parameters are transferred from the physical layer to the DT layer at one time to participate in the construction of DT model. Real-time weather data is transmitted from the physical layer to the DT layer, which is used for subsequent PV power prediction.

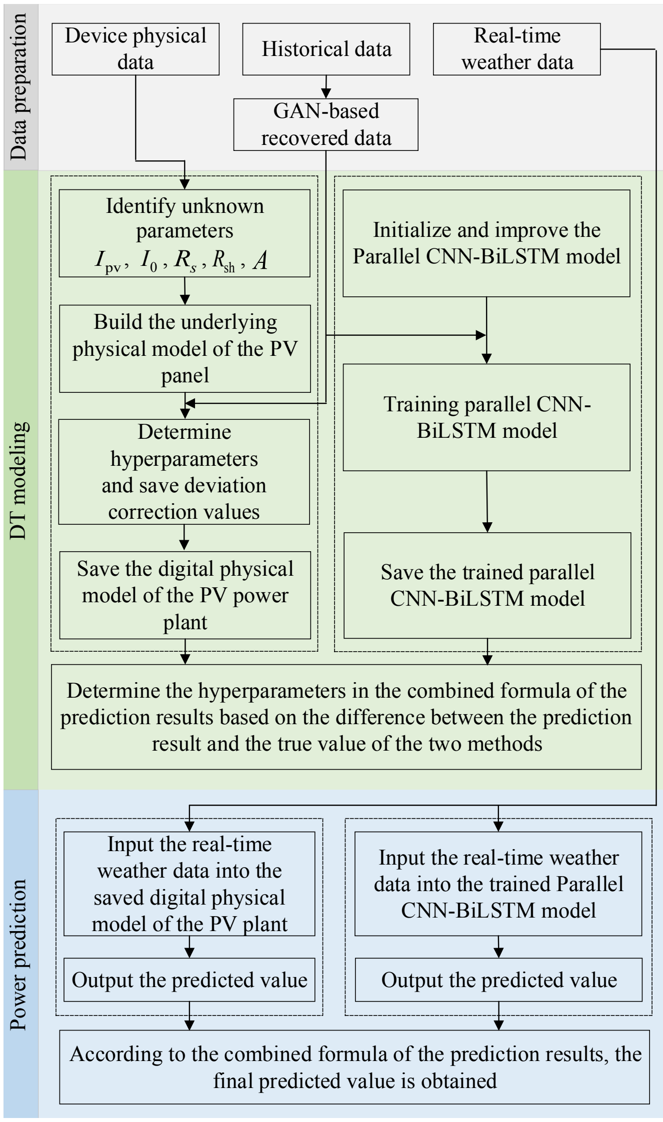

Fig. 2. Flowchart of the DT empowered prediction method.

1) Tensor modeling of historical data: Historical weather and power data collected from PV sites are combined and modeled as a tensor.

Firstly, c adjacent vectors $D_{xp1}, D_{xp2}, \ldots, D_{xpc}$ are established for D_{xp}, with a time interval of one sampling interval, i.e., 15 minutes, The adjacent vectors $X_1, X_2, \ldots, X_c$ and $P_1, P_2, \ldots, P_c$ adjacent to X and P are also established. For example, $D_{xp2} = \{X_2, P_2\} = [D_{xp}(2), D_{xp}(3), \ldots, D_{xp}(l+1)]^{\mathrm{T}}$ has two adjacent vectors $D_{xp1} = \{X_1, P_1\} = [D_{xp}(1), D_{xp}(2), \ldots, D_{xp}(l)]^{\mathrm{T}}$ and $D_{xp3} = \{X_3, P_3\} = [D_{xp}(3), D_{xp}(4), \ldots, D_{xp}(l+2)]^{\mathrm{T}}$, where l represents the time step of the vector. It means that each vector contains data of l sampling time points. The corresponding adjacent vectors $M_1, M_2, \ldots, M_c$ of the mask matrix can also be obtained by using its procedure.

We represent S_i as the ith training sample input into the GAN. Then, we have

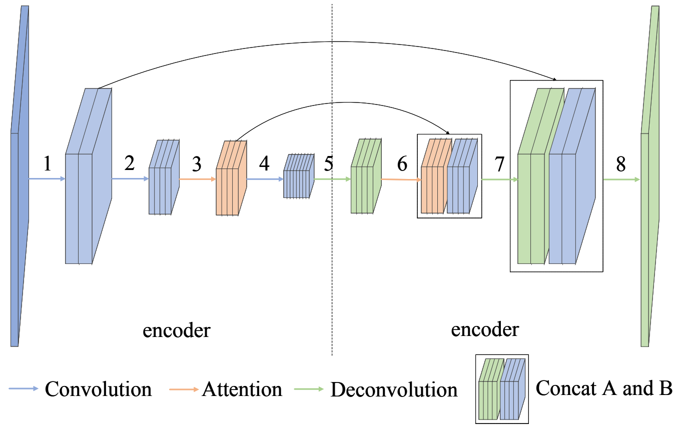

Fig. 3. Generator Network Structure.

$$S_i = \left[D_{xpi}, D_{xp(i+1)}, \ldots, D_{xpj}, P_j, P_j, P_j, \ldots\right], \quad (4)$$

where $S_i \in \mathbb{R}^{l \times l}$ and $S_i \in S$; S represents all training samples input into GAN; $i < j$, $j - i + 1 = \lfloor l/(n_x+1) \rfloor$; $f = l\%(n_x+1)$ is the number of padding vectors P_j, where n_x represents the number of meteorological factors. An binary mask matrix M with the same shape as S is created to mark the positions of missing elements. For the missing elements in S, the corresponding elements in M are set to 0; meanwhile, the rest of elements are set to 1.

After modeling the historical data into a tensor, the problem of historical data recovery becomes the recovery of missing elements in the tensor.

2) Data recovery: To achieve effective data recovery, we employ the GAN consisting of a generator and a discriminator, which is capable of learning the temporal features of the data and capturing the intrinsic relationship between meteorological data and power data. The generator uses a CNN-based encoder-decoder structure. The encoder takes the missing data set $S_0 = S \odot M$ as input and generates the latent feature representation of S_0, where $\odot$ represents the dot product operator. Then, the decoder obtains the latent feature representation and outputs $\overline{S}$, which includes the recovered part of the missing data. Furthermore, to maximize the utilization of the reliable data that is already presented in set S_0 during the data generation process, a U-net is adopted in the generator to enhance feature extraction. The discriminator takes the restored matrix $\overline{S}$ and the original complete matrix S as inputs. The generator is trained to generate the restored matrix $\overline{S}$, while discriminator is trained to judge whether the quality of the missing data recovery is realistic enough. The employed generator and discriminator network structures are shown in Fig. 3 and Fig. 4, respectively. Through the game between the generator and the discriminator, effective data recovery can be achieved.

3) The loss function of the GAN: Based on the description of the model structure, the loss function for the generator and discriminator is proposed.

