IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 30, 2018, accepted August 30, 2018, date of publication September 10, 2018, date of current version October 12, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2869424

Deep Power Forecasting Model for Building Attached Photovoltaic System

LIUFENG DU[1,2], LINGHUA ZHANG[1,3], AND XIYAN TIAN[2]

[1]College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]School of Mechanical and Electrical Engineering, Henan Institute of Science and Technology, Xinxiang 453003, China
[3]Jiangsu Engineering Research Center of Communication and Network Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Corresponding author: Linghua Zhang (zhanglh@njupt.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771258, in part by the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant KYCX18_0884, and in part by the Landmark Innovation Project of Henan Institute of Science and Technology.

ABSTRACT Geographical dispersion and output power fluctuations are the major barriers to efficient utilization and grid connection of building attached photovoltaic (BAPV). To eliminate these negative factors, a reliable energy management system and an accurate power forecasting model are necessary. In this paper, we first design an energy management micro-grid based on the energy Internet, which aims to tackle the problems faced by the grid-connected BAPV through the effective dual-flow management of energy and information. In the context of the proposed micro-grid, we propose a deep power forecasting model that employs a convolutional neural network to find the nonlinear relationship between meteorological information and BAPV power, while the data fed to the model are obtained through the 2-D Fourier transform of meteorological data. We evaluate the proposed model based on real-world meteorological and power data sets. Numerical results highlight the superiority of our forecasting model in terms of accuracy and reliability.

INDEX TERMS Building attached photovoltaic, power forecasting, convolutional neural network, energy Internet, two-dimensional Fourier transform.

I. INTRODUCTION

The booming of residential areas with a large number of high-rise buildings brings with it huge increases in energy consumption and greenhouse gas emissions [1]–[3]. To effectively achieve energy conservation and emission reductions, renewable energy technologies are being applied to the building energy system, among which photovoltaic (PV) is most widely used because of its advantages of cleanliness, sufficiency and universality, and the building attached photovoltaic (BAPV) generation system [4] is installed on a large number of buildings. More and more new residential buildings in China are being equipped with the PV system, as shown in Fig. 1. The BAPV system can directly feed generated energy to buildings, which is not only clean from a green energy perspective, but also has a lower degree of transmission loss [5].

However, similar to other renewable energy generation systems (REGS), geographical dispersion and low yield hinder the grid connection of BAPV systems [6], [7]. Moreover, due to the uncertainty of the weather the output power of PV is stochastic and uncontrollable, which also limits the development of this clean energy [8]. To address these issues, our work focuses on two aspects of BAPV systems.

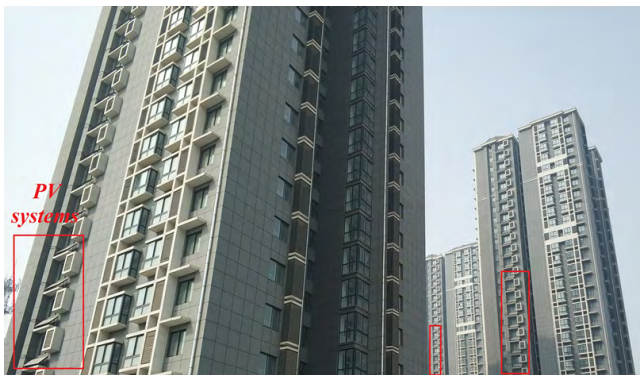


FIGURE 1. High-rise buildings equipped with the PV in a north China-based residential area.

In the first study, we attempt to change the traditional energy management model of the BAPV system based on the Energy Internet (EI) [3] scenario. The EI integrates

renewable energy and IT technologies to enable the efficient utilization and intelligent management of the REGS. In [7], an EI architecture with hierarchical integration is proposed, which provides a feasible method for REGS grid connection. Inspired by its infrastructure and design ideas, this paper proposes a BAPV managed micro-grid (BEMG) that aims to control several BAPV systems efficiently and overcome the challenges PV generation faces in local load supply and grid connection as much as possible. Moreover, for the internal energy scheduling or external energy interaction of the micro-grid, the controllability and stability are dependent on not only the infrastructure but the accurate output estimation for a large number of BAPV systems, and a reliable power forecasting model is one of the important conditions to ensure the effective implementation of the proposed BEMG. In the context of the proposed infrastructure, we conduct the second study of this paper which is the key point of our work.

Due to the inherent properties of PV, the impact of a large number of grid-connected BAPV systems on the operation of power systems cannot be ignored. Therefore, to eliminate the negative factors of the grid-connected BAPV systems and then take corresponding technical measures to suppress or compensate the fluctuations caused by them, an accurate day-ahead planning for the output power is necessary. Compared to the fossil energy-based power generation systems (such as various types of thermal power plants), whose output power can be artificially controlled, the REGS has obvious randomness and fluctuation due to the natural factors. For the BAPV system, the various meteorological elements (MEs) are the most critical among the nonlinear factors affecting power [9]. With the increasing accuracy of numerical weather prediction (NWP), the statistics-based power forecasting approach, which combines the historical ME information related to PV systems with a mathematical model, has become the most reliable approach available at present [10]. However, due to the variety and complexity of meteorological data, selecting the MEs that are highly related to PV power and preprocessing them according to the forecasting model structure are the primary problems to be solved.

The model employed completely deconstructs the complex coupling in the processed MEs by analyzing a large amount of data. Therefore, to correctly characterize the nonlinear relationship between input and output, a competent mathematical model is essential. With the development of hardware and algorithms, deep learning [11] has been successful in many areas [12]–[14]. For REGS power forecasting, deep neural networks have been applied to some scenarios [15], [16]. These deep-based models leverage multiple hidden layers to enable them to better deal with multi-variable input cases without excessive manual extraction of data features. Such an advantage makes them capable of taking into account more factors that influence the output, thus guaranteeing a lower error rate.

In view of this, we propose a deep power forecasting model for BAPV based on a convolutional neural network (CNN) [17] architecture, in an attempt to address the issues caused by the uncertainty of PV systems through reliable pre-estimation. We first selected and normalized the key MEs that influence power as input data, and put them into a matrix. Then, we refined the meteorological features in the matrix using the Fourier transform before feeding it to the CNN-based model. Finally, we trained and tested the forecasting model based on real-world data.

The contributions of our work are summarized as follows:

- We design a micro-grid, or BEMG, that can effectively manage several BAPV systems, while also eliminating the negative impacts BAPV systems face in local load supply and grid connection as much as possible.
- To overcome the heterogeneity of the MEs as forecast input data, we normalize the selected MEs in a uniform matrix format, and then we further enhance the features of the matrix through the Fourier transform.
- In accordance with the characteristics of the generated input data, we explore and build the CNN-based power forecasting model for the BAPV system and train the model in exploiting the data collected from the real world.
- We evaluate the proposal based on real-world data sets, and find that the results meet expectations. The performance comparison with other statistics-based approaches also highlights the superiority of our model.

The remaining of the paper is organized as follows. Section II describes the related work. Section III designs a management micro-grid for the BAPV system. Section IV introduces the method of producing forecast input data. Section V builds and trains the CNN-based power forecasting model. In Section VI, the evaluation results based on real-world data are given. Conclusions are summarized in Section VII.

II. RELATED WORK

How to efficiently and stably utilize renewable energy such as PV, wind energy has been drawing a lot of attention from industry and academia. Under the scenario with a high penetration of REGS, designing a stable management infrastructure and establishing an accurate power forecasting model are the most direct and effective measures against the uncertainty of renewable energy [18]. Aiming at these two aspects, researchers have made some valuable explorations.

Architecture design for the REGS, including the BAPV management infrastructure, is included in the research field of EI in general. The idea of EI has been proposed for more than a decade, but its related research is still in the pilot phase [3], [19]. In the background of large-scale REGS grid connection, Bui *et al.* [20] studied the issues of improving energy efficiency and optimizing distribution and proposed a planning for energy management based on the concept of Internet. Wang *et al.* [3], [21] studied the dual-flow management, architecture and communication standards for grid-connected renewable energy, and explored the design of EI information systems that take into account data security.

