Sky Image Prediction Model Based on Convolutional Auto-Encoder for Minutely Solar PV Power Forecasting

Sky Image Prediction Model Based on Convolutional Auto-Encoder for Minutely Solar PV Power Forecasting

Yuwei Fu, Hua Chai, Zhao Zhen, *Member, IEEE*, Fei Wang , *Senior Member, IEEE*, Xunjian Xu, Kangping Li , *Member, IEEE*, Miadreza Shafie-Khah, *Senior Member, IEEE*, Payman Dehghanian , *Senior Member, IEEE*, and João P. S. Catalão , *Senior Member, IEEE*

***Abstract*—The precise minute time scale forecasting of an individual PV power station output relies on accurate prediction of cloud distribution, which can lead to dramatic fluctuation of PV power generation. Precise cloud distribution information is mainly achieved by ground-based total sky imager, then the future cloud distribution can also be achieved by sky image prediction. In previous studies, traditional digital image processing technology (DIPT) has been widely used in predicting sky images. However, DIPT has two deficiencies: relatively limited input spatiotemporal information and linear extrapolation of images. The first deficiency makes the input spatiotemporal information not rich enough, while the second creates the prediction error from the beginning. To avoid these two deficiencies, convolutional autoencoder (CAE) based sky image prediction models are proposed due to the spatiotemporal feature extraction ability of two-dimensional (2-D) CAEs and 3-D CAEs. For 2-D CAEs and 3-D CAEs, four architectures are given respectively. To verify the effectiveness of the proposed models, two typical DIPT methods, including particle image velocimetry and Fourier phase correlation theory are introduced to build the benchmark models. Besides, five different scenarios are also set and the results show that the proposed models outperform the benchmark models in all scenarios.**

***Index Terms*—Convolutional autoencoder (CAE), minute time scale, sky image, solar PV power forecasting, spatiotemporal feature.**

Yuwei Fu and Hua Chai are with the Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China (e-mail: yuweifu@ncepu.edu.cn; chhncepu@sina.com).

Zhao Zhen is with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China, and also with the Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China (e-mail: zhenzhao@ncepu.edu.cn).

Fei Wang is with the Department of Electrical Engineering, North China Electric Power University, Baoding 071003, China, with the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China, and also with the Hebei Key Laboratory of Distributed Energy Storage and Microgrid, North China Electric Power University, Baoding 071003, China (e-mail: feiwang@ncepu.edu.cn).

Xunjian Xu is with the State Grid Hunan Electric Power Company Ltd. Disaster Prevention and Reduction Center,, and also with the State Key Laboratory of Disaster Prevention and Reduction for Power Grid Transmission and Distribution Equipment, Changsha 410000, China (e-mail: bearxxj@126.com).

Kangping Li is with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China (e-mail: kangpingli@mail.tsinghua.edu.cn).

Miadreza Shafie-Khah is with the School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland (e-mail: mshafiek@univaasa.fi).

Payman Dehghanian is with the Department of Electrical and Computer Engineering, George Washington University, Washington, DC 20052 USA (e-mail: payman@gwu.edu).

João P. S. Catalão is with the Faculty of Engineering of the University of Porto and INESC TEC, 4200-465 Porto, Portugal (e-mail: catalao@fe.up.pt).

I. Introduction

A. Literature Review

WITH the increase of solar PV integrated capacity [1] and wind power integrated capacity [2], the difficulty of stable operation of the power system is also increasing. The reason behind this is that renewable energy is aggressively promoted by governments due to its clean, low-cost, and inexhaustible characteristics [3]. However, both solar PV power and wind power have inherent and prominent uncertainty caused by weather conditions and meteorological factors, which makes a big difference between their fluctuation random process and the usual load curve of consumers [4]. Thus, when it comes to large-scale solar PV power integrated into power grids, to reduce the operation difficulty generated by the difference, the key lies in weakening the uncertainty of solar PV power, one kind of effective methods is solar PV power forecasting [5]. Besides, accurate solar PV power forecasting under some cases is important to electricity price forecasting, even though the latter is more difficult than the former [6].

The output of a single PV power plant is mainly influenced by the surface irradiance [7], while the surface irradiance is dominantly impacted by the clouds with complex distribution over the plant [8]. When the clouds change drastically on the minute time scale, the surface irradiance will show significant nonlinear fluctuation [9]. The larger the capacity of a single station, the more serious the adverse effect of this fluctuation on power grids. Accurate minute time scale solar PV power forecasting

is propitious to the stability of power system operation, the consumption of solar PV power, and electricity market operation [10]. Such methods use ground-based sky images commonly [11] and can be carried out in two steps: the first step is cloud distribution prediction, which means sky image prediction; the second step is to establish a mapping relationship from cloud distribution to surface irradiance, then to the output. This article only focuses on the first step.

At present, the traditional sky image prediction research adopts digital image processing technology (DIPT). The research content is to use two adjacent images with high image resolution to calculate the cloud displacement vector, and then use the cloud displacement vector to linearly extrapolate the current image to get a prediction image [12]. The above three images get the same time resolution. The research can be divided into two categories: the first is based on image gray information and the second is based on the image Fourier frequency domain. The first category includes the scale-invariant feature transform [13], optical flow [14], and particle image velocimetry (PIV) [15], among which PIV is widely used. However, the stability and accuracy of these methods are poor to some extent [16]. The second category of methods can describe the difference between images by using mathematical expressions in less computing resources and shorter computing time [17], and the Fourier phase correlation theory (FPCT) is widely used. However, the DIPT-based methods need an idealized hypothesis: the cloud distribution and shape in two adjacent images are nearly the same on the minute time scale [18], which leads to two problems affecting the prediction performance: one is the relatively limited length of the input image sequence, which means the input spatiotemporal information maybe not rich enough; the other is the idealized linear assumption, which means the prediction error could be introduced into the models from the beginning.

B. Novelty

In addition to DIPT, sky image prediction can also draw lessons from the video prediction models based on deep learning. The common between sky image prediction and video prediction is that they have the same modeling idea, that is, using historical image sequences to predict future image sequences. The difference between them lies in spatiotemporal resolutions of image sequences, due to their different data sources. The temporal resolution of image sequences used in video prediction is very high, such as usually 24 frames per second. Dynamic information in videos often focuses on people or objects in real life, so the spatial resolution of the image is very high, and the people or objects in a video sequence usually do not undergo violent movement and change. The temporal resolution of image sequences used for sky image prediction is usually up to the second level at most, and the observed target is cloud. However, because the spatial scale of sky images is significantly larger than that of video prediction, cloud distributions in two adjacent sky images with longer time interval than video prediction is usually close. For that reason, sky image sequences can also achieve the smooth and continuous vision effect like video frame sequences, so sky image prediction can learn from video prediction methods.