The loss function of generator includes the adversarial loss and the recovery loss. The adversarial loss is defined based

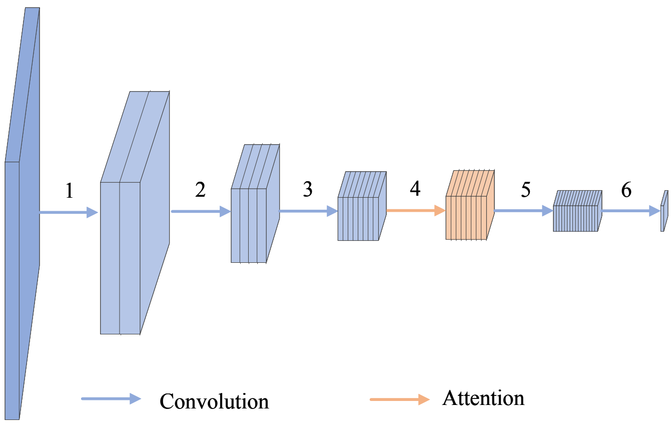

Fig. 4. Discriminator Network Structure.

on the output of discriminator, which represents the quality of recovery of missing data, i.e.,

$$L_a = -\mathbb{E}_{S,M}[D(\bar{S})]. \tag{5}$$

The recovery loss is defined as the masked root mean squared error between $\overline{S}$ and S. Since L_a has already dealt with the missing data part, the recovery loss mainly focuses on the part of intact data. The mathematical expression of the recovery loss is given by

$$L_r = \mathbb{E}_{S,M}[\|S \odot M - \bar{S} \odot M\|]. \tag{6}$$

Next, the loss function of the generator is defined as

$$L_G = L_a + L_r. \tag{7}$$

The objective of the discriminator is to maximize the discriminative value of real historical data and minimize the discriminative value of the generator's output. Therefore, the loss function of the discriminator is defined as

$$L_D = -\mathbb{E}_S[D(S)] + \mathbb{E}_{S,M}[D(\bar{S})]. \tag{8}$$

B. DT Modeling Phase

In order to build a virtual model at the DT layer that can accurately reflect the process of PV power generation in the real world, we receive the device parameters and the historical dataset after the data recovery from the transmission layer. First, we construct a digital physical model to simulate the internal mechanism of PV panel power generation. Then, a parallel CNN-BiLSTM model is built and trained to extract the inherent characteristics of meteorological factors and PV power generation. Eventually, a combination formula is applied to connect the two models to form the DT model.

1) Digital physical model: This part is composed of the underlying physical model and the power deviation correction module. Specifically, the PV power plant is a device designed to convert solar radiation into direct current electricity. It primarily consists of solar cells, which are semiconductor thin films that directly generate electricity when exposed to sunlight of a specific irradiance. These solar cells can produce voltage and current when connected in a circuit. The power output of PV cells varies due to fluctuations in weather conditions. Solar radiation plays a crucial role in determining the power output, with higher temperatures to reduce the efficiency of power generation components, while strong winds can help to reduce the temperature of PV cells, thereby increasing power generation. This behavior can be effectively modeled using an equivalent circuit.

The formula for describing the output current of a single diode equivalent circuit is given by

$$I = I_{pv} - I_{sh} - I_d, \tag{9}$$

where I_{pv} is the photocurrent generated by the battery due to incident solar radiation; I_{sh} is the short-circuit current caused by leakage at the edge of the battery and the formation of metal bridges; I_d is the diode current that comes from the Shockley equation. The mathematical expressions of I_{sh} and I_d are given by

$$I_{sh} = \frac{I \cdot R_s + V}{R_{sh}}, \tag{10}$$

$$I_d = I_0 \cdot \left(e^{\frac{q\cdot(V+I\cdot R_s)}{A\cdot b\cdot T_m}} - 1\right), \tag{11}$$

where V represents the voltage drop across the battery due to incident solar radiation; R_s is the series resistance; R_{sh} is the shunt resistance; I_0 is the reverse saturation current; q is the electron charge; A is the ideality factor of the diode; b is the Boltzmann constant; T_m is the actual temperature of the PV module, defined as

$$T_m = T + \frac{G}{\mu_0 + \mu_1 \cdot \nu}, \tag{12}$$

where T is the ambient temperature; G is the real-time irradiance; μ_0 is the irradiance-induced shading effect; μ_1 is the effect of wind speed; and ν is the real-time wind speed.

The generated power of the solar cell, denoted as P_0, is calculated as.

$$P_0 = V \cdot I. \tag{13}$$

There exist five unknown parameters, i.e., I_{pv}, I_0, R_s, R_{sh}, and A. By establishing five equations based on the short-circuit current I_{SC}, open-circuit voltage U_{OC}, maximum power $P_{0m} = U_m \cdot I_m$, $\frac{\mathrm{d}P}{\mathrm{d}V} = 0$ at the maximum power point, and $\frac{\mathrm{d}I}{\mathrm{d}V} = -\frac{1}{R_{sh}}$ at the short-circuit point, the unknown parameters can be obtained. With those components, a physical model of the PV power station can be constructed. The input data are the environmental temperature T, real-time irradiance G, and real-time wind speed ν, while the output data are the output current I, voltage drop across the battery terminals V, and power generation P_0.

Based on the predicted PV power data obtained from the aforementioned underlying physical model, the model considers only environmental temperature, real-time irradiance, and real-time wind speed as inputs. However, this approach fails to account for the complex practical conditions of the PV power station and other weather factors, leading to certain deviations in the predicted results. To address this issue, a deviation correction process is introduced. In this process, the similarity in PV power output under the influence of external climate conditions is taken into account, considering different seasons and sampling times within a day. By calculating and storing the difference between the output power of the underlying physical model and the actual historical power, it is possible to determine a correction value. This correction value is then used to adjust the predicted power from the underlying physical model, resulting in more accurate prediction results within the digital physical model. To implement the deviation correction, the historical weather data that has been restored through the use of GAN is employed as input to the underlying physical model. Let $P_0 = [P_0(1), P_0(2), \cdots, P_0(n_T)]^{\mathrm{T}}$ represent the output power, and P denote the actual power. The difference between the predicted power and the actual power of the underlying physical model can be calculated as follows:

$$\theta = P_0 - P = [\theta(1), \theta(2), \cdots, \theta(n_T)]^{\mathrm{T}}, \quad \theta \in \mathbb{R}^{n_T}. \tag{14}$$

According to (14), the revised value, denoted as $E = [E_1, E_2, \cdots, E_{n_T}]^{\mathrm{T}}$, can be calculated as

$$E_0 = 0, \tag{15}$$

$$E_i = \frac{\beta \cdot E_{i-1} + (1-\beta) \cdot \theta(i)}{1-\beta^i}, \tag{16}$$

where β is an adjustable hyperparameter between 0 and 1.

As the historical dataset used for constructing the digital physical model typically contains a large amount of data, spanning more than one year, it is essential to fully utilize this dataset while ensuring the stability of the revised value calculation. To achieve this, the calculation result of the revised value is averaged on a yearly basis, resulting in $\bar{E} = \left[\bar{E}_1, \bar{E}_2, \ldots, \bar{E}_{365t}\right]^{\mathrm{T}}$. The formula for calculating the elements in the $\bar{E}$ array is as follows:

$$\bar{E}_j = \begin{cases} \frac{1}{m}\sum_{i=1}^{m} E_{j+(i-1)\cdot 365t} & 1 \le j \le j_0 \\ \frac{1}{m-1}\sum_{i=1}^{m-1} E_{j+(i-1)\cdot 365t} & j_0 \le j \le 365t, \end{cases} \tag{17}$$

where $j_0 = n_T \% (365 \times t)$; $m = \lceil n_T / (365 \times t) \rceil$; t represents the number of data sampling times per day.