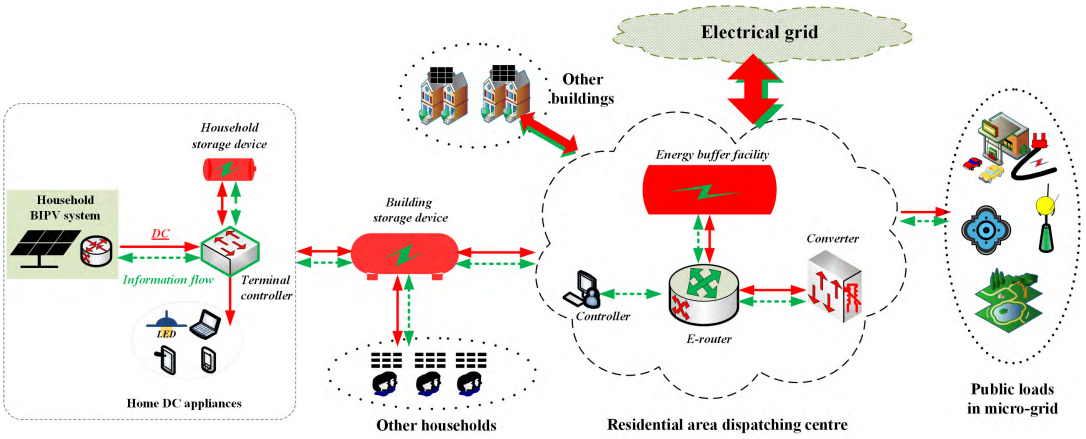


FIGURE 2. Energy management micro-grid for the BAPV system.

In [7], we designed a management architecture for the bottom renewable energy unit, and a hierarchically integrated EI architecture was proposed to provide reliable and efficient dual-flow management. However, although these contributions provide a good inspiration for this study, they lack a more targeted solution for BAPV systems, and we will attempt to tackle the problem.

According to the principle, the power forecasting of PV, or BAPV, can be roughly divided into physical methods and statistical ones. The physical methods [22], [23] do not rely on historical data but require detailed information on the physical facilities of the generation systems. The accuracy of such methods is greatly affected by the established model, the influencing factors and internal mechanisms of the PV modules, however, are complex, leaving accurate models difficult to build. Therefore, the physical methods are not widely used in practical project due to their accuracy and generalization. The statistics-based methods capture the intrinsic relationship between PV output and various influencing factors through the analysis of relevant historical information, by which to achieve power estimation. Models employed for statistical analysis include Markov chain [24], support vector regression (SVR) [9] and artificial neural networks (ANN) [25], as well as hybrids models [10], etc., while historical data are generally meteorological and power information. Sanjari and Gooi [24] took temperature, irradiance and historical power as input data, and proposed a prediction model of 15-min ahead output probability distribution for PV based on the high-order Markov chain. However, the work did not consider the impact of different climate as an important factor on power. Yang *et al.* [10] proposed a hybrid method based on NWP and SVR, which trained input data using models that correspond to different weather types. Although the work fully considered the impact of climate on power and was more practical, the MEs selected were fewer and the time resolution was also lower (1 hour). Similar shortcomings appeared in [9] as well. In addition, some schemes based on MEs and ANNs either did not refine the features of input data [26] or did not optimize the model used accordingly [27], resulting in unsatisfactory accuracy.

Deep learning has been successfully applied in various fields [13], [17], [28], which has inspired researchers engaged in REGS power forecasting to start experimenting with this advanced technology. Zhang *et al.* [15] applied deep belief network (DBN) to short- and long-term wind speed predictions to provide a reliable reference for wind power generation. Gensler *et al.* [16] used the historical power data of 21 PV facilities and the corresponding NWP data to test four deep neural networks employed for power forecasting, and the evaluation results showed the superiority of the deep-based models. Wang *et al.* [29] proposed a PV power prediction model that combines CNN and wavelet decomposition, which used historical power data as training input and achieved significantly better results than the other three models. The aforementioned studies indicate that the deep-based models can exactly find the nonlinear relationship between input and output. Accordingly, in this work, we will fully explore the relationship between meteorological information and BAPV power by leveraging the nonlinear fitting ability of deep convolutional network, and attempt to provide a new approach for the research on BAPV power forecasting.

III. BEMG FRAMEWORK

Similar to most REGS, PV generation also has the disadvantages of geographical dispersion and output uncertainty. Thus, the traditional centralized and unified management structure is difficult to adapt for the grid connection and efficient exploitation of large-scale BAPV systems, and only by informatizing the energy system, and thus controlling the energy flow through the information flow, can we overcome the problems caused by the inherent properties of REGS.

In this section, from the point of view of "using and storing on the spot and surplus grid connection" [7], We design an energy management infrastructure, as shown in Fig. 2, to achieve efficient use of the BAPV systems and to ensure

a smooth and safe transition from the island pattern to grid connection pattern.

In the illustration, every single household BAPV system in the residential area is an energy unit that depends on its own PV generation to produce energy. The generated energy is first fed to the household DC loads after being processed by the terminal controller, which includes the filter circuit and charge controller. We note that, since the installed capacity of the PV system studied in our work is small, the inverter system is not considered part of the household energy unit. The surplus energy is stored in the home storage device [30] or transmitted to the residential area dispatching center (RADC) of the BEMG via the building storage device, in order to participate in redistribution or connect into the grid. The dual-flow of energy and information of the micro-grid is a two-way transmission, and each energy unit can also pay to use the energy from others within the micro-grid for a fee through the RADC.

As the control center of the BEMG, the RADC manages the micro-grid in two patterns:

(i) In the island pattern, the RADC uses the E-router as an interactive node to coordinate the demands of its members through controllers and building storage devices, and to ensure the balance of energy supply and demand among internal units first and foremost. In addition, with the converter including the inverter and solid-state transformer the RADC can supply power for public loads such as streetlights, irrigation systems and electric vehicles in the area.

(ii) In the connection pattern, the RADC aggregates the surplus energy by controlling the energy buffer facility, and then utilizes the converter to incorporate the energy into the power grid. Furthermore, the RADC can exploit a large number of storage facilities in the micro-grid, such as building storage devices, electric vehicles, and so forth, to carry out energy reserves at low electricity prices.

The designed BEMG strives to actualize the ''using and storing on the spot'' with the help of the scheduling of the RADC, and attempts to minimize the transmission losses by enabling the generated energy to be consumed first by the units inside the micro-grid. In addition, with the dual-flow control pattern and advanced power electronics [19], the surplus energy of the micro-grid can be connected to the backbone grid flexibly, which also benefits the maximization of users' interests by storing low-cost energy.

However, whether operating in the island pattern or connection pattern, the premise of the BEMG to reach the expectation is that its operation is guaranteed to be controllable and smooth. In the island pattern, to achieve efficient peak load shifting in the micro-grid, it is necessary to forecast the output capacity of the BAPV in each period to reasonably schedule the various load demands in the micro-grid; in the connection pattern, the total output of the micro-grid also needs to be estimated in advance to achieve stable and smooth energy interaction, all of which requires the support of a reliable power forecasting model. Therefore, to eliminate the impact of the output on the forecast results, we explore the power forecasting technique that targets the BAPV system in the following sections.

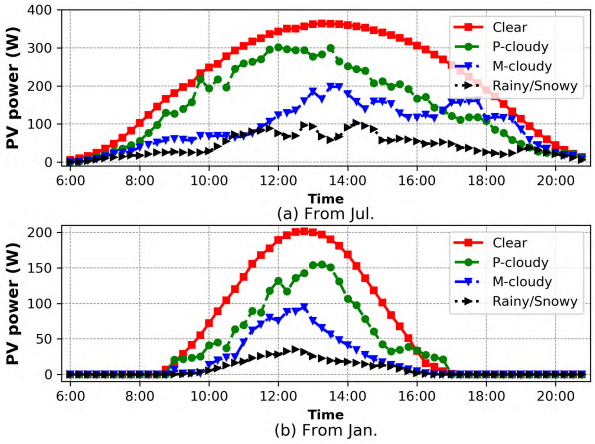


FIGURE 3. One-day power curves of PV under different conditions.

IV. MODEL INPUT DATA

In this section, we first study the MEs that influence the output power by eliciting the nature of BAPV systems, and selecting the key ones as the forecast input. Secondly, the selected MEs are preprocessed in a matrix format to match the proposed CNN-based model. Finally, to enhance the features of the input data, we transform it into 2-channel data that is more suitable for the model through the Fourier transform.

A. KEY RELATED FACTORS

Although the factors affecting the PV output are various and complex, the meteorological factor is the most important, and the output power of a PV system varies greatly with different types of weather. These include clear, cloudy, rainy, or snowy and smoggy. Fig. 3 depicts the daily power curves for the four weather types in January and July. The PV system data are taken from a small-scale independent imitating BAPV system (or OPVS, shown in Fig. 4) that we set up for collecting data and testing, with a sampling interval of 15 minutes. Here, ''P-cloudy'' represents partly cloudy and ''M-cloudy'' denotes mostly cloudy, and note that we also classify overcast or smoggy as the M-cloudy type.

In Fig. 3 (a) and (b), the output curves are smooth and efficient on clear days, and are fluctuant and inefficient under cloudy or rainy conditions, with the yield in the rain or snow being lowest. These curves illustrate the variety of PV power conditions occurring with different weather types. In addition, although the PV power is significantly diverse under the same weather type in different months because of differences in the solar angle and irradiation duration, as can be seen in the data for clear days, the peak in Fig. 3(a) is around 1.8 times that in Fig. 3(b) and it does not affect the trend of the power curve in the same way as the MEs.