Video prediction models are usually complex and deep convolutional autoencoders (CAEs) in which convolutions include two-dimensional (2-D) convolution and 3-D convolution. Specifically, over 20 computational layers are used in prediction CAEs at most in [19]. In [20], five computational layers are used in prediction CAEs at most. In reference [21], over ten computational layers are used in prediction CAEs. Convolutional neural network is a deep learning model widely used in the field of computer science [22]. An autoencoder is an unsupervised deep learning model. For video prediction models, their input and output are image sequences and can be flexibly set lengths according to the need, which means that models can easily input more than two images. Thus, more abundant spatiotemporal information can be input and then prediction performance of models can be improved. Moreover, such models can fit the complex nonlinear relationship between input and output automatically [23], which means spatiotemporal features can be effectively learned without idealization, thus avoiding the introduction of correlative errors.

C. Contributions

First, the first kind of models, CAE-based sky image prediction models, are first established. Furthermore, convolutions have 2-D and 3-D forms, and CAE models have four structures. As controls, then the second kind of models, DIPT-based sky image prediction models, are established using PIV and FPCT. Finally, five different prediction time resolutions are considered on the practical dataset, and the performance of CAEs and DIPT are compared under each scenario. The results under five scenarios indicate the performance of CAEs is superior to that of DIPT models. To sum up, the main contributions in this article include the following.

1) CAE-based sky image prediction models are proposed to overcome the deficiencies of DIPT-based sky image prediction models, including limited input image sequence length and linear image extrapolation.
2) To obtain accurate CAE-based sky image prediction models, 2-D convolution and 3-D convolution are used regarding computation layers, and four different architectures are set regarding model structures.
3) The comparison between CAE based sky image prediction models and two widely used DIPT-based methods (PIV and FPCT) is simulated under five different prediction scenarios. The superiority of the proposed models can be verified by the results.

D. Paper Organization

The rest of this article is organized as follows. Section II introduces data processing. Section III shows the methodology of the above two kinds of models. Section IV presents assessment metrics, simulation description, results and discussion. Finally, Section V concludes the article.

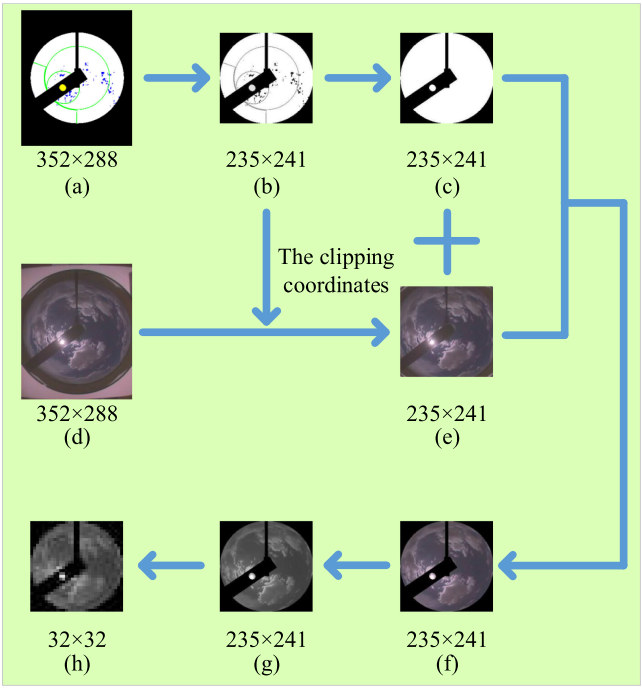


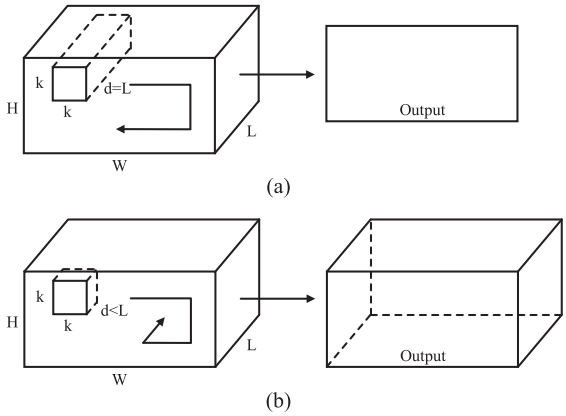
Fig. 1. Sky image processing.

II. Dataset Processing

The sky images used in this article are from the PV power station in Alamosa, Colorado, which are available on the National Oceanic and Atmospheric Administration website [24]. Their coordinated universal time (UTC) time span is from May 22, 2015 to May 31, 2015. For each image, the time resolution is 0.5 min and the size is 352 × 288. After data cleansing, the processing of an image is shown as Fig. 1, with the image resolution marked below each image.

A sky image has a corresponding cloud analyzed image, both images generated simultaneously, as shown in Fig. 1(d) and (a), respectively. A cloud analyzed image can identify the sky, clouds, the sun and the shadow band and so on, by which the nonsky information of a sky image can be filtered out effectively and the region of interest can be obtained. Specifically, first, the rectangular picture having the white area is obtained as Fig. 1(b) and the clipping coordinates are retained. Second, Fig. 1(c) is obtained by binarizing the pixel values of the Fig. 1(b). Third, Fig. 1(e) is obtained by using the clipping coordinate to cut the Fig. 1(d). Fourth, Fig. 1(f) is obtained by pixel-wise calculation between Fig. 1(c) and (e). Fifth, Fig. 1(f) is transformed into the grayscale one as Fig. 1(g). Finally, Fig. 1(h) is obtained by downsampling Fig. 1(g). It should be noted that limited to the hardware condition, and in order to reduce the training difficulty of CAE-based sky image prediction models, the sky images used are grayscale images with resolution of 32×32.