After obtaining $\bar{E}$, the power value of the corrected output power at the j-th sampling time point, representing the predicted power of the digital physical model, can be calculated as follows:

$$\widetilde{P}_D(j) = P_0(j) - \bar{E}_j. \tag{18}$$

2) Parallel CNN-BiLSTM model: To capture the underlying relationships among diverse meteorological data and the temporal dependencies within the data, we propose a parallel CNN-BiLSTM model, as depicted in Fig. 5. The parallel CNN-BiLSTM network can simultaneously process different parts of the input data and fully leverage the capabilities of parallel computing. This significantly enhances computational efficiency and speeds up both model training and inference processes. Furthermore, the parallel CNN and BiLSTM layers facilitate the extraction and integration of data features concurrently. This enables us to capture information pertaining to various aspects of the data and to provide a potent model representation, thereby improving prediction accuracy. To be specific, the CNN component is employed to extract intrinsic features between different data types within a defined time step. Meanwhile, the BiLSTM is utilized to capture deeper temporal features by considering information from both the "forward" and "backward" directions. The parallel architecture of the CNN and BiLSTM allows independent extraction of intrinsic features from various data types and deeper temporal features from the input data. These features are then concatenated into a final feature vector, which is used for predicting PV power generation.

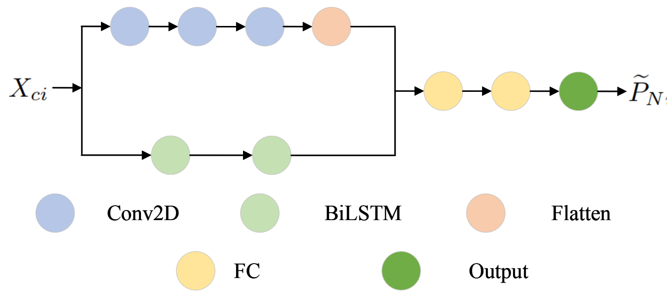

Fig. 5. the structure of the parallel CNN-BiLSTM model.

Tensor modeling was conducted on the recovered historical meteorological data and historical power data as X_c and P_c respectively. Meanwhile, the predicted power of the neural network model is defined as $\widetilde{P}$:

$$X_c = [X_{c1}, X_{c2}, \ldots, X_{cn_T}], \tag{19}$$

$$P_c = [P_{c1}, P_{c2}, \ldots, P_{cn_T}], \tag{20}$$

$$\widetilde{P}_N = \left[\widetilde{P}_{N1}, \widetilde{P}_{N2}, \ldots, \widetilde{P}_{Nn_T}\right], \tag{21}$$

$$X_{ci} = \left[x_{c(i-L)}, x_{c(i-L+1)}, \cdots, x_{c(i-2)}, x_{c(i-1)}\right]^{\mathrm{T}}, \tag{22}$$

where $X_{ci} \in \mathbb{R}^{L \times n_x}$ represents the input data required to predict the power at the ith sampling time point, L represents the time step of the input data of the parallel CNN-BiLSTM model, meteorological factors, e.g., temperature, wind speed, solar irradiance and humidity. Note that X_{ci} contains the time series data. After normalization, it is viewed as the grayscale image and served as input of Conv2D. In addition, X_{ci} is flattened and served as input of BiLSTM. x_{ci} represents the meteorological data collected at the ith sampling time point; P_{ci} represents the power at the ith sampling time point; $\widetilde{P}_{Ni}$ represents the predicted power by the parallel CNN-BiLSTM model. The loss function is set to

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\tilde{P}_{Ni} - P_{ci}\right)^2}. \tag{23}$$

where N represents the batch of samples for each training.

3) The combination formula of power prediction results: In order to leverage the advantages of the digital physical model and the parallel CNN-BiLSTM model, we design the combination formula that is a linear combination of the prediction results form the two models. The combined result is used as the final predicted PV power. We define P_D and P_N to represent the predicted values of the digital physical model and the parallel CNN-BiLSTM model. The difference θ_1 between the real power and the predicted power from the digital physical model as well as the difference θ_2 between the real power and the predicted power from the parallel CNN-BiLSTM model can be calculated as

$$\theta_1 = P_D - P = \left[\theta_1(1), \theta_1(2), \ldots, \theta_1(n_T)\right]^{\mathrm{T}}, \tag{24}$$

$$\theta_2 = P_N - P = \left[\theta_2(1), \theta_2(2), \ldots, \theta_2(n_T)\right]^{\mathrm{T}}, \tag{25}$$

In order to reduce the amount of data, maximize the use of recovered historical data and avoid the contingency of calculation results, the above two difference values are averaged annually. The calculation formula is given by

$$\bar{\theta}_\kappa(j) = \begin{cases} \frac{1}{m}\sum_{i=1}^{m}\theta(j+(i-1)\cdot 365t) & 1 \le j \le j_0 \\ \frac{1}{m-1}\sum_{i=1}^{m-1}\theta(j+(i-1)\cdot 365t) & j_0 \le j \le 365t, \end{cases} \tag{26}$$

where $\kappa = 1$ or 2.

According to (26), we can get the averaged differences $\bar{\theta}_1$ and $\bar{\theta}_2$, given by

$$\bar{\theta}_1 = \left[\bar{\theta}_1(1), \bar{\theta}_1(2), \ldots, \bar{\theta}_1(365t)\right]^{\mathrm{T}}, \tag{27}$$

$$\bar{\theta}_2 = \left[\bar{\theta}_2(1), \bar{\theta}_2(2), \ldots, \bar{\theta}_2(365t)\right]^{\mathrm{T}}, \tag{28}$$

where $\bar{\theta}_1(j) \in \bar{\theta}_1$ and $\bar{\theta}_2(j) \in \bar{\theta}_2$.

The combined formula of the power prediction results is defined as

$$\hat{P}(j) = w_1(j)\, P_D(j) + w_2(j)\, P_N(j), \tag{29}$$

where w_1 and w_2 represent the weight coefficients of the physical model predicted power and the parallel CNN-BiLSTM model, respectively. The mathematical definitions of the weight coefficients are designed as

$$w_1(j) = \frac{k\bar{\theta}_2^2(j)}{\bar{\theta}_1^2(j) + k\bar{\theta}_2^2(j)}, \tag{30}$$

$$w_2(j) = \frac{\bar{\theta}_1^2(j)}{\bar{\theta}_1^2(j) + k\bar{\theta}_2^2(j)}, \tag{31}$$

where $k > 0$ is a hyperparameter.

C. Power Prediction Phase

After finishing the phases of data preparation and DT modelling, we proceed the final power prediction phase. Taking the real-time weather data as input, we use the digital physical model and the parallel CNN-BiLSTM model to calculate the prediction results $\hat{P}_D$ and $\hat{P}_N$. Then, the final predicted power $\hat{P} = w_1\hat{P}_D + w_2\hat{P}_N$ is obtained through the calculation of the combined formula of the power prediction results.