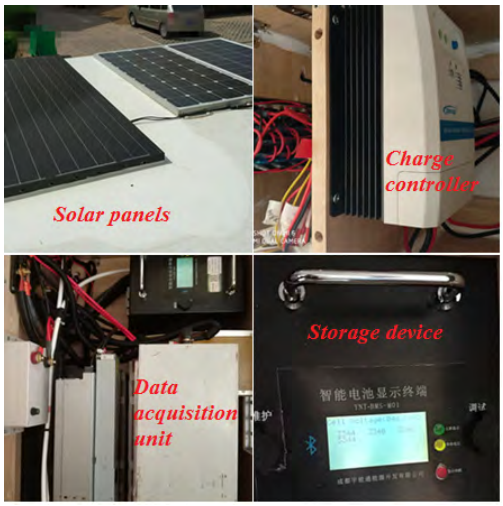


FIGURE 4. Our data collection and testing platform (located at 35.28N, 113.95E, 75mASL, capacity with 480Wp).

The illustrations reflect the diversity of PV power under different weather types, especially with non-clear conditions, and reveal the randomness of PV output. Therefore, we need to build forecasting models for different weather conditions, which also take into account the influence of specific MEs.

According to previous research [9], [10], [24], MEs, including solar irradiance, temperature, humidity, and so on, play an important role in the process of PV power generation. Among these factors, solar irradiance is particularly important. In contrast to PV farms located in open suburbs, the BAPV system usually exists in densely populated areas. Therefore, the air quality index (AQI) has a greater impact on the yield of BAPV compared with the PV farms [31], which is one important consideration during the selection process.

To accurately measure the influence of the key MEs on PV power, we used the Pearson correlation coefficient expressed in Eq. (1) to quantitatively describe the relationship between each element and the output power, as follows:

$$\rho_{W,P} = \frac{\sum_{n=0}^{N-1}(w_n - \bar{w})(p_n - \bar{p})}{\sqrt{\sum_{n=0}^{N-1}(w_n - \bar{w})^2} \cdot \sqrt{\sum_{n=0}^{N-1}(p_n - \bar{p})^2}} \quad (1)$$

where set W represents the values of MEs quantified between [0,1] (details are in the next subsection); set P is the PV data that our platform collected in January; and $\bar{w}$ and $\bar{p}$ are the sample means. N denotes the total of sampling time points during one day. The correlation coefficients under different weather types are shown in Tab. 1.

Tab. 1 shows that among the factors that are positively correlated with PV power under the same weather type, the correlation between irradiance and output is the highest, followed by temperature and humidity, while the influence of wind speed is smaller. The AQI and humidity have a greater impact but are negative. In addition, the same element

TABLE 1. Correlation coefficients under different weather types.

MEs \ Type	Clear	P-cloudy	M-cloudy	Rainy
Direct irradiance	0.93	0.85	0.80	0.72
Temperature	0.65	0.59	0.54	0.58
AQI	-0.24	-0.39	-0.53	-0.23
Relative humidity	-0.25	-0.35	-0.38	-0.47
Wind speed	0.13	0.17	0.20	-0.09

under different weather types has different influences on the PV output.

To sum up, in this work we take into account the relevance of different MEs and the influence of different weather types on PV output power, and select irradiance, temperature, AQI, humidity and wind speed as the key factors for determining PV power. In the following section, the preprocessing methods are outlined.

B. INPUT DATA PREPROCESSING

Nine kinds of MEs were selected as forecast input data in accordance with their relevance to PV power in Tab. 1, which is expressed as a feature set:

$$\mathbb{M}^{D_{-1}(n)} = \{H^{D(n)}, I_{swa}^{D(n)}, I_{lwa}^{D(n)}, I_{net}^{D(n)}, T_{rea}^{D(n)}, T_{\max}^{D(n)}, T_{\min}^{D(n)}, A^{D(n)}, W^{D(n)}\} \quad (2)$$

where $D_{-1}(n)$ denotes the nth decision point of the day D-1 that to carry out the forecasting, $\forall n \in \{0, 1, 2, \ldots, N-1\}$, D is the forecast day. According to Chinese standard [32], the daily forecast period in our work is set from 6:00 to 20:45 and the time interval is set to 15 minutes, hence the decision stage is split into 60 points along the horizontal time, i.e. $N = 60$. The elements in $\mathbb{M}^{D_{-1}(n)}$ indicate that the MEs captured in advance according to the weather prediction of the day D. $H^{D(n)}$ denotes relative humidity; $I_{swa}^{D(n)}$, $I_{lwa}^{D(n)}$ and $I_{net}^{D(n)}$ indicate three kinds of solar irradiance respectively: incident short wave, long wave and net radiation; $T_{rea}^{D(n)}$, $T_{\max}^{D(n)}$ and $T_{\min}^{D(n)}$ indicate that the real-time, maximum, and minimum temperature, respectively; $A^{D(n)}$ is AQI and $W^{D(n)}$ is real-time wind speed.

The exact values of the above MEs are all downloaded from the various NWP systems' databases. and the sequence of the elements is based on the processing mechanism of the convolutional network, with the purpose of exploiting the data information properly.

Obviously, the MEs are heterogeneous data with various measurements. Thus, before being put into the same matrix, they need to be normalized in a uniform format. We take $T_{rea}^{D(n)}$ as an example to give the method for all elements, as follows:

$$\widehat{T}_{rea}^{D(n)} = \mathbf{R}\left(\frac{T_{rea}^{D(n)} - \min(T_{rea})}{\max(T_{rea}) - \min(T_{rea})}\right) \quad (3)$$

where $\max(T_{rea})$ and $\min(T_{rea})$ represent the historical maximum and minimum values, respectively. $\mathbf{R}(\cdot)$ is to prevent $T_{rea}^{D(n)}$ from overflowing with extreme values.

Using Eq. (3), the temperature, as well as other elements, are mapped into [0,1] of dimensionless. The set of normalized feature scalars is expressed as:

$$\begin{aligned} \mathrm{M}_{sca}^{D_{-1}(n)} &= [\widehat{M}_0^{D(n)}, \widehat{M}_1^{D(n)}, \ldots, \widehat{M}_{k-1}^{D(n)}, \ldots \widehat{M}_{K-1}^{D(n)}]_{1\times K} \\ &= [\widehat{H}^{D(n)}, \widehat{I}_{swa}^{D(n)}, \ldots, \widehat{T}_{rea}^{D(n)}, \ldots, \widehat{W}^{D(n)}]_{1\times 9} \end{aligned} \quad (4)$$

Then, we grade each scalar $\widehat{M}_{k-1}^{D(n)}$ by its value and convert it into a sparse vector, which is expressed as:

$$\widehat{M}_{k-1}^{D(n)} \Rightarrow m_{k-1}^{D(n)} = [0, 0, \ldots, \underset{\uparrow}{1}, \ldots, 0]_{1\times L} \quad (5)$$
$$\qquad\qquad\qquad\qquad\qquad\qquad\quad \uparrow 0 \qquad\quad \uparrow l-1$$

where $m_{k-1}^{D(n)}$ denotes a sparse vector resulting from the scalar. l out of L levels indicates the grade of ME at n, and larger l indicates a larger scalar value, hence L is called the element resolution. Through this conversion, the value of each scalar is reflected by the position of "1" in the sparse vector. To demonstrate the vectorization process, we need to introduce a function:

$$f_{vec}(x, l) = \sum_{\tau=0}^{l} [\varepsilon(x - \frac{\tau}{L}) - \varepsilon(x - \frac{\tau+1}{L})] \quad (6)$$

where $\varepsilon(\cdot)$ is the Step Function, and $\forall l \in \{0, 1, 2, \ldots, L-1\}$. With the function $f_{vec}(\cdot)$, we can obtain the feature matrix $\mathrm{M}_{nor}^{D_{-1}(n)}$, as follows:

$$\begin{aligned} \mathrm{M}_{nor}^{D_{-1}(n)} &= \begin{bmatrix} f_{vec}(\widehat{M}_0^{D(n)}, l) \\ f_{vec}(\widehat{M}_1^{D(n)}, l) \\ \vdots \\ f_{vec}(\widehat{M}_{K-1}^{D(n)}, l) \end{bmatrix}_{l=\{0,1,\ldots,L-1\}} \\ &= [m_0^{D(n)}; m_1^{D(n)}; \ldots; m_{k-1}^{D(n)}; \ldots m_{K-1}^{D(n)}]_{K\times L} \end{aligned} \quad (7)$$

This process is shown in detail in **Algorithm 1**. After the above processing, complex and heterogeneous original MEs are converted into a simple unified matrix format. However, the sparsity of the matrix can weaken the diversity of features under the training mechanism of the convolutional network. Therefore, we further probe into the methods to enhance the diversity of features.