Finally, 16456 image sequences are obtained in the dataset. Each image sequence contains 20 consecutive images, the first 10 are used to construct input and the last 10 are used to construct output. The dataset is divided according to the time order, the training set accounts for 80%, the validation set accounts for 10%, and the test set accounts for 10%. For CAE-based sky image prediction models, the training set is used for model training, the validation set is used for model selection, and the test set is used for model test. DIPT based sky image prediction models only uses the test set.



H: Height, d: Height of a kernel,
L: Height of input frames, W: Width,
k: Side length of a kernel

Fig. 2. Convolution operation. (a) 2D convolution. (b) 3D convolution.

III. Methodology

A. CAE-Based Sky Image Prediction Models

The proposed CAEs consist of convolutional layers and transposed convolutional layers. The two types of layers are the same in the calculation theory, the main difference between them is generally that a convolutional layer generates a down-sampled feature map while a transposed convolutional layer generates an up-sampled one [25]. Their equations are described as follows

$$\mathbf{Y}_j = f\left(\sum_i \mathbf{X}_i * \mathbf{W}_{ij} + \mathbf{B}_j\right) \tag{1}$$

where $\mathbf{W}_{ij}$ denotes the ith weight kernel in the jth layer, $\mathbf{X}_i$ represents the input corresponding to $\mathbf{W}_{ij}$, $\mathbf{B}_j$ is the bias of the jth layer, $\mathbf{Y}_j$ is the output feature map of the jth layer and $f(\cdot)$ is an activation function.

Convolution operations including 2-D convolution and 3-D convolution is depicted as Fig. 2 [26]. A 2-D convolution kernel has two directions to move, so 2-D convolution applied on multiple frames stacked together generates a feature map. A 3-D convolution kernel has three directions to move, so 3-D convolution applied on multiple frames generates an output volume [26]. Under the same conditions, 3-D convolution is finer than 2-D convolution in terms of spatiotemporal feature extraction, but which means longer computation time.

For the structure of an auto-encoder, there are two kinds of descriptions: one contains an input layer, hidden layers and an output layer, the other contains an encoder, a bottleneck

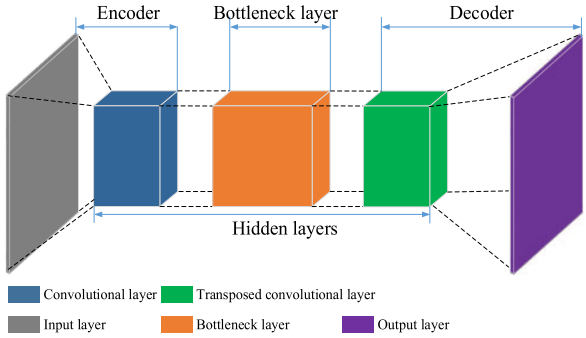


Fig. 3. Architecture of a CAE.

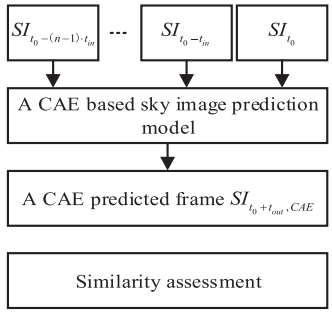


Fig. 4. Flowchart of a CAE-based sky image prediction model.

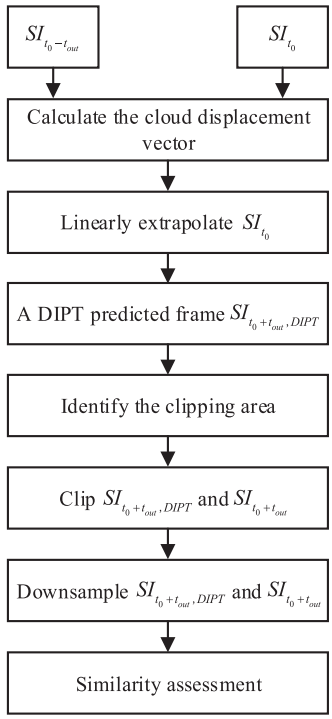


Fig. 5. Flowchart of a DIPT-based sky image prediction model.

layer and a decoder. For the CAEs proposed in this article, hidden layers consists of convolutional layers or transposed convolutional layers, and an output layer is a convolutional layer. The output of each layer is successively processed by batch normalization [27] and an activation function named as LeakyReLU [28]. A CAE with three hidden layers is shown in Fig. 2, whose structure conforms to the remaining CAEs in this article. In Fig. 3, the encoder contains an input layer and a convolutional layer, the bottleneck layer is a convolutional layer, and the decoder contains a transposed convolutional layer and an output layer.

In CAE-based sky image prediction models, an input image sequence information is processed when passing through an encoder and a bottleneck layer successively. Then, the processed information is used by a decoder to predict a future frame. This process can be described as follows

$$\mathrm{f}_{BL} = \mathrm{f}_{CE}(\mathbf{X}) \tag{2}$$

$$\hat{\mathbf{X}} = \mathrm{f}_{CD}(f_{BL}) \tag{3}$$

where $\mathbf{X}$ is a sky image sequence; $f_{CE}(\cdot)$ represents a convolutional encoder; $f_{BL}(\cdot)$ represents the features generated by a bottleneck layer; $f_{CD}(\cdot)$ represents a convolutional decoder; and $\hat{\mathbf{X}}$ is a predicted image.

According to (2) and (3), if t_0 is the time point to implement image prediction and n is the length of an input image sequence, the type of models can be described as follows:

$$SI_{t_0+t_{\mathrm{out}},CAE} = f_{CAE}([SI_{t_0-(n-1)\cdot t_{in}}, \ldots, SI_{t_0-t_{in}}, SI_{t_0}]) \tag{4}$$

where t_{out} is the time resolution of the predicted sky image; $SI_{t_0+t_{\mathrm{out}},CAE}$ represents the predicted sky image generated by a CAE; $f_{CAE}(\cdot)$ represents a CAE-based sky image prediction model; SI_{t_0} is the sky image at t_0; t_{in} is the time resolution of input images, namely 0.5 min; $n = 10$.

The flowchart of a CAE-based sky image prediction model is demonstrated as Fig. 4. For CAEs, their input images, from $SI_{t_0-(n-1)\cdot t_{in}}$ to SI_{t_0}, are the same as Fig. 1(h).