Remark 1: The data augmentation method is a kind of data preprocessing technique for expanding training data through a series of transformations and extensions of the original data set to generate new training samples. Distinguished from the data augmentation methods, the DT focuses on creating the data counterpart of the physical systems to provide simulation and analysis. In the aspect of solving the prediction problem, the data augmentation method enables the extension of training data to deal with the data imbalance and improve the prediction accuracy. In this paper, the DT is used to create digital physical models that reflect the intrinsic mechanisms of physical systems, and use machine learning models to capture hidden features that are difficult to analyze based on physical models. This enables the integration of physical knowledge and data-driven approaches to achieve accurate modeling and prediction of real systems. In this paper, we have complete real data set without the requirement of generating new data set. Thus, we consider the introduction of the DT to increase the prediction accuracy.

IV. Simulations

A. Preparation

1) Dataset: The real dataset comes from the Global Intelligent Evolution Simulation Experiment Platform and Engineering Demonstration Application Project of Distributed Information Energy System at Northeastern University. This dataset contains historical records of relevant information on power generation and weather conditions. Specifically, it covers the period between 2016 and 2018 and includes data recorded from 8:00 to 17:00 daily. The sampling interval is 15 minutes. The data types include temperature, wind speed, solar irradiance, relative humidity, and PV output power. The first 24 months and the last 12 months of the historical data set are taken as training and testing samples, respectively. The time dimension of the data, the number of meteorological factors, and the data sampling frequency of per day are $n_T = 40515$, $n_x = 12$, and $t = 37$, respectively. To handle the missing and abnormal data, invalid data is identified and set to 0 in the mask matrix M. In order to eliminate data dimensions and enhance data features, the historical data set is normalized and then inputted into the GAN for data recovery to improve the quality of the data set.

2) Network parameters: The structures of the neural networks of the generator and discriminator are listed in Table I and Table II, respectively. The generator takes input data with a time step of $l = 92$, $n_x + 1 = 13$, resulting in $j - i + 1 = 7$ and $f = 1$. The convolutional layers in the generator network employ SAME padding with a stride of $s = 2$. Similarly, the convolutional layers in the discriminator network also adopt SAME padding, with a stride of $s = 2$, except for the last convolutional layer, which has a stride of $s = 1$. The Adam optimizer is used for the GAN with the activation function of Leaky Relu and keep-probability of 0.8. The input data of the parallel CNN-BiLSTM model in the DT layer has a time step of $L = 12$ and $n_x = 12$ meteorological factors. We chose the batch size as 64, epochs as 50, and the learning rate as 0.0002. The network parameters are shown in Table III, where the convolutional layer has no padding (i.e., $p = 0$) with a stride of $s = 1$.

3) Performance evaluation metrics for prediction: We evaluate the accuracy of PV power prediction models by using the root mean square error (RMSE) and the mean absolute error (MAE), defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_i - \hat{P}_i\right)^2}, \tag{32}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|P_i - \hat{P}_i\right|, \tag{33}$$

This article has been accepted for publication in a future issue of this journal, but has not been edited. Content will change prior to final publication.

TABLE I
PARAMETERS OF THE GENERATOR NETWORK

Layer	Part	Kernel size	Number
1	Convolution	$4*4$	16
2	Convolution	$3*3$	32
3	Attention		
4	Convolution	$4*4$	64
5	Deconvolution	$4*4$	64
6	Attention		
7	Deconvolution	$3*3$	64
8	Deconvolution	$4*4$	32

TABLE II
PARAMETERS OF THE DISCRIMINATOR NETWORK

Layer	Part	Kernel size	Number
1	Convolution	$3*3$	8
2	Convolution	$3*3$	16
3	Convolution	$5*5$	32
4	Attention		
5	Convolution	$3*3$	64
6	Convolution	$4*4$	1

TABLE III
PARAMETERS OF THE PARALLEL CNN-BILSTM MODEL

Part	Kernel size or Hidden size	Number of convolutional kernels
Conv2D 1	$4*4$	6
Conv2D 2	$4*4$	6
Conv2D 3	$3*3$	8
BiLSTM 1	64	
BiLSTM 2	64	
FC 1	128	
FC 2	64	

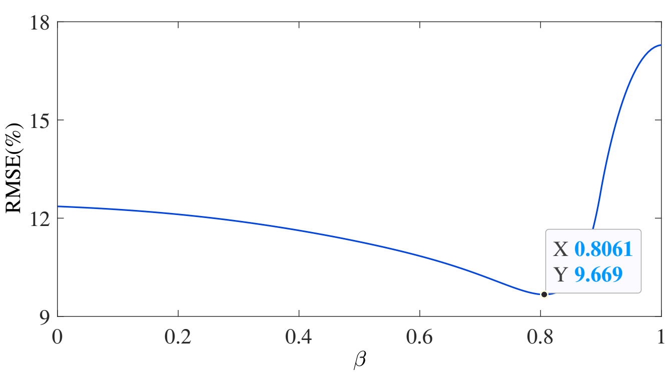

Fig. 6. RMSE-β curve.

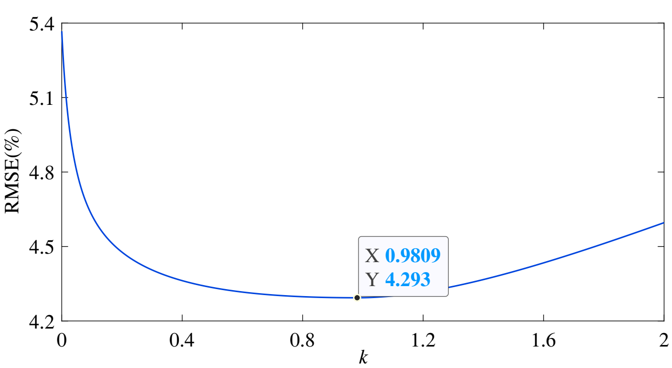

Fig. 7. RMSE-k curve.

TABLE IV
TYPICAL DAY

Weather Style	Winter	Spring	Summer	Autumn
Sunny	9/1/2018	7/4/2018	13/7/2018	27/10/2018
	10/1/2018	8/4/2018	14/7/2018	28/10/2018
	11/1/2018	9/4/2018	15/7/2018	29/10/2018
	12/1/2018	10/4/2018	16/7/2018	30/10/2018
Rainy	2/1/2018	16/4/2018	2/7/2018	14/10/2018
	5/1/2018	29/4/2018	3/7/2018	15/10/2018
	19/1/2018	17/5/2018	4/7/2018	16/10/2018
	21/1/2018	27/5/2018	24/7/2018	21/10/2018
Extreme	3/1/2018	7/3/2018	28/6/2018	5/11/2018
	4/1/2018	5/4/2018	6/7/2018	7/11/2018
	27/1/2018	23/4/2018	22/7/2018	8/11/2018
	19/2/2018	5/5/2018	17/8/2018	26/11/2018

where P_i is the measured PV power at the ith sampling time; $\hat{P}_i$ is the corresponding predicted value; n is the total number of samples.