C. INPUT DATA TRANSFORMATION

First, we modify the feature matrix, before performing the transformation. Considering the background of the Fourier theory [33], we treat the $m_{k-1}^{D(n)}$ as an L-point discrete spectrum, and further use the frequency to indicate the meteorological scalar. Since the lowest frequency and the highest of the digital frequency are 0 and π within one period of $[0, 2\pi)$ respectively, in order to exactly reflect the meteorological data in terms of frequency and to avoid the impact of symmetry on representation, we add L zeros at the end of $m_{k-1}^{D(n)}$

Algorithm 1 Process of Solving the Feature Matrix

Input: Raw meteorological data set $\mathbb{M}^{D_{-1}(n)}$ and their corresponding historical data including maximum and minimum, N, K and L.
Output: Feature matrix $\mathrm{M}_{nor}^{D_{-1}(n)}$.
Initialization: Matrix of $K \times L$: M_{nor}.

```
for n = 1, 2, ..., N do
    Import a data set M^{D_{-1}(n)} and the historical values.
    for k = 1, 2, ..., K do
        M_n ← M^{D_{-1}(n)}(k)
        Substituting M_n and the historical values into the Eq. (3) to
        solve the feature scalar M̂.
        R(·) implementation:
        if M̂ > 1 || M̂ < 0 then
            M̂ = 1 or M̂ = 0
        end if
        for l = 1, 2, ..., L do
            Substituting M̂ into the Eq. (6):
            m ← Σ_{τ=0}^{l} [ε(M̂ − τ/L) − ε(M̂ − (τ+1)/L)]
            M_nor(k, l) = m
        end for
    end for
    M_nor^{D_{-1}(n)}{n} = M_nor
end for
```

to compress the position of 1 into $[0, \pi]$ over the frequency domain, as follows:

$$m_{k-1}^{D(n)} \rightarrow [0, \ldots, 1, \ldots, \underset{\uparrow}{0}, 0, \ldots]_{1\times 2L} \quad (8)$$
$$\qquad\qquad\qquad\qquad\qquad\quad L-1$$

Second, the extended matrix $\mathrm{M}_{nor}^{D_{-1}(n)}$(shape is $K\times 2L$) is mapped into the time domain through the 2-dimensional inverse Fourier transform (2D-IFT). For simplicity, we omit the superscript $D_{-1}(n)$ and express it as:

$$\begin{aligned} \mathrm{M}(r, c) &= \sum_{k=0}^{K-1} \sum_{l=0}^{2L-1} \mathrm{M}_{nor}(k, l) e^{[j2\pi(\frac{kr}{K} + \frac{lc}{2L})]} \\ &= \sum_{k=0}^{K-1} \left[\sum_{l=0}^{2L-1} \mathrm{M}_{nor}(k, l) e^{(j2\pi \frac{lc}{2L})} \right] e^{(j2\pi \frac{kr}{K})} \\ &= \sum_{k=0}^{K-1} \left[\tilde{\mathrm{M}}(k, c) \right] e^{(j2\pi \frac{kr}{K})} \end{aligned} \quad (9)$$

where $\tilde{\mathrm{M}}(k, c)$ denotes the transformation along the column direction. For the Fourier transform, the zero-fill operation does not improve the frequency resolution, that is, no further details of the feature can be obtained. Thus, to reduce the redundant operation of our model, we only obtain the even

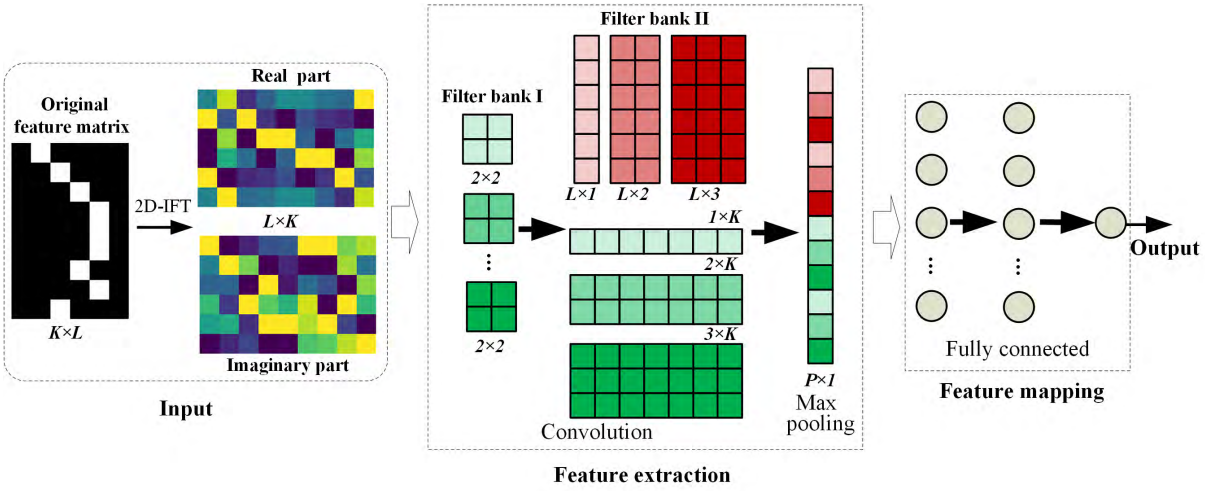


FIGURE 5. Architecture of the proposed model for power forecasting.

component of $\tilde{\mathrm{M}}(k, c)$, as follows:

$$\begin{aligned}\tilde{\mathrm{M}}(k,c) &= \sum_{l=0}^{2L-1} \mathrm{M}_{nor}(k,l)e^{(j2\pi\frac{lc}{2L})} \\ &= \sum_{l=0}^{L-1} \mathrm{M}_{nor}(k,l)e^{(j2\pi\frac{lc}{2L})} + \sum_{l=L}^{2L-1} \mathrm{M}_{nor}(k,l)e^{(j2\pi\frac{lc}{2L})} \\ &= \sum_{l=0}^{L-1} \left[\mathrm{M}_{nor}(k,l) + e^{(jc\pi)}\mathrm{M}_{nor}(k,l+L)\right]e^{(j2\pi\frac{lc}{2L})}\end{aligned} \tag{10}$$

Let $\tilde{c} = \{0, 1, \ldots, L-1\}$, and when $c = 2\tilde{c} = \{0, 2, 4\ldots\}$, the even component $\tilde{\mathrm{M}}(k, 2\tilde{c})$ can be expressed as:

$$\tilde{\mathrm{M}}(k,2\tilde{c}) = \sum_{l=0}^{L-1} [\mathrm{M}_{nor}(k,l) + \mathrm{M}_{nor}(k,l+L)]e^{(j2\pi\frac{l\tilde{c}}{L})} \tag{11}$$

Substituting Eq. (11) into Eq. (9), and we get a transformed feature matrix of $K \times L$, which is expressed as follows:

$$\mathrm{M}(r,c)|_{c=\{0,2,4,\ldots\}} = \sum_{k=0}^{K-1} \left[\tilde{\mathrm{M}}(k,2\tilde{c})\right]e^{(j2\pi\frac{kr}{K})} \tag{12}$$

where $\mathrm{M}(r, c)$ is the matrix whose features are enhanced by the Fourier transform.

Finally, considering that $\mathrm{M}(r, c)$ is a complex matrix, we decompose it into a real part and an imaginary part, which not only avoids complex data processing but also increases the dimensionality of the data features, as follows:

$$\begin{aligned}\mathrm{M}(r,c)|_{K\times L} &= \mathrm{M}_{rea}(r,c) + j\mathrm{M}_{ima}(r,c) \\ &\Rightarrow \begin{cases} \sum_{k=0}^{K-1} \left[\tilde{\mathrm{M}}(k,2\tilde{c})\right] \cos(2\pi\frac{kr}{K}) \\ \sum_{k=0}^{K-1} \left[\tilde{\mathrm{M}}(k,2\tilde{c})\right] \sin(2\pi\frac{kr}{K}) \end{cases}\end{aligned} \tag{13}$$

With the preprocessing of meteorological data, we can obtain 2-channel input data that stems from the MEs selected. After this, we embarked on the building of the CNN-based forecasting model.

V. FORECASTING MODEL

The core purpose of the CNN is feature learning, which aims to capture a hierarchical and independent characterization of input objects through a layered network. A typical CNN is mainly composed of convolutional layers, pooling layers, and fully connected layers, where the convolution cooperates with the pooling to extract features layer by layer and finally generates output via several fully connected neurons [17], [34]. Combining this scenario and the purpose of this paper, we introduce the proposed CNN-based forecasting model in the following sections.