B. DIPT-Based Sky Image Prediction Models

PIV first divides both input images into small block regions, and then computes the cloud displacement vector by matching these blocks [15]. For FPCT, it uses fast Fourier transform to realize the interconversion between image space domain and frequency domain [17]. Fourier transform of two input images is calculated and then used to calculate the cross-power spectrum. Finally, the cloud displacement vector is obtained by the inverse Fourier Transform of the cross power spectrum.

PIV and FPCT are used to construct two determined models which generate a predictive image by linear extrapolation of the image at t_0. The type of models can be described as follows:

$$SI_{t_0+t_{\mathrm{out}},DIPT} = f_{DIPT}([SI_{t_0-t_{in}}, SI_{t_0}]) \tag{5}$$

where $SI_{t_0+t_{\mathrm{out}},DIPT}$ is the predicted sky image generated by a DIPT based method; $f_{DIPT}(\cdot)$ is a DIPT based sky image prediction model; besides, $n = 2$, $t_{\mathrm{out}} = t_{in}$.

The flowchart of a DIPT based sky image prediction model is demonstrated as Fig. 5. For DIPT models, there are a few things to note. The first is that $SI_{t_0-t_{\mathrm{out}}}$ and SI_{t_0} are high quality images like Fig. 1(g), and both of them are used to predict a near future

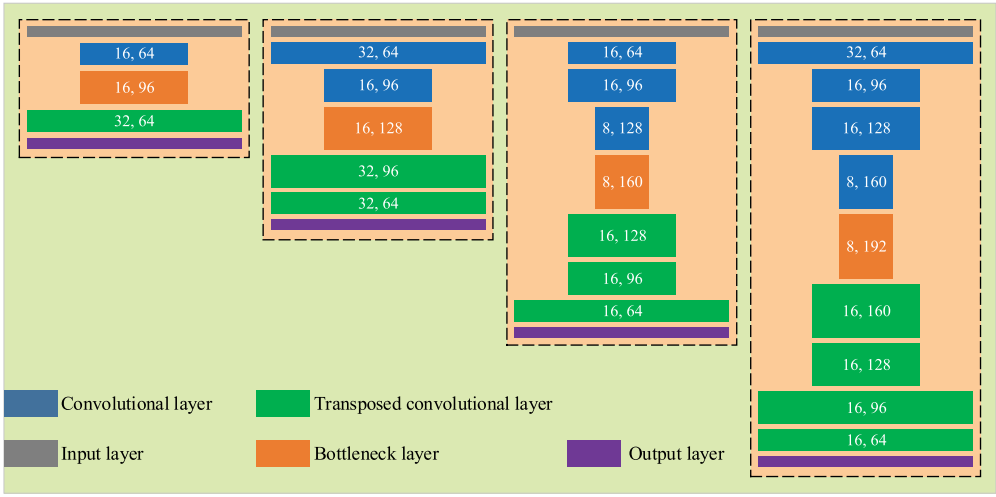


Fig. 6. Four architectures of CAEs.

frame. The second is that a predicted frame $SI_{t_0+\mathrm{t_{out}},DIPT}$ is a clipped image due to linear image extrapolation, and the same operation is done to its ground truth image $SI_{t_0+\mathrm{t_{out}}}$. Therefore their sizes are little less than SI_{t_0}. The image clipping process can be referred to [18]. The third point is that $SI_{t_0+\mathrm{t_{out}},DIPT}$ and $SI_{t_0+\mathrm{t_{out}}}$ are be downsampled to the size of 32×32 for similarity assessment.

IV. Simulation

A. Assessment Metrics

Structural similarity (SSIM) and mean square error (MSE) are introduced to evaluate the performances of sky image prediction models. For two images, SSIM whose value range is [0, 1] [29] is used to measure the similarity of structural information; while MSE is used to measure the similarity of gray values. The larger SSIM of two images means that their structures are more similar; the smaller MSE means that their pixels values are more similar.

If x and y are two images with same resolution of $M \times N$, their SSIM is the product of luminance comparison $l(x,y)$, the contrast comparison $c(x,y)$ and the structure comparison $s(x,y)$, the three comparisons are expressed as

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{6}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{7}$$

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \tag{8}$$

where μ_x and μ_y, respectively, represent the mean value of all the pixels in x and y, σ_x, and σ_y respectively represent the standard deviation, σ_{xy} is covariance between x and y. C_1, C_2 and C_3 are three constants to avoid denominator very close to zero, and they can be respectively described as follows: $C_1 = (K_1L)^2$, $C_2 = (K_2L)^2$ and $C_3 = C_2/2$. In general, $K_1 = 0.01$, $K_2 = 0.03$ and $L = 255$. Finally, SSIM is described as

$$SSIM(x,y) = l(x,y) \cdot c(x,y) \cdot s(x,y). \tag{9}$$

For x and y, MSE is described as follows:

$$MSE(x,y) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[x(i,j) - \mathrm{y}(i,j)]^2 \tag{10}$$

where $x(i,j)$ and $\mathrm{y}(i,j)$ represent the pixel values of coordinate (i,j) in x and y, respectively.

B. Simulation Description

All CAE models were built by Spyder and Keras, and they were run on a computer equipped with the Linux operating system, i7 CPU, 16 GB RAM, and a NVIDIA GeForce GTX 1080Ti GPU. All DIPT models were built by Spyder, OpenCV and OpenPIV, and they were run on a computer equipped with the Window10 operating system, i5 CPU, and 12 GB RAM.

In this article, the four different structures are set for 2-D CAEs and 3-D CAEs, respectively, as shown in Fig. 6. The number of hidden layers (NHL) of each structure is 3, 5, 7, and 9, respectively. There are pairs of numbers in Fig. 6. For example, in the first hidden layer of a three-layer 2-D CAE, 16 is its output feature map side length, and 64 is the number of its output feature maps.

For a sky image prediction model, the mean value on the test set will be used as its performance evaluation. Since Adam optimizer performing a random gradient descent is used during the training process [30], each evaluation of a CAE is slightly different. To mitigate this randomness, each CAE is evaluated ten times in each situation, the median of ten evaluation values is taken as the performance evaluation value. Furthermore, each PIV-based or FPCT-based sky image prediction model is evaluated only once in each situation due to its determination. Two types of models in this article predict an image. As can be seen from Table I, five prediction scales are set in this article, each corresponding to one time resolution of the predicted images.