4) Hyperparameters determination: The hyperparameter β of the deviation correction module in the digital physical model and the hyperparameter k in the combination formula of power prediction results are determined by using the grid search method. The decision principle of β and k is that the higher the accuracy of the predicted power, the better the hyperparameter selection. It means that the hyperparameter should be selected to minimize the RMSE. The searching results are shown in Fig. 6 and Fig. 7, respectively. Specifically, it follows from Fig. 6 that the optimal value of hyperparameter β is 0.8061. When β is 0, RMSE is 12.360. As β increases to 0.8061, RMSE decreases to 9.669. As β further increases to 1, RMSE increases to 17.291. According to Fig. 7, the optimal value of hyperparameter k is 0.9809. When $k=0$, RMSE=5.367. As k increases to 0.9809, RMSE decreases to 4.293. As k further increases to 2, RMSE increases to 4.596.

B. Performance Evaluation and Comparison Analysis

In this case study, we focus on evaluating the performance of the proposed DT empowered prediction method by comparing with several baselines. The baselines include CNN [11], LSTM [14], CNN-LSTM [18], and GCN [22]. We compare the prediction accuracy of those methods for different weather types (i.e., sunny, rainy, and extreme weather), and different seasons (i.e., spring, summer , autumn, and winter). Meanwhile, typical days are selected as shown in Table IV.

The prediction results by executing those methods on typical days are displayed in Fig. 8. The performance evaluation metrics, including all prediction results of the testing set, are listed in Table V. In order to comprehensively compare the predictive performance of the proposed method, Table VI and Table VII list the values of RMSE and MAE after performing the proposed method and baselines in different weather types and seasons. Meanwhile, Table VIII presents the RMSE and MAE values of the proposed method and baselines on all testing samples. The following conclusions can be drawn:

1) The proposed method obtains the lowest RMSE and MAE values compared to the baseline model, regardless of the season and weather conditions. The lowest RMSE value means

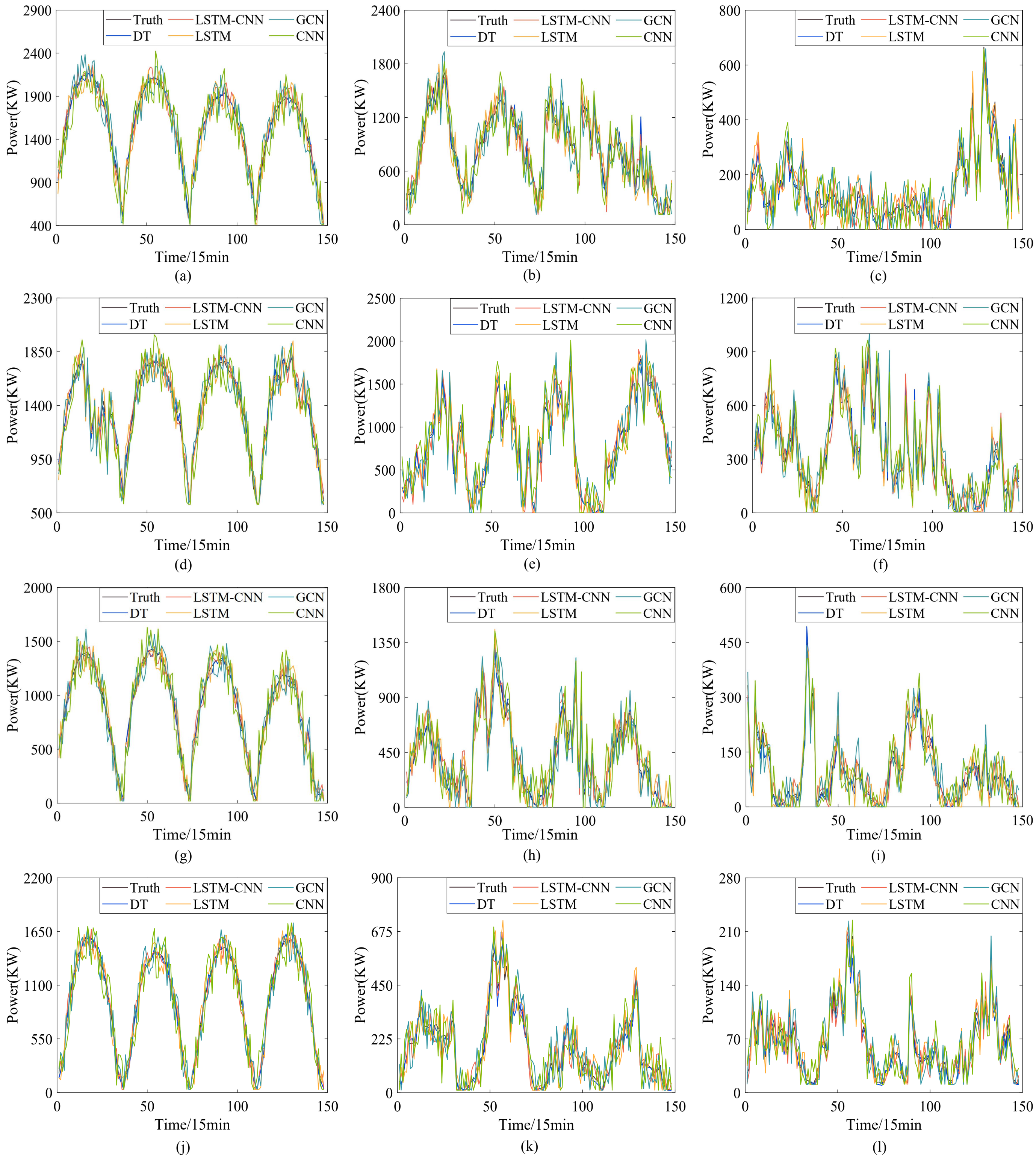

Fig. 8. The prediction PV power results of the proposed DT empowered method and baseslines: (a) Sunny weather in spring; (b) Rainy weather in spring; (c) Extreme weather in spring; (d) Sunny weather in summer; (e) Rainy weather in summer; (f) Extreme weather in summer; (g) Sunny weather in autumn; (h) Rainy weather in autumn; (i) Extreme weather in autumn; (j) Sunny weather in winter; (k) Rainy weather in winter; (l) Extreme weather in winter.

that the prediction performance of the proposed method is the most stable and the error fluctuation range is small. The lowest MAE value denotes that the difference between the predicted results of the proposed method and the actual observed values is the least.

2) In comparison to the LSTM and CNN models, the proposed method is capable of extracting spatiotemporal features from the dataset more effectively and has stronger abilities in mining data features. Compared to the CNN-LSTM model, the proposed method considers not only the inherent hidden

This article has been accepted for publication in a future issue of this journal, but has not been edited. Content will change prior to final publication.