A. ARCHITECTURE DESIGN

The architecture of the CNN-based model to serve our background demand is shown in Fig. 5.

The model is divided into three parts: input part, feature extraction, and the mapping part. First, we treat the 2-dimensional meteorological matrix as a 2-channel "image" of $L \times K$, which is fed to is fed to the first filter bank.

The second step, involving the 2×2 linear filters which are exploited to activate the classification features of the input data, performs convolution on the two "images" and the results are sent to the second filter bank. Unlike in classic network structures such as LeNet-5 [11], we do not implement the pooling operation following the first convolution.

The filter bank II, which is based on the 2D-IFT mechanism, consists of two types of filters: one is for the columns of the matrix and its purpose is to extract the abstract features of the meteorological elements, and to take into account the relationship with other elements. The other aims to extract the interaction information from among the MEs using the

rows of the matrix. Filters in the filter bank II are doubled. Each filter performs a convolution operation and generates row or column vectors that contain only 1-dimensional features. Then, we perform maximum pooling on each output of the previous layer, from which only a maximum value is retained in each vector, and these maximum values form a vector by concatenating.

The final feature mapping layer takes these feature vectors as inputs and employs a feedforward network with two hidden layers, where each layer has 20 neurons, to obtain the estimated value.

Note that the shape of the model's filters, neurons and so on are confirmed through trial and error, and although they are not necessarily the absolute best solution, this model architecture is optimal with a maximum probability in our scenario.

B. LEARNING ALGORITHM

Given the contributions of a large amount of literature, the classic algorithms for CNN, such as the feedforward, backpropagation and so forth [34], are quite mature. Therefore, here we only give some brief theoretical introductions based on the proposed model architecture.

The 2-channel input obtained at the decision point n according to Eq. (13) is normalized and represented as M, then the convolution output of the first layer $\{y_j^1\}$ can be expressed as:

$$y_j^1 = \mathbf{ReLU}\left(\sum_{ch} \mathrm{M}[:,:,ch] \otimes w_{ch,j}^1 + b_j^1\right) \tag{14}$$

where $\mathbf{ReLU}(\cdot)$ denotes rectified linear unit (ReLU) function, and $\otimes$ denotes the convolution operator; $w_{ch,j}^1 \in \mathbb{R}^{2\times 2}$ is the jth filter for channel ch, and the number of filters in bank I is 16 in our work; $b_j^1 \in \mathbb{R}$ denotes the bias term of this layer. Note that this convolution is boundary-padded; that is, padding is "SAME".

The operation of the second layer is valid convolution, and the output of the second layer $\{y_j^2\}$ are fed to pooling layer for down-sampling, expressed as follows:

$$y_j^3 = \mathbf{ReLU}(\lambda_j^3 \cdot \max(y_j^2) + b_j^3) \tag{15}$$

where $max(\cdot)$ indicates the max pooling, λ_j is a multiplicative bias which is generally set to constant along with b_j.

After pooling and concatenating, the feature vector of $P\times 1$ that is obtained is fully connected to the neurons of mapping module, then the final output estimate y_{est} is expressed as:

$$y_{est} = f_{act}\left[\sum_{P}(X_{P\times 1} \times \mathrm{W}_{1\times J}^1) + \mathrm{B}_{1\times J}^1\right] \times \mathrm{W}_{J\times 1}^2 + \mathrm{B}^2 \tag{16}$$

where $f_{act}(\cdot)$ is sigmoid function; $\sum_P$ represents the sum of each column of the matrix; X denotes the extracted feature vector, and W and B are the connection coefficient and bias of each layer, respectively.

The learning objective of the model is to minimize the total error, which represents the deviation between the estimated and the true value. We define the total error E_{tot} as:

$$E_{tot} = \frac{1}{2}\sum_{n=0}^{N_s}(y_{tru}^n - y_{est}^n)^2 \tag{17}$$

where y_{tru}^n denotes the true power at decision point n, and the samples for test total N_s. Focusing on this goal, our model will complete the training effort by leveraging the existing gradient descent-based algorithms [36].

C. FORECAST PROCEDURE

Fig. 6 shows the procedure of forecast implementation.

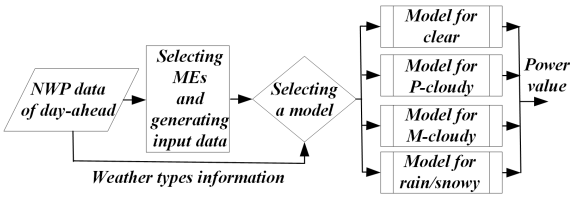


FIGURE 6. Block diagram for the forecasting model.

According to the abovementioned analysis, the models are divided into four categories in our work: sunny, P-cloudy, M-cloudy (smoggy), and rainy/snowy types, where each sub-model is built according to the CNN-based architecture in Fig. 5. During the training stage, due to the significant impact of different climate conditions on output power, we trained the four models separately according to the classified historical meteorological information; during the testing and decision stage, the meteorological feature matrices are fed to different sub-models according to their corresponding weather types.

The determination of the weather types in the test or training phase is achieved through the day-ahead weather forecast, which does not require additional classification techniques [10], and the MEs used in this process are reanalyzed data [37], [38], all of which means that in this paper we assume the NWP is approximately error-free.

VI. NUMERICAL EVALUATION

In this section, we provide details on the performance evaluation of the proposed scheme. Our forecasting model is built on *TensorFlow 1.3* [39], along with the other models picked for comparison. To achieve a fair comparison, the data used by each of the models in our evaluation is the same.

In addition, we selected the relative error (RE), root mean squared error (RMSE) and mean absolute percentage error (MAPE), which are the commonly used in power forecasting as the evaluation metrics [9], [15], [40]. The MAPE

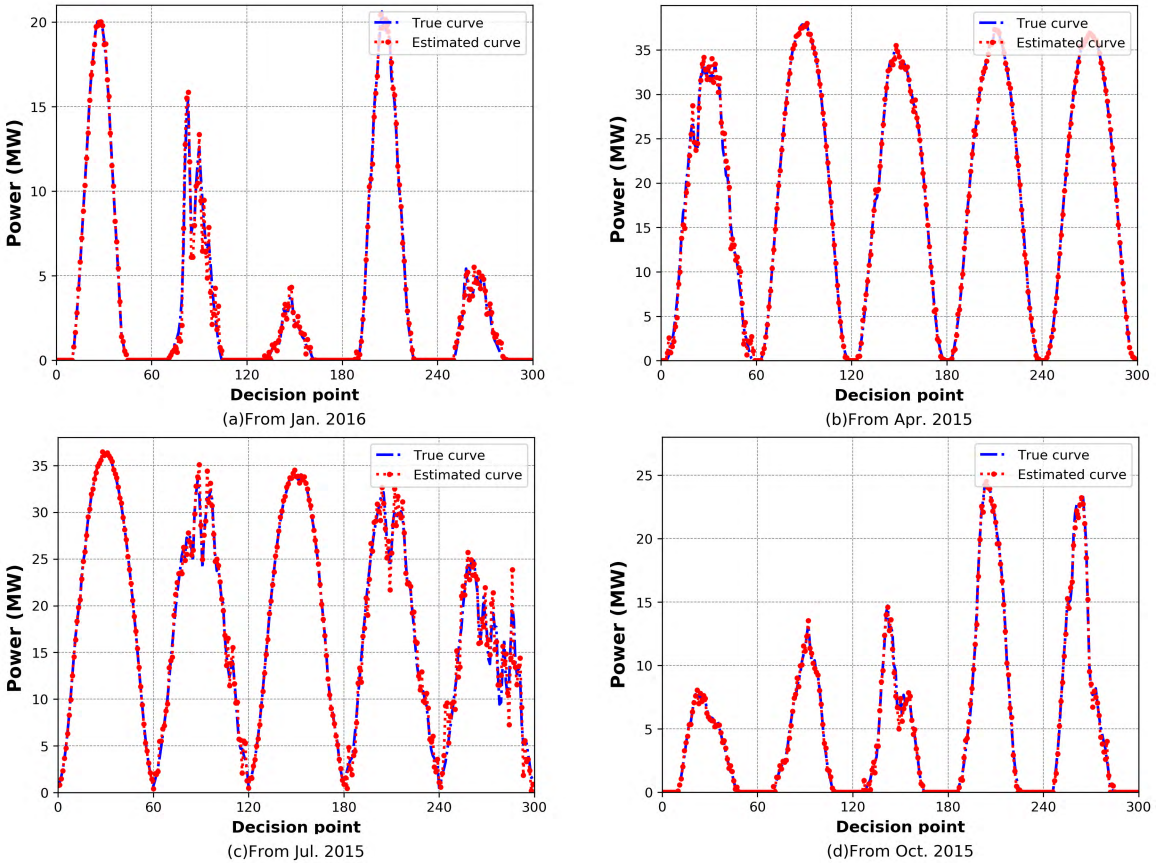


FIGURE 7. True and estimated output power curves.

is expressed as:

$$MAPE = \frac{100\%}{N_s} \sum_{n_s=1}^{N_s} \frac{|p_{est}^{n_s} - p_{tru}^{n_s}|}{p_{avg}}$$
$$p_{avg} = \frac{1}{N_s} \sum_{n_s=1}^{N_s} p_{tru}^{n_s} \quad (18)$$

where N_s is the total number of samples, $P_{est}^{n_s}$ and $P_{tru}^{n_s}$ are estimated and true power value, respectively.