TABLE I
INPUT AND OUTPUT SETTINGS OF THE TWO TYPES OF MODELS SCENARIO

Scenario Number	CAE based models (n=10)		DIPT based models (n=2)	
	t_{in}/min	t_{out}/min	t_{in}/min	t_{out}/min
1	0.5	0.5	0.5	0.5
2	0.5	1	1	1
3	0.5	2	2	2
4	0.5	3	3	3
5	0.5	4	4	4

TABLE II
SSIM VALUES OF TWO TYPES OF MODELS

t_{out}/min	Best 2D CAEs		Best 3D CAEs		PIV	FPCT
	Values	*NHL*	*Values*	*NHL*	*Values*	*Values*
0.5	**0.9953**	5	**0.9953**	5	0.9830	0.9831
1	0.9889	5	**0.9890**	5	0.9584	0.9586
2	**0.9692**	5	0.9678	5	0.9244	0.9245
3	**0.9475**	5	0.9444	5	0.9039	0.9038
4	**0.9291**	3	0.9259	5	0.8895	0.8895

TABLE III
MSE VALUES OF TWO TYPES OF MODELS

t_{out}/min	Best 2D CAEs		Best 3D CAEs		PIV	FPCT
	Values	*NHL*	*Values*	*NHL*	*Values*	*Values*
0.5	**5.12**	5	5.22	5	7.19	7.18
1	14.14	3	15.24	3	12.32	**12.30**
2	**11.84**	3	12.12	3	18.23	18.24
3	15.90	3	**15.88**	3	21.91	21.83
4	**15.07**	3	15.17	3	24.41	24.39

C. Results and Discussion

The results of SSIM and MSE from the two types of models are given in Tables II and III, respectively, where bold numbers represent the optimal values in each situation. The ratios of optimal evaluations between CAE models and DIPT models in five scenarios are compared, as shown in Fig. 7. The comparison results of SSIM and MSE from the two types of models are shown in Figs. 8 and 9, respectively. From these two Tables and three figures, we can find some results.

The first is that the prediction performance of the two types of models deteriorates with the decrease of output time resolution by and large. The reason is that as the output time resolution decreases, the spatiotemporal correlation between input and output also decreases, which makes the prediction difficult and finally leads to the obvious deterioration.

Second, as shown in Fig. 7, for SSIM ratios, each value is greater than 1, which indicates that with the decrease of output time resolution, the ability of CAE models to predict image

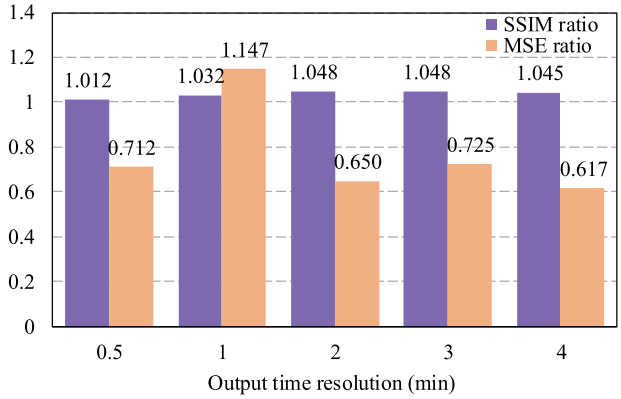


Fig. 7. Ratios of optimal evaluations between two types of models.

structure information is always better than that of DIPT models. For instance, the SSIM of CAE models is 1.2% higher than that of DIPT models when $t_{\text{out}} = 0.5$ min, while the SSIM value of CAE models is 4.5% higher than that of DIPT models when $t_{\text{out}} = 4$ min. For MSE ratios, except for $t_{\text{out}} = 1$ min, the values are significantly less than 1, which indicates that the ability of CAE models to predict the gray information of image pixels is basically stronger than that of DIPT models. For instance, when $t_{\text{out}} = 0.5$ min, the MSE of CAE models is 71.2% of that of DIPT models, while that of CAE model is 61.7% when$t_{\text{out}} = 4$ min. According to the analysis of the second results, we can find that when increasing the input image sequence length and introducing the relatively powerful nonlinear relationship fitting capacity by CAEs, the sky image prediction can be improved significantly in given cases. In other words, CAE models are basically superior to DIPT models regarding 5 scenarios.

The third is that 2-D CAEs and 3D CAEs are close regarding best performance values, for example, when $t_{\text{out}} = 1$ min in Table II and $t_{\text{out}} = 3$ min in Table III. However, in different structures under different situations, the best performances of 2-D CAEs are basically better than that of 3-D CAEs. This result indicates that although 3-D convolution has finer spatiotemporal feature extraction than 2-D convolution, complex operations may not lead to better performance.

The fourth is that due to fluctuating performances, it is necessary to train CAE models for many times and compare the performances of different structures to find better models, and this is exactly the inherent disadvantage of such stochastic learning models. Besides, DIPT models get very close results, and their curves basically overlap in Figs. 8 and 9.

As depicted in Tables II and III, the best models are 2-D CAEs containing five and three hidden layers, respectively, when $t_{\text{out}} = 0.5$ min and $t_{\text{out}} = 4$ min, but the optimal model cannot be judged directly from the values of SSIM and MSE in the remaining three situations. However, the relatively good models in the remaining three situations can be obtained. The above analysis is expected to provide reference for future research.