TABLE V
COMPARISONS OF THE PROPOSED DT METHOD WITH BASELINES

Season	Weather Style	Evaluation Indicators	DT	CNN	LSTM	CNN-LSTM	GCN
$Spring$	Sunny	RMSE	5.4841	13.8077	9.1976	7.6780	11.36767
		MAE	3.8838	7.3492	5.3634	5.3930	5.2880
	Rainy	RMSE	6.9357	12.8645	10.2409	8.0464	11.7896
		MAE	5.0822	7.9312	5.9123	5.2222	6.2912
	Extreme	RMSE	3.6891	9.3243	7.4335	5.6493	8.7926
		MAE	2.3016	5.3195	4.3060	3.2325	5.0059
$Summer$	Sunny	RMSE	4.2208	7.0925	6.6126	5.5649	7.7595
		MAE	3.2375	4.3243	4.2353	3.4159	5.1515
	Rainy	RMSE	5.4557	12.8795	10.3968	8.6488	13.4175
		MAE	3.5505	8.0703	5.9831	4.8015	7.7678
	Extreme	RMSE	4.2045	13.2859	9.5298	7.7692	12.3455
		MAE	2.7993	7.4703	5.4042	4.4423	7.0225
$Autumn$	Sunny	RMSE	3.9254	10.3079	6.8957	6.2815	8.5103
		MAE	2.8185	5.6227	4.4287	4.0632	4.7555
	Rainy	RMSE	3.3121	11.1165	8.2280	6.9399	9.8849
		MAE	2.2533	6.3080	47879	4.4101	5.4118
	Extreme	RMSE	2.7799	6.1071	4.8034	3.6499	5.7009
		MAE	1.5666	3.4169	2.6687	2.0296	3.1732
$Winter$	Sunny	RMSE	4.9257	12.4006	7.9110	6.6768	9.4003
		MAE	3.6179	7.7115	5.3117	4.3161	6.2213
	Rainy	RMSE	3.0111	6.4340	5.3785	4.7058	6.0889
		MAE	1.7395	4.1226	3.3040	2.5032	3.8803
	Extreme	RMSE	3.0069	5.6353	3.7990	3.3466	5.1164
		MAE	1.9092	3.3758	2.6538	2.1242	3.0018

TABLE VI
COMPARISONS IN DIFFERENT WEATHER TYPES

Weather Style	Evaluation Indicators	DT	CNN	LSTM	CNN-LSTM	GCN
Sunny	RMSE	4.6787	11.7701	7.7518	6.6092	9.3590
	MAE	3.3787	6.2900	4.5375	4.2887	5.2065
Rainy	RMSE	4.7853	11.4384	8.6185	7.0423	10.4900
	MAE	2.7853	6.4790	4.9233	4.1279	5.8891
Extreme	RMSE	3.3973	8.7974	6.5928	5.2062	8.2042
	MAE	2.1973	5.3658	3.8859	3.1043	4.7576

TABLE VII
COMPARISONS IN DIFFERENT SEASONS

Season	Evaluation Indicators	DT	CNN	LSTM	CNN-LSTM	GCN
Spring	RMSE	5.2201	11.2352	8.2694	6.1115	9.8017
	MAE	3.4201	7.1627	4.7939	4.1282	5.6451
Summer	RMSE	4.5776	10.9193	8.6484	7.4751	9.6631
	MAE	3.1776	6.6673	5.3002	4.1911	6.7146
Autumn	RMSE	3.1688	8.9542	6.3748	4.9967	7.7242
	MAE	2.1688	5.3829	3.5284	3.2643	4.4611
Winter	RMSE	3.6216	7.7566	5.1221	4.9871	6.6926
	MAE	2.4216	4.9165	3.7249	2.8665	4.1703

TABLE VIII
COMPARISONS ON ALL TEST SAMPLES

Evaluation Indicators	DT	CNN	LSTM	CNN-LSTM	GCN
RMSE	4.2934	9.8195	6.9598	5.8476	9.1588
MAE	2.7841	6.2591	4.4019	3.7675	5.3338

TABLE IX
ABLATION ANALYSIS UNDER DIFFERENT WEATHER TYPES

Weather Style	Evaluation Indicators	DT	Model-1	Model-2
Sunny	RMSE	4.6787	14.1536	6.1925
	MAE	3.3787	8.7514	3.2086
Rainy	RMSE	4.7853	7.8433	6.6547
	MAE	2.7853	4.8217	3.8305
Extreme	RMSE	3.3973	3.7055	3.2356
	MAE	2.1973	2.4293	2.3866

features of weather and power data but also takes into account the practical conditions of photovoltaic panels and other devices. For the GCN, it relies primarily on the adjacency relationships of nodes, which limits information propagation and leads to lower prediction accuracy. Consequently, the prediction accuracy of the proposed method is significantly superior to that of baselines.

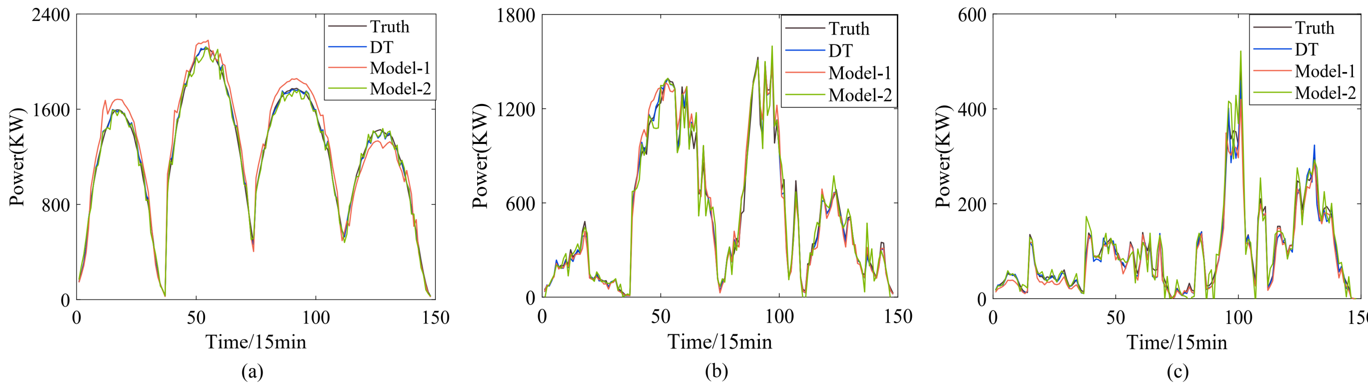

Fig. 9. The prediction PV power results under ablation analysis:(a) Sunny; (b) Rainy; (c) Extreme.

TABLE X
ABLATION STUDIES IN DIFFERENT SEASONS

Season	Evaluation Indicators	DT	Model-1	Model-2
Spring	RMSE	5.2201	11.3062	5.3032
	MAE	3.4201	7.2975	3.2701
Summer	RMSE	4.5776	9.7968	6.1744
	MAE	3.1776	5.8107	3.8261
Autumn	RMSE	3.1688	6.2832	3.9064
	MAE	2.1688	4.2208	2.2982
Winter	RMSE	3.6216	9.5294	4.9965
	MAE	2.4216	5.0249	2.7553

TABLE XI
ABLATION STUDIES IN ALL TESTING SAMPLES

Evaluation Indicators	DT	Model-1	Model-2
RMSE	4.2934	9.6687	5.3674
MAE	2.7841	5.5808	3.2338

C. Ablation Analysis

In order to further demonstrate the effectiveness of the proposed method, ablation studies are conducted in this case study. Fig. 9 displays the prediction results for three different weather types, where one typical day is selected for each weather type in four seasons, i.e., winter, spring, summer, and autumn. Table IX and Table X compare the prediction accuracy of the combined prediction results of the proposed method with two different sub-methods, i.e., the digital physical model (Model-1) and the parallel CNN-BiLSTM model (Model-2), in different weather types and seasons. Table XI displays the prediction performance of different methods in the entire testing samples.