A. DATA DESCRIPTION

The power data set was obtained from the corrected measurements of *ELIA*'s Brussels PV power plant [41], which is similar to the BAPV system we built in terms of climate and geography. The data set covers the period from March 2015 to February 2016 with a 15 minutes sampling rate, and thus the data volume is $366 \times 96 = 35,136$. During the evaluation, the last five days of each month's data are used to test the model and the rest as training data.

The corresponding MEs are downloaded from *ECMWF* Public Datasets [37] by employing *Python 3.6*, and the AQI data from [42]. Since the sampling interval of the MEs obtained is one hour, and considering that the MEs are usually the continuous time variables, we employed a linear interpolation-based method, which has been proven to effectively improve the resolution of NWP [43], [44], to fill the MEs to match the interval of the power data used, i.e., the meteorological data and power data are synchronized in horizontal time.

Moreover, to take into account the generalization of the model, the test power data also includes data collected from our OPVS, which totals 11, 808 and are collected in January, April, July and October of 2017, and the corresponding weather data comes from CMDC [38]. Before being fed into the forecasting model, all of the input data are generated according to the method of Section IV, and power data is mapped to the interval [0,1] through normalization. The daily forecast period is set from 6:00 to 20:45 in our work, and thus the number of decision points in one day is 60.

B. FORECASTING RESULTS

We first set the meteorological feature matrix to a standard data format (SDF) of $K = 9$, $L = 6$, (nine elements are selected and each is divided into six grades). Therefore, the meteorological data for forecasting are processed into the 6×9 2-channel matrices before being fed to the input layer of the model. The model structure, filter banks, and fully connected neurons are set according to Section V, and the optimizer employs ADAM.

Based on the above, we completed the training effort, and tested the trained model using the related data from the last five days of a total of four months. The power data from *ELIA*'s database and the results are shown in Fig. 7.

Through the analysis in Section III, we have learned that the weather type has a notable influence on power output; that is, the output is stable and smooth in stable weather, while volatile weather causes power fluctuations. In the four illustrations of Fig. 7, curves with a large peak value and smoothness reflect clear weather, which represents an approximate linear case and the estimated power curves fit well with the true curves. For fluctuating output caused by unstable climates, including M-cloudy and smoggy weather, the error of the proposed curves is relatively large and this indicates that our model still has room for improvement when dealing with strongly nonlinear cases. On a rainy or snowy day, although the power output has a small peak due to low solar radiation, as seen in Fig. 7 (a) and (d), the true curve fluctuates less because the climate is consistent and thus the error is small as well.

To assess the generalization of the forecasting model, three months of data gathered from the OPVS were also used for testing. The results are shown in Fig. 8, 9 and 10.

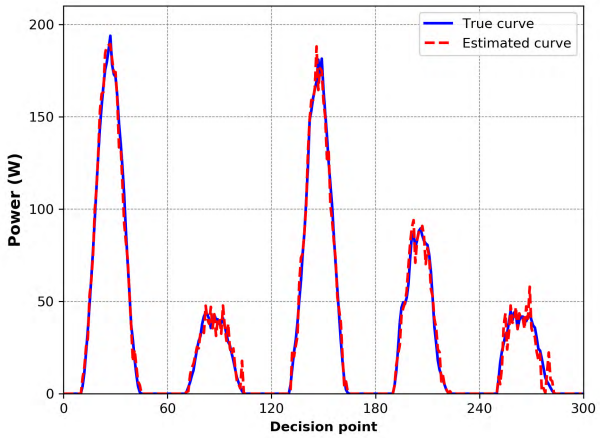


FIGURE 8. True and estimated output power curves for Jan. 2017.

From Fig. 8-10, the related data from other installations (or regions) can also be adequately estimated by the trained model, especially in clear type weather. In this case, for unstable climate conditions the error is larger. Overall, although they are slightly worse than the test results using the same database data in Fig. 7, Fig. 8-10 still illustrate a better generalization of the proposed method. The main cause for this is that, with the proposed data preprocessing method, the forecasting model can discover the inner link between MEs and PV output exactly and can thereby create correct mapping for them.

The above evaluation shows the effectiveness and accuracy of the proposed model. Next, we specifically analyzed the factors or parameters affecting the estimate.

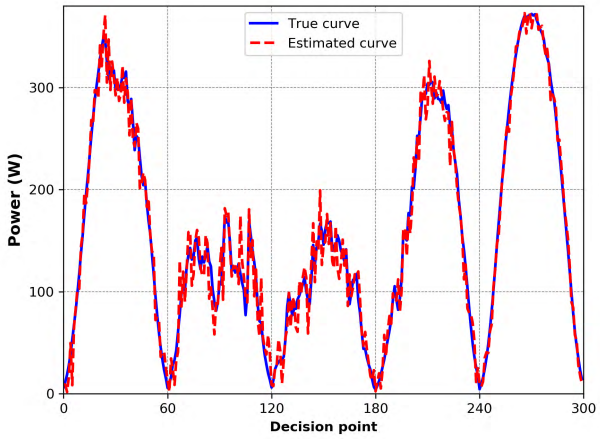


FIGURE 9. True and estimated output power curves for Jul. 2017.

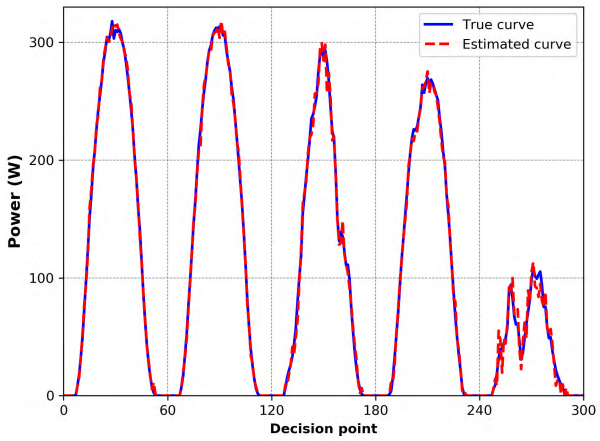


FIGURE 10. True and estimated output power curves for Oct. 2017.

C. PERFORMANCE ANALYSIS

To probe the influence of some key factors on the accuracy of the forecasting model, we took weather type, factor resolution, and the model optimizer as objects and performed different evaluations separately. The data for the first two tests comes from OPVS, and for the last test comes from *ELIA*.

Fig. 11 illustrates the cumulative distribution of the MAPE of our model under four types of climatic conditions. We selected clear, P-cloudy, and M-cloudy for 30 days from the historical NWP data to be used randomly as input, and used the related data from April as a mixed input. Each MAPE is calculated on a statistical scale for one day ($N_s = 60$). From the illustration, it can be seen that the overall error of the model on clear days is the smallest, with about 90% of MAPE being lower than 0.012. The total error of the M-cloudy type is the largest, with the error of more than half of the estimated values being higher than 3.5%, which is still attributed to the deficiency of our model in a strongly nonlinear scenario. However, this is a common weakness among these statistics-based forecasting models [24], [45]. For the mixed data, the number of the estimates for which MAPE was below 2% was 23 out of 30, and this result outperforms the industry

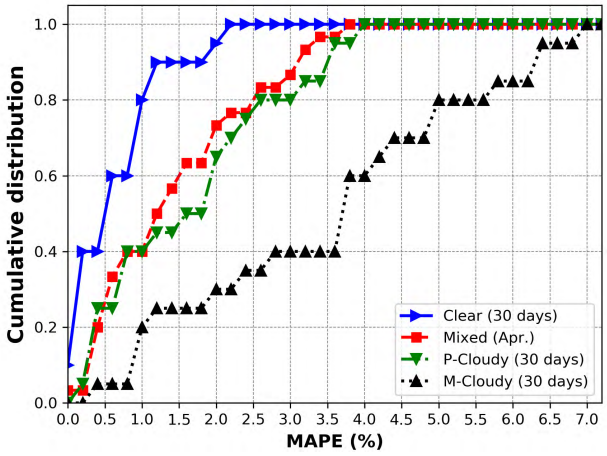


FIGURE 11. Error distribution under different weather types.