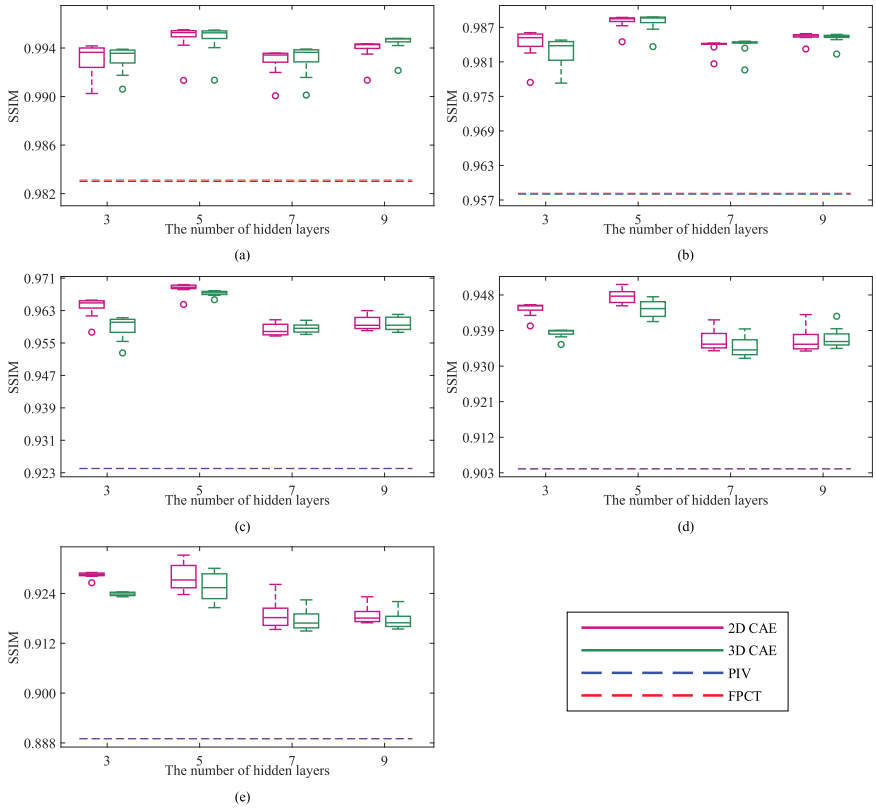


Fig. 8. SSIM comparison of the two types of models. (a) SSIM values when $t_{\mathrm{out}} = 0.5$ min. (b) SSIM values when $t_{\mathrm{out}} = 1$ min. (c) SSIM values when $t_{\mathrm{out}} = 2$ min. (d) SSIM values when $t_{\mathrm{out}} = 3$ min. (e) SSIM values when $t_{\mathrm{out}} = 4$ min.

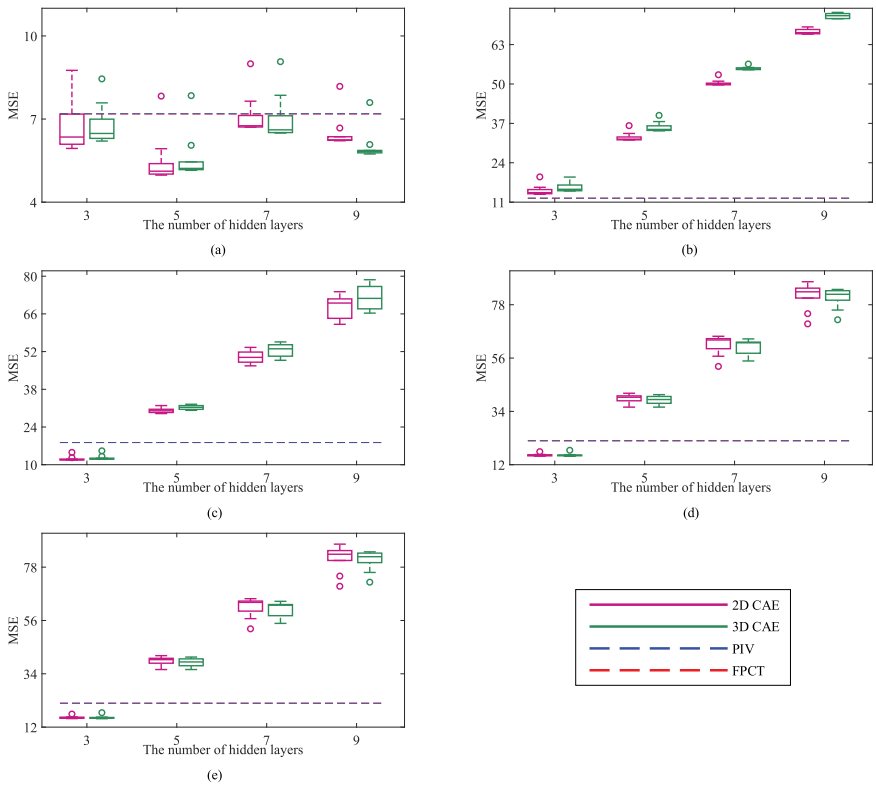


Fig. 9. MSE comparison of the two types of models. (a) MSE values when $t_{\mathrm{out}} = 0.5$ min. (b) MSE values when $t_{\mathrm{out}} = 1$ min. (c) MSE values when $t_{\mathrm{out}} = 2$ min. (d) MSE values when $t_{\mathrm{out}} = 3$ min. (e) MSE values when $t_{\mathrm{out}} = 4$ min.

V. Conclusion

As the dynamic cloud distribution can cause significant generation fluctuation in a very short time, it is very important to achieve accurate sky image prediction for the minute time scale output forecasting of PV power station. Recently, most related works use DIPT to achieve cloud distribution prediction. However, DIPT uses relatively limited spatiotemporal information and linear image extrapolation, which may lead to inaccurate prediction result of cloud. To avoid these shortages of DIPT-based sky image prediction methods, 2-D CAE-based and 3-D CAE-based sky image prediction models are proposed. In this article, four structures are designed for CAE-based models, and the benchmark models are built by using two typical DIPT methods: PIV and FPCT. Besides, five different output time resolutions are considered. Through simulation, it can be seen that whether it is to predict the image structure information or the image pixels, CAE models are basically better than DIPT models, and 2-D CAE models are basically better than 3-D CAE models. It can be seen that 2-D CAEs have the best image prediction ability for spatiotemporal information due to comparatively long input image sequence length and powerful nonlinear relationship fitting capacity.

The above research cloud also be expected to support further studies in irradiance and PV power forecasting, such as building a mapping relationship from a predicted cloud distribution or its features to the corresponding surface irradiance to realize irradiance forecasting.

In fact, the impacts caused by deep penetration of renewable energy on generation flexibility and operating costs of power grid [31] are significant and reflect in many aspects not only include power forecasting [32]–[37] for supply-demand balancing and energy trading under various scenarios [38]–[40], but also refer to load forecasting/load pattern [41]–[44], demand response applications [45]–[48], aggregator aggregated capacity forecasting and multi-aggregator scheduling with plenty distributed PV systems [49]–[52]. The abovementioned research topics will be conducted in the future works.

References

[1] Installed Capacity Trends, IRENA. [Online]. Available: https://www.irena.org/solar</b, Accessed: Oct. 1, 2019.