The results indicate that the combined version achieves the highest prediction accuracy compared to the digital physical model and the parallel CNN-BiLSTM model. This is because the proposed method takes advantages of both the physical characteristics of PV power station and the inherent data features between meteorological and power data. This approach enables better simulation of real-world PV power generation processes and achieves accurate prediction of PV power generation.

V. Conclusion

In the paper, we have proposed a DT empowered PV power prediction framework to achieve reliable data transmission and power prediction with high accuracy. We have designed the use of GAN for data recovery from historical data, which is capable of significantly improving the quality of constructing a DT virtual power station. This enhances the reliability of mapping from the physical space to the digital space. We have propsoed a novel DT empowered prediction method. By integrating the digital physical model and the parallel CNN-BiLSTM model, the proposed method effectively enhances the prediction accuracy for PV power generation. Finally, the testing results on the real data set from Northeastern University show that the proposed method can achieve higher prediction accuracy that the baselines in different scenarios. In future work, we would like to investigate the integration of federated learning to enhance the privacy of the proposed method.

References

[1] A. Shafi, H. Sharadga, and S. Hajimirza, "Design of Optimal Power Point Tracking Controller Using Forecasted Photovoltaic Power and Demand," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 3, pp. 1820-1828, Jul. 2020, 10.1109/TSTE.2019.2941862.

[2] H. Zhang, Y. Li, and D. W. Gao, et al., "Distributed Optimal Energy Management for Energy Internet," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3081-3097, Dec. 2017, 10.1109/TII.2017.2714199.

[3] Y. Li, H. Zhang, and X. Liang, et al., "Event-Triggered-Based Distributed Cooperative Energy Management for Multienergy Systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2008-2022, Apr. 2019, 10.1109/TII.2018.2862436.

[4] X. Zhang, Y. Li, and S. Lu, et al., "A Solar Time Based Analog Ensemble Method for Regional Solar Power Forecasting," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 1, pp. 268-279, Jan. 2019, 10.1109/TSTE.2018.2832634.

[5] K. Hu, S. Cao, and L. Wang, et al., "A new ultra-short-term photovoltaic power prediction model based on ground-based cloud images," *Journal of Cleaner Production*, vol. 200, pp. 731-745, Nov. 2018, 10.1016/j.jclepro.2018.07.311.

[6] B. Kim, D. Suh, and M. -O. Otto, et al., "A novel hybrid spatio-temporal forecasting of multisite solar photovoltaic generation," *Remote Sensing*, vol. 13, no. 13, pp. 2605, Jul. 2021, 10.3390/rs13132605.

[7] K. Doubleday, S. Jascourt, and W. Kleiber, et al., "Probabilistic Solar Power Forecasting Using Bayesian Model Averaging," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 325-337, Jan. 2021, 10.1109/TSTE.2020.2993524.

[8] L. F. Tratar and E. Strmcnik, "The comparison of Holt-Winters method and Multiple regression method: a case study," *Energy*, vol. 109, pp. 266-276, Aug. 2016, 10.1016/j.energy.2016.04.115.

[9] H. T. C. Pedro and C. F. M. Coimbra, "Assessment of forecasting techniques for solar power production with no exogenous inputs," *Solar Energy*, vol. 86, no. 7, pp. 2017-2028, Jul. 2012, 10.1016/j.solener.2012.04.004.

[10] Y. Tang, K. Yang, and S. Zhang, et al., "Photovoltaic power forecasting: A hybrid deep learning model incorporating transfer learning strategy," *Renewable and Sustainable Energy Reviews*, vol. 162, no. 11, Jul. 2022, 10.1016/j.rser.2022.112473.

[11] J. Yan, L. Hu, and Z. Zhen, et al., "Frequency-Domain Decomposition and Deep Learning Based Solar PV Power Ultra-Short-Term Forecasting Model," *IEEE Transactions on Industry Applications*, vol. 57, no. 4, pp. 3282-3295, Jul.-Aug. 2021, 10.1109/TIA.2021.3073652.

[12] H. Li, Z. Ren, and Y. Xu, et al., "A Multi-Data Driven Hybrid Learning Method for Weekly Photovoltaic Power Scenario Forecast," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 91-100, Jan. 2022, 10.1109/TSTE.2021.3104656.

[13] F. Wang, J. Li, and Z. Zhen, et al., "Cloud Feature Extraction and Fluctuation Pattern Recognition Based Ultrashort-Term Regional PV Power Forecasting," *IEEE Transactions on Industry Applications*, vol. 58, no. 5, pp. 6752-6767, Sept.-Oct. 2022, 10.1109/TIA.2022.3186662.

[14] Y. Zhang, C. Qin, and A. K. Srivastava, et al., "Data-Driven Day-Ahead PV Estimation Using Autoencoder-LSTM and Persistence Model," *IEEE Transactions on Industry Applications*, vol. 56, no. 6, pp. 7185-7192, Nov.-Dec. 2020, 10.1109/TIA.2020.3025742.

[15] L. Cheng, H. Zang, and Z. Wei, et al., "Short-term Solar Power Prediction Learning Directly from Satellite Images With Regions of Interest," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 629-639, Jan. 2022, 10.1109/TSTE.2021.3123476.

[16] R. Zhang, H. Ma, and T. K. Saha, et al., "Photovoltaic Nowcasting With Bi-Level Spatio-Temporal Analysis Incorporating Sky Images," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 3, pp. 1766-1776, Jul. 2021, 10.1109/TSTE.2021.3064326.

[17] J. Li, C. Zhang and B. Sun, "Two-Stage Hybrid Deep Learning With Strong Adaptability for Detailed Day-Ahead Photovoltaic Power Forecasting," *IEEE Transactions on Sustainable Energy*, vol. 14, no. 1, pp. 193-205, Jan. 2023, 10.1109/TSTE.2022.3206240.

[18] S. Chai, Z. Xu, and Y. Jia, et al., "A Robust Spatiotemporal Forecasting Framework for Photovoltaic Generation," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5370-5382, Nov. 2020, 10.1109/TSG.2020.3006085.

[19] J. Simeunović, B. Schubnel, and P. -J. Alet, et al., "Spatio-Temporal Graph Neural Networks for Multi-Site PV Power Forecasting," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 2, pp. 1210-1220, Apr. 2022, 10.1109/TSTE.2021.3125200.

[20] L. Cheng, H. Zang, and Z. Wei, et al., "Solar Power Prediction Based on Satellite Measurements – A Graphical Learning Method for Tracking Cloud Motion," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 2335-2345, May. 2022, 10.1109/TPWRS.2021.3119338.

[21] M. Zhang, Z. Zhen, and N. Liu, et al., "Optimal Graph Structure Based Short-Term Solar PV Power Forecasting Method Considering Surrounding Spatio-Temporal Correlations," *IEEE Transactions on Industry Applications*, vol. 59, no. 1, pp. 345-357, Jan.-Feb. 2023, 10.1109/TIA.2022.3213008.