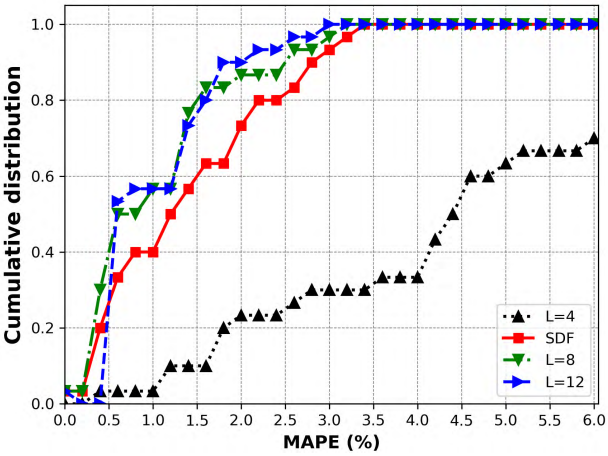


FIGURE 12. Influence of different element resolutions (*L*).

standards of ultra-short-term forecasting (see 5.3 in [32]), which is about 5%. Although these results depend on having a stable climate during April, they do demonstrate that the overall performance of our model is better.

Fig. 12 illustrates the impact of diverse element resolutions on the forecast result, using test data taken from April. As in the other conditions, the overall error decreases with increasing L. When $L = 4$, (MEs are divided into four grades only) the number of errors exceeding industry standards is close to 40%. When $L = 8$ and 10, the performance of the model keeps improving, which is due to the higher resolution representing more detailed information. However, above the SDF, the estimation value is not greatly improved. The reason for this is that if the element resolution is increased without adding to its number, the model will suffer from the Barrel Theory [46]. Thus, for a tradeoff between complexity and error, the input matrices of the following evaluation are all SDF.

Fig. 13 illustrates the training process of models that employ two typical optimizers respectively, from which the two characteristics of the proposed model are revealed:

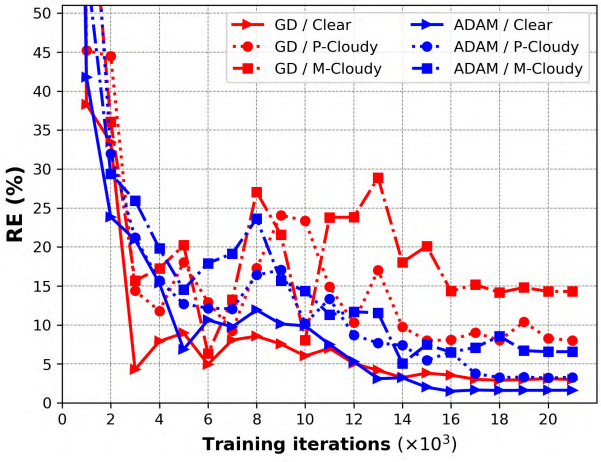


FIGURE 13. Performance of the optimizers under different weather types.

TABLE 2. Performance of models under different weather types.

Type \ Criterion	MAPE (%)			
	SVR	ANN	DBN	**Ours**
Clear	2.97	3.02	2.66	2.49
P-cloudy	4.32	4.35	3.78	3.50
M-cloudy	12.76	8.63	5.31	4.87
Rainy/Snowy	4.55	4.41	3.90	3.46
Avg	6.15	5.10	3.91	3.59

(i) Under the same climate, compared to a gradient descent (GD) optimizer, adaptive moment estimation (ADAM) has better convergence and a smaller relative error, and it performs better against the nonlinear cases. This is due to the learning rate of each iteration of ADAM having a definite boundary after completing the bias correction, which makes the convergence more stable. GD has fluctuations and large errors, but the convergence rate is faster. (ii) As mentioned above, the weather type has a significant effect on the model. Taking into account that the forecast decision occurs in non-real-time, the optimizer selects ADAM.

D. PERFORMANCE COMPARISON

To highlight the superiority of the proposed model, three well-known statistics-based forecasting models (based on SVR, ANN and DBN) were applied to the same case studies. T he structure of the SVR used for comparison was taken from previous research by Yang *et al.* [10]; the traditional ANN structure is similar to that used by Cervone *et al.* [25] but sets up 20 hidden neurons; and DBN refers to the research by Zhang *et al.* [15], and the number of nodes in a 3-layered RBM is $20 \times 20 \times 20$. In view of the structure of the three models, the input data were only quantized using Eq. (3), the labeled power data was taken from the OPVS, and the calculation parameter was $N_s = 30$ (days) $\times$ 60.

Tab. 2 lists the MAPEs generated by the four models in different climates. Compared with the other models, the

TABLE 3. Comparison of different models.

Criterion / Model	RMSE (W)					MAPE (%)				
	January	April	July	October	Avg	January	April	July	October	Avg
SVR	69.74	35.01	72.07	35.32	**53.05**	8.39	4.10	9.05	4.37	**6.49**
ANN	44.39	34.14	47.73	35.72	**40.50**	5.24	4.08	5.47	4.22	**4.72**
DBN	35.40	31.18	37.85	34.64	**34.78**	4.93	3.49	4.98	4.04	**4.21**
Ours	32.86	30.74	35.13	31.89	**32.66**	4.37	3.24	4.56	3.47	**3.91**

proposed model has the lowest average relative error and highest accuracy under each weather type. In addition, all four models perform well on clear days and have large errors with strongly nonlinear cases, while the proposed model's performance is relatively optimal on M-cloudy days. The reasons for these superiorities stem from two factors: one is that the proposed input data processing method makes the meteorological features more detailed, and the other is that feature extraction based on the combination of CNNs and multi-hidden layer output mapping enables the model to depict nonlinear relationships accurately, and is not easily trapped into local optima. SVR and ANN perform well only when the objective function is highly smooth, and DBN also has an excellent ability to discover non-linear relationships, but since only 1-dimensional meteorological data without detailed features can be fed into these models, various errors are higher than in our model.

In Tab. 3, the RMSE and MAPE criteria are provided to compare the power forecasting results over four months for the models. Similarly, the two criteria of the proposed model outperform the others for each month. For the test data of the OPVS with a capacity of 480Wp, the average absolute error of our model is about 32.66W, while the error relative to the total capacity is only about 6.8%, which is nearly 1.5 times lower than the industry standard of 10%.

VII. CONCLUSION

The efficient utilization of renewable energy sources is currently a very active research area and has important practical significance. This work targets the BAPV system, and designs a micro-grid infrastructure that is beneficial to BAPV grid connection with the background of EI. A scheme for forecasting power using meteorological information and CNN technique is proposed, which aims to guard against the uncertainty of PV power.

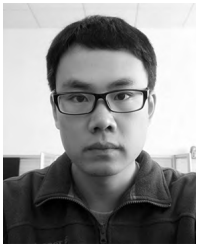
Real-world power data and their corresponding meteorological data trained and tested the model, a nd the results show that with detailed processing for meteorological data features, along with the outstanding feature abstraction and fitting ability of the CNN-based model, the proposed model exhibits satisfactory forecast accuracy and generalization. Compared with the models based on SVR, ANN and DBN, the model is superior. Therefore, it can be concluded that the proposed model is feasible and promising in the field of REGS power forecasting.

There are still several problems that need to be considered regarding the proposed model; one of which is that this paper does not involve more MEs. Based on previous research and the results of this work, we can suggest with some confidence that selecting more MEs will bring more rewards. The other issue that needs to be emphasized is that our forecasting technique relies on the accuracy of NWP, and therefore if we can integrate a high-accuracy NWP system with the model, then the proposed model will be more practical.