[2] Y. Wu, S. Chang, and P. Mandal, "Grid-connected wind power plants: A survey on the integration requirements in modern grid codes," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 5584–5593, Nov./Dec. 2019.

[3] Y. Wu, J. Lin, and H. Lin, "Standards and guidelines for grid-connected photovoltaic generation systems: A review and comparison," *IEEE Trans. Ind. Appl.*, vol. 53, no. 4, pp. 3205–3216, Jul./Aug. 2019.

[4] T. Khalili, A. Jafari, M. Abapour, and B. Mohammadi-Ivatloo, "Optimal battery technology selection and incentive-based demand response program utilization for reliability improvement of an insular microgrid," *Energy*, vol. 169, pp. 92–104, Feb. 2019.

[5] F. Wang, Z. Zhen, B. Wang, and Z. Mi, "Comparative study on KNN and SVM based weather classification models for day ahead short term solar PV power forecasting," *Appl. Sci.*, vol. 8, no. 2, p. 28, Dec. 2017, doi: 10.3390/app8010028.

[6] F. Wang *et al.*, "Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting," *Int. J. Elect. Power Energy Syst.*, vol. 105, pp. 529–540, Feb. 2019.

[7] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, "A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework," *Energy Convers. Manage*, vol. 212, May. 2020, Art. no. 112766.

[8] Z. Zhen *et al.*, "Deep learning based surface irradiance mapping model for solar PV power forecasting using sky image," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3385–3396, Jul./Aug. 2020.

[9] F. Wang *et al.*, "A minutely solar irradiance forecasting method based on real-time sky image-irradiance mapping model," *Energy Convers. Manage*, vol. 220, Sep. 2020, Art. no. 113075.

[10] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016.

[11] J. Zhang, R. Verschae, S. Nobuhara, and J. Lalonde, "Deep photovoltaic nowcasting," *Sol. Energy*, vol. 176, pp. 267–276, Dec. 2018.

[12] Z. Zhen, Z. Wang, F. Wang, Z. Mi, and K. Li, "Research on a cloud image forecasting approach for solar power forecasting," *Energy Procedia*, vol. 142, pp. 362–368, Dec. 2017.

[13] M. Lourenco, J. P. Barreto, and F. Vasconcelos, "sRD-SIFT: Keypoint detection and matching in images with radial distortion," *IEEE Trans. Robot.*, vol. 28, no. 3, pp. 752–760, Jun. 2012.

[14] S. Dev, F. M. Savoy, Y. H. Lee, and S. Winkler, "Short-term prediction of localized cloud motion using ground-based sky imagers," in *Proc. IEEE Region 10 Conf.*, 2016, pp. 2563–2566.

[15] A. Zaher, S. Thil, J. Nou, A. Traoré, and S. Grieu, "Comparative study of algorithms for cloud motion estimation using sky-imaging data," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 5934–5939, Jul. 2017.

[16] Z. Zhen, Z. Xuan, F. Wang, R. Sun, N. Duić, and T. Jin, "Image phase shift invariance based multi-transform-fusion method for cloud motion displacement calculation using sky images," *Energy Convers. Manage*, vol. 197, Oct. 2019, Art. no. 111853.

[17] F. Wang *et al.*, "Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting," *Energy Convers. Manage*, vol. 157, pp. 123–135, Feb. 2018.

[18] Z. Zhen *et al.*, "Pattern classification and PSO optimal weights based sky images cloud motion speed calculation method for solar PV power forecasting," *IEEE Trans. Ind. Appl.*, vol. 55, no. 4, pp. 3331–3342, Jul./Aug. 2019.

[19] S. Aigner and M. Körner, "FutureGAN: Anticipating the future frames of video sequences using spatio-temporal 3d convolutions in progressively growing GANs," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 42, no. 2, pp. 3–11, Jan. 2018.

[20] M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," in *Proc 4th Int. Conf. Learn. Represent. ICLR 2016 - Conf. Track Proc.*, Feb. 2016, pp. 1–14.

[21] Z. W. Liu, R. A. Yeh, X. O. Tang, Y. M. Liu, and A. Agarwala, "Video frame synthesis using deep voxel flow," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 4473–4481.

[22] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018.

[23] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[24] Nat. Ocean. Atmos. Admin, Washington, DC, USA, 2019, Accessed: May 1, 2019. [Online]. Available: https://www.noaa.gov/</b

[25] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," pp. 1–31, 2016.

[26] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3D convolutional networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 4489–4497.

[27] S. Ioffe, and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *J. Pract.*, vol. 10, no. 6, pp. 730–743, 2015.

[28] "Rectifier nonlinearities improve neural network acoustic models," [Online]. Available: https://ai.stanford.edu/∼amaas/papers/relu_hybrid_icml2013_final.pdf</b, Accessed: Oct. 1, 2019.

[29] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, May 2004.

[30] D. P. Kingma, and J. L. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. ICLR 2015-Conf. Track Proc.*, Jan. 2017, pp. 1–15.

[31] Y. Wu, W. Tan, S. Huang, Y. Chiang, C. Chiu, and C. Su, "Impact of generation flexibility on the operating costs of the taiwan power system under a high penetration of renewable power," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, pp. 2348–2359, May/Jun. 2020.

[32] H. Chai *et al.*, "Convolutional Auto-encoder based sky image prediction model for minutely solar PV power forecasting," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, 2020, pp. 1–7.