[22] T. Yao, J. Wang, and Y. Wang, et al., "Very short-term forecasting of distributed PV power using GSTANN," *CSEE Journal of Power and Energy Systems*, Oct. 2022, 10.17775/CSEEJPES.2022.00110.

[23] L. Cheng, H. Zang, and T. Ding, et al., "Multi-Meteorological-Factor-Based Graph Modeling for Photovoltaic Power Forecasting," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 3, pp. 1593-1603, Jul. 2021, 10.1109/TSTE.2021.3057521.

[24] J. Choi, J. -I. Lee, and I. -W. Lee, et al., "Robust PV-BESS Scheduling for a Grid With Incentive for Forecast Accuracy," *IEEE Transactions on Sustainable Energy*, vol. 13, no. 1, pp. 567-578, Jan. 2022, 10.1109/TSTE.2021.3120451.

[25] Q. Li, Y. Xu, and B. S. H. Chew, et al., "An Integrated Missing-Data Tolerant Model for Probabilistic PV Power Generation Forecasting," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4447-4459, Nov. 2022, 10.1109/TPWRS.2022.3146982.

[26] W. Liu, C. Ren and Y. Xu, "PV Generation Forecasting With Missing Input Data: A Super-Resolution Perception Approach," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 2, pp. 1493-1496, Apr. 2021, 10.1109/TSTE.2020.3029731.

[27] Y. Nie, A. S. Zamzam and A. Brandt, "Resampling and data augmentation for short-term PV output prediction based on an imbalanced sky images dataset using convolutional neural networks," *Solar Energy*, vol. 224, pp. 341-354, Aug. 2021, 10.1016/j.solener.2021.05.095.

[28] T. Polasek and M. Čadík, "Predicting photovoltaic power production using high-uncertainty weather forecasts," *Applied Energy*, vol. 339, Jun. 2023, 10.1016/j.apenergy.2023.120989.

[29] S. Goudarzi, A. Asif and H. Rivaz, "Fast Multi-Focus Ultrasound Image Recovery Using Generative Adversarial Networks," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1272-1284, Aug. 2020, 10.1109/TCI.2020.3019137.

[30] L. Han, K. Zheng, and L. Zhao, et al., "Content-Aware Traffic Data Completion in ITS Based on Generative Adversarial Nets," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11950-11962, Oct. 2020, 10.1109/TVT.2020.3007025.

[31] Y. Li and Y. Zhang, "Digital twin for industrial internet," *Fundamental Research*, Feb. 2023, 10.1016/j.fmre.2023.01.005.

[32] A. Marot, A. Kelly, and M. Naglic, et al., "Perspectives on Future Power System Control Centers for Energy Transition," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 2, pp. 328-344, Mar. 2022, 10.35833/MPCE.2021.000673.

[33] S. Mihai, M. Yaqoob, and D. V. Hung, et al., "Digital Twins: A Survey on Enabling Technologies, Challenges, Trends and Future Prospects," *IEEE Communications Surveys and Tutorials*, vol. 24, no. 4, pp. 2255-2291, Sept. 2022, 10.1109/COMST.2022.3208773.

[34] Y. Wu, K. Zhang, and Y. Zhang, "Digital Twin Networks: A Survey," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 13789-13804, Sept. 2021, 10.1109/JIOT.2021.3079510.

[35] C. Wang and Y. Li, "Digital-Twin-Aided Product Design Framework For IoT Platforms," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9290-9300, Jun. 2022, 10.1109/JIOT.2021.3100796.

[36] H. Elayan, M. Aloqaily, and M. Guizani, et al., "Digital Twin for Intelligent Context-Aware IoT Healthcare Systems," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16749-16757, Dec. 2021, 10.1109/JIOT.2021.3051158.

[37] D. Liu, Y. Du, and W. Chai, et al., "Digital Twin and Data-Driven Quality Prediction of Complex Die-Casting Manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 8119-8128, Nov. 2022, 10.1109/TII.2022.3168309.

[38] H. Xu, A. Berres, and S. B. Yoginath, et al., "Smart Mobility in the Cloud: Enabling Real-Time Situational Awareness and Cyber-Physical Control Through a Digital Twin for Traffic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 3, pp. 3145-3156, Mar. 2023, 10.1109/TITS.2022.3226746.

[39] L. Cascone, M. Nappi, and F. Narducci, et al., "DTPAAL: Digital Twinning Pepper and Ambient Assisted Living," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 2, pp. 1397-1404, Feb. 2022, 10.1109/TII.2021.3090363.

Xiaoyu Zhang received the B. S. degree in automation from the North China Electric Power University, Beijing, China, in 2018, respectively. She is currently a graduate student in electrical engineering at Northeastern University, Shenyang, China. Her main research interests include machine learning, digital twin, and data prediction.

Yushuai Li received the B. S. degree in electrical engineering and automation, and the Ph.D. degree in control theory and control engineering from the Northeastern University, Shenyang, China, in 2014 and 2019, respectively. He is currently a Marie Curie Researcher at the Department of Informatics, University of Oslo, Norway. He received the Best Paper Awards from Journal of Modern Power Systems and Clean Energy and 2023 International Conference on Cyber-energy Systems and Intelligent Energies (ICCSIE). His main research interests include distributed optimization and control, machine learning, digital twin, and their applications in integrated energy and transportation systems.

Tianyi Li received the Ph.D. degree from Aalborg University, Denmark, in 2022. She is currently an assistant professor with the Department of Computer Science, Aalborg University. Her research concerns spatio-temporal data management and analytics, knowledge integration, graph neural network. and digital energy.

Yonghao Gui received the B.S. degree in automation from Northeastern University, Shenyang, China, in 2009, and the M.S. and Ph.D. degrees in electrical engineering from Hanyang University, Seoul, South Korea, in 2012 and 2017, respectively. From 2017 to 2022, he was with Aalborg University, Aalborg, Denmark, as a Postdoctoral Researcher and an Assistant Professor. In 2022, he was with Vestas, Aarhus, Denmark. He has been working with Oak Ridge National Laboratory, Oak Ridge, TN, USA. His research interests include control of power electronics in power systems, renewable energy integration, and smart grids.

Qiuye Sun received the Ph.D. degree in 2007. He is currently a full Professor with Northeastern University and obtained Special Government Allowances from the State Council in China. He has authored or coauthored over 200 papers, authorized over 100 invention patents, and published over 10 books or textbooks. He is an Associate Editor of IEEE Trans NNLS, IET Cyber-Physical Systems, and so on. His current research interests include optimization analysis technology of power distribution network, network control of Energy Internet, Integrated Energy Systems and Microgrids.

David Wenzhong Gao received his M.S. and Ph.D. degrees in electrical and computer engineering, specializing in electric power engineering, from Georgia Institute of Technology, Atlanta, USA, in 1999 and 2002, respectively. He is now with the Department of Electrical and Computer Engineering, University of Denver, Colorado, USA. He is an Associate Editor for IEEE Journal of Emerging and Selected Topics in Power Electronics, and Journal of Modern Power Systems and Clean Energy. He was an editor of IEEE Transactions on Sustainable Energy. He is the General Chair for the 48th North American Power Symposium (NAPS 2016) and the IEEE Symposium on Power Electronics and Machines in Wind Applications (PEMWA 2012).