REFERENCES

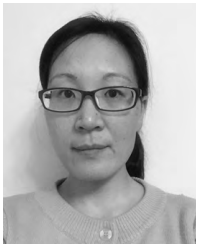
[1] M. Gao, K. Wang, and L. He, "Probabilistic model checking and scheduling implementation of an energy router system in energy Internet for green cities," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1501–1510, Apr. 2018.
[2] H. Ma *et al.*, "Analysis of typical public building energy consumption in northern China," *Energ. Buildings*, vol. 136, pp. 139–150, Feb. 2017.
[3] K. Wang *et al.*, "A survey on energy Internet: Architecture, approach, and emerging technologies," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2403–2416, Sep. 2018.
[4] B. P. Jelle, C. Breivik, and H. D. Røkenes, "Building integrated photovoltaic products: A state-of-the-art review and future research opportunities," *Sol. Energy Mater. Solar Cells*, vol. 100, pp. 69–96, May 2012.
[5] M. Sechilariu, B. Wang, and F. Locment, "Building integrated photovoltaic system with energy storage and smart grid communication," *IEEE Trans. Ind. Electron.*, vol. 60, no. 4, pp. 1607–1618, Apr. 2013.
[6] A. Etxeberria, I. Vechiu, H. Camblong, J. M. Vinassa, and H. Camblong, "Hybrid energy storage systems for renewable energy sources integration in microgrids: A review," in *Proc. IEEE IPEC*, Oct. 2010, pp. 532–537.
[7] L. Du, L. Zhang, X. Tian, and J. Lei, "Efficient forecasting scheme and optimal delivery approach of energy for the energy Internet," *IEEE Access*, vol. 6, pp. 15026–15038, 2018.
[8] K. Wang *et al.*, "Wireless big data computing in smart grid," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 58–64, Apr. 2017.
[9] J. Shi, W.-J. Lee, Y. Liu, Y. Yang, and P. Wang, "Forecasting power output of photovoltaic systems based on weather classification and support vector machines," *IEEE Trans. Ind. Appl.*, vol. 48, no. 3, pp. 1064–1069, May/Jun. 2012.
[10] H.-T. Yang, C.-M. Huang, Y.-C. Huang, and Y.-S. Pai, "A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 917–926, Jul. 2014.
[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[12] X. He *et al.*, "Green resource allocation based on deep reinforcement learning in content-centric IoT," *IEEE Trans. Emerg. Topics Comput.*, to be published, doi: 10.1109/TETC.2018.2805718.
[13] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction," in *Proc. 31st AAAI Conf. Artif. Intell.*, Feb. 2017, pp. 1655–1661.
[14] K. Wang, Y. Shao, L. Shu, G. Han, and C. Zhu, "LDPA: A local data processing architecture in ambient assisted living communications," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 56–63, Jan. 2015.
[15] C.-Y. Zhang, C. L. P. Chen, M. Gan, and L. Chen, "Predictive deep Boltzmann machine for multiperiod wind speed forecasting," *IEEE Trans. Sustain. Energy*, vol. 6, no. 4, pp. 1416–1425, Oct. 2015.

[16] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep Learning for solar power forecasting—An approach using AutoEncoder and LSTM neural networks," in *Proc. IEEE SMC*, Oct. 2016, pp. 002858–002865.
[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2012, pp. 1097–1105.
[18] K. Wang *et al.*, "Distributed energy management for vehicle-to-grid networks," *IEEE Netw.*, vol. 31, no. 2, pp. 22–28, Mar. 2017.
[19] Y. Xu, J. Zhang, W. Wang, A. Juneja, and S. Bhattacharya, "Energy router: Architectures and functionalities toward energy Internet," in *Proc. IEEE SmartGridCommun*, Oct. 2011, pp. 31–36.
[20] N. Bui, A. P. Castellani, P. Casari, and M. Zorzi, "The Internet of energy: A Web-enabled smart grid system," *IEEE Netw.*, vol. 26, no. 4, pp. 39–45, Jul. 2012.
[21] K. Wang, H. Li, Y. Feng, and G. Tian, "Big data analytics for system stability evaluation strategy in the energy Internet," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1969–1978, Aug. 2017.
[22] T. Ma, H. Yang, and L. Lu, "Solar photovoltaic system modeling and performance prediction," *Renew. Sustain. Energy Rev.*, vol. 36, pp. 304–315, Aug. 2014.
[23] E. Lorenz, J. Hurka, D. Heinemann, and H. G. Beyer, "Irradiance forecasting for the power prediction of grid-connected photovoltaic systems," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 2, no. 1, pp. 2–10, Mar. 2009.
[24] M. J. Sanjari and H. B. Gooi, "Probabilistic forecast of PV power generation based on higher order Markov chain," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2942–2952, Jul. 2017.
[25] G. Cervone, L. Clemente-Harding, S. Alessandrini, and L. D. Monache, "Short-term photovoltaic power forecasting using artificial neural networks and an analog ensemble," *Renew. Energy*, vol. 108, pp. 274–286, Aug. 2017.
[26] A. Fentis, L. Bahatti, M. Mestari, and B. Chouri, "Short-term solar power forecasting using support vector regression and feed-forward NN," in *Proc. 15th IEEE NEWCAS*, Jun. 2017, pp. 405–408.
[27] M. G. De Giorgi, P. M. Congedo, and M. Malvoni, "Photovoltaic power forecasting using statistical methods: Impact of weather data," *IET Sci. Meas. Technol.*, vol. 8, no. 3, pp. 90–97, May 2014.
[28] H. Jiang, K. Wang, Y. Wang, M. Gao, and Y. Zhang, "Energy big data: A survey," *IEEE Access*, vol. 4, pp. 3844–3861, Aug. 2016.
[29] H. Wang *et al.*, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Convers. Manage.*, vol. 153, pp. 409–422, Dec. 2017.
[30] L. Yao, B. Yang, H. Cui, J. Zhuang, J. Ye, and J. Xue, "Challenges and progresses of energy storage technology and its application in power systems," *J. Modern Power Syst. Clean Energy*, vol. 4, no. 4, pp. 519–528, Oct. 2016.
[31] A. Peled and J. Appelbaum, "Enhancing the power output of PV modules by considering the view factor to sky effect and rearranging the interconnections of solar cells," *Prog. Photovolt, Res. Appl.*, vol. 25, no. 9, pp. 810–818, Sep. 2017.
[32] *Technical Requirements for Connecting Photovoltaic Power Station to Power System*, Chinese standard GB/T 19964-2012, Jun. 2013.
[33] R. N. Bracewell, *The Fourier Transform & its Applications*, 3rd ed. New York, NY, USA: McGraw-Hill, 2000.
[34] K. Wang, J. Mi, C. Xu, Q. Zhu, L. Shu, and D. J. Deng, "Real-time load reduction in multimedia big data for mobile Internet," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 12, no. 5s, Dec. 2016, Art. no. 76.
[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE CVPR*, Jun. 2016, pp. 770–778.
[36] X. He, K. Wang, H. Huang, and B. Liu, "QoE-driven big data architecture for smart city," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 88–93, Feb. 2018.
[37] ECMWF. *Public Datasets: ERA-Interim and CAMS Near-real-Time*. Accessed: Oct. 2017. [Online]. Available: http://apps.ecmwf.int/datasets/
[38] CMDC. *Share data: NWP*. Accessed: Feb. 2018. [Online]. Available: http://data.cma.cn/site/index.html
[39] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX conf. OSDI*, Nov. 2016, pp. 265–283.
[40] K. Wang, Y. Shao, L. Shu, C. Zhu, and Y. Zhang, "Mobile big data fault-tolerant processing for ehealth networks," *IEEE Network*, vol. 30, no. 1, pp. 36–42, Jan./Feb. 2016.
[41] ELIA. *Solar-PV Power Generation Data*. Accessed: Oct. 2017. [Online]. Available: http://www.elia.be/en/grid-data/power-generation
[42] EEA. *AQ e-Reporting*. Accessed: Oct. 2017. [Online]. Available: https://www.eea.europa.eu/data-and-maps/data/aqereporting-2
[43] R. Perez *et al.*, "Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe," *Sol. Energy*, vol. 94, pp. 305–326, Aug. 2013.
[44] R. Zhang and H. Yang, "Dynamic building energy consumption forecast using weather forecast interpolations," in *Proc. IEEE SmartGridComm*, Nov. 2015, pp. 671–676.
[45] J. Antonanzas *et al.*, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016.
[46] K. Wang, Z. Ouyang, R. Krishnan, L. Shu, and L. He, "A game theory-based energy management system using price elasticity for smart grids," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1607–1616, Dec. 2015.



LIUFENG DU received the B.S. degree in electronic information engineering from Zhengzhou University in 2003 and the M.S. degree in signal and information processing from Tianjin University in 2011. He is currently pursuing the Ph.D. degree in signal and information processing with the Nanjing University of Posts and Telecommunications, Nanjing, China. He is also a Lecturer with the Henan Institute of Science and Technology, Xinxiang, China. His research interests include signal processing in the Internet of Things and energy management technology.

LINGHUA ZHANG received the M.S. degree in signal and information processing from Southeast University in 1990 and the Ph.D. degree in signal and information processing from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2005. She is currently a Full Professor with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications. Her research interests lie in the area of voice communication, Internet of Things, and signal processing in wireless communication.



XIYAN TIAN received the B.S. degree in applied electronic technology from Henan Normal University in 2002 and the M.S. degree in communication and information system from Zhengzhou University in 2012. She is currently a Lecturer with the School of Mechanical and Electrical Engineering, Henan Institute of Science and Technology, Xinxiang, China. Her research interests include wireless communication, sparse representation theory, and wireless sensor networks.