[33] Y. Wu, P. Sun, T. Wu, J. Hong, and M. Y. Hassan, "Probabilistic wind-power forecasting using weather ensemble models," *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 5609–5620, Nov./Dec. 2018.
[34] M. Yang, S. Fan, and W. Lee, "Probabilistic short-term wind power forecast using componential sparse bayesian learning," *IEEE Trans. Ind. Appl.*, vol. 49, no. 6, pp. 2783–2792, Nov./Dec. 2013.
[35] J. Ma, M. Yang, X. Han, and Z. Li, "Ultra-short-term wind generation forecast based on multivariate empirical dynamic modeling," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1029–1038, Mar./Apr. 2018.
[36] M. Bozorg *et al.*, "Bayesian bootstrap quantile regression for probabilistic photovoltaic power forecasting," *Prot. Control Mod. Power Syst.*, vol. 5, no. 1, p. 21, Dec. 2020, doi: 10.1186/s41601-020-00167-7.
[37] M. Madhiarasan, "Accurate prediction of different forecast horizons wind speed using a recursive radial basis function neural network," *Prot. Control Mod. Power Syst.*, vol. 5, no. 1, p. 22, Dec. 2020, doi: 10.1186/s41601-020-00166-8.
[38] Z. Zhang, H. Tang, P. Wang, Q. Huang, and W. Lee, "Two-stage bidding strategy for peer-to-peer energy trading of nanogrid," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1000–1009, Mar./Apr. 2020.
[39] H. Xu, Y. Lin, X. Zhang, and F. Wang, "Power system parameter attack for financial profits in electricity markets," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3438–3446, Jul. 2020.
[40] Z. Ding, W. Lee, and J. Wang, "Stochastic resource planning strategy to improve the efficiency of microgrid operation," *IEEE Trans. Ind. Appl.*, vol. 51, no. 3, pp. 1978–1986, May/Jun. 2015.
[41] X. Wang, W. Lee, H. Huang, R. L. Szabados, Y. Wang, and P. V. Olinda, "Factors that impact the accuracy of clustering-based load forecasting," *IEEE Trans. Ind. Appl.*, vol. 52, no. 5, pp. 3625–3630, Sep./Oct. 2016.
[42] K. Li *et al.*, "Meta-heuristic optimization based two-stage residential load pattern clustering approach considering intra-cluster compactness and inter-cluster separation," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3375–3384, Jul./Aug. 2020.
[43] F. Wang *et al.*, "Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns," *Energy Convers. Manage*, vol. 171, pp. 839–854, Sep. 2018.
[44] S. Yan *et al.*, "Time-frequency features combination-based household characteristics identification approach using smart meter data," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, pp. 2251–2262, May/Jun. 2020.
[45] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalao, "Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6972–6985, Nov. 2018.
[46] K. Li *et al.*, "Impact factors analysis on the probability characterized effects of time of use demand response tariffs using association rule mining method," *Energy Convers. Manage*, vol. 197, Oct. 2019, Art. no. 111891.
[47] X. Lu, K. Li, H. Xu, F. Wang, Z. Zhou, and Y. Zhang, "Fundamentals and business model for resource aggregator of demand response in electricity markets," *Energy*, vol. 204, May 2020, Art. no. 117885.
[48] F. Wang *et al.*, "Day-ahead optimal bidding and scheduling strategies for DER aggregator considering responsive uncertainty under real-time pricing," *Energy*, vol. 213, Dec. 2020, Art. no. 118765.
[49] F. Wang *et al.*, "Smart households' aggregated capacity forecasting for load aggregators under incentive-based demand response programs," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1086–1097, Mar./Apr. 2020.
[50] V. Gupta, R. Kumar, and B. K. Panigrahi, "User willingness based decentralized EV charging management in multi-aggregator scheduling," *IEEE Trans. Ind. Appl.*, vol. 56, no. 5, pp. 5704–5715, Sep./Oct. 2020.
[51] Q. Chen *et al.*, "Dynamic price vector formation model-based automatic demand response strategy for PV-assisted EV charging stations," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2903–2915, Nov. 2017.
[52] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duić, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Appl. Energy*, vol. 253, Nov. 2019, Art. no. 113595.

Yuwei Fu received the B.S. degree in electrical engineering from Nanchang University, Jiangxi, China, in 2019. She is currently working toward the M.S. degree at the Department of Electrical Engineering at North China Electric Power University, Baoding, China.

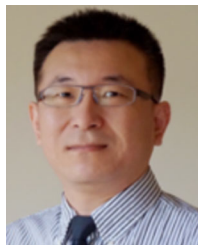
Her research interests include power forecasting, demand response and smart grid.

Hua Chai received the B.S. degree in electrical engineering and automation from Inner Mongolia University of Technology, Huhhot, China, in 2018. He is currently working toward the M.S. degree at the Department of Electrical Engineering at North China Electric Power University, Baoding, China.

His research interest is solar power forecasting.

Zhao Zhen (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from North China Electric Power University (NCEPU), Baoding, China, in 2012 and 2018, respectively.

He is currently an Assistant Professor with the Department of Electrical Engineering, NCEPU, Baoding, China, and also a Postdoc Researcher with the Department of Electrical Engineering, Tsinghua University, Beijing, China. His research interests include solar irradiance and PV power forecasting, wind speed and power forecasting.



Fei Wang (Senior Member, IEEE) received the B.S. degree from Hebei University, Baoding, China in 1993, the M.S. and Ph.D. degrees in electrical engineering from North China Electric Power University (NCEPU), Baoding, China, in 2005 and 2013, respectively.

He is currently a Professor with the Department of Electrical Engineering, NCEPU and the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, Baoding and Beijing, China. He is the Director of Smart Energy Network Integrated Operation Research Center and the Leader of "Double First-Class" research team project with NCEPU. He was a Visiting Professor with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA, from 2016 to 2017. He was a Researcher with the Department of Electrical Engineering, Tsinghua University, Beijing, China, from 2014 to 2016. He supervised more than 60 Post-docs, Ph.D. and M.Sc. students. He has authored or coauthored more than 220 publications, including 80 journal papers. His research interests include renewable energy power, electricity price and electricity load forecasting, demand response and electricity market, smart grid, microgrid, and integrated energy system.

Dr. Wang is an Editor for the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, an Editor for the IEEE POWER ENGINEERING LETTERS, an Associate Editor for the *IET Renewable Power Generation*, and an Editor of the IEEE OPEN ACCESS JOURNAL OF POWER AND ENERGY AND PROTECTION AND CONTROL OF MODERN POWER SYSTEMS. He was the Guest Editor for the Special Issue on "Demand Side Management and Market Design for Renewable Energy Support and Integration" of *IET Renewable Power Generation*. He is the Expert Member of IEC SC8A/WG2. He was the recipient of the 2020 Science and Technology Progress First Award of Hebei Province, the 2018 Technical Invention First Award of Hebei Province, the 2018 Patent Third Award of Hebei Province, the 2014 Natural Sciences Academic Innovation Achievement Award of Hebei Province, the 2018 China Electric Power Science and Technology Progress Third Award, and the 2014 Outstanding Doctoral Dissertation Award of NCEPU. He was the General Chair of the 2017 International Seminar of Renewable Energy Power Forecasting and Absorption Technology and 2018 International Seminar of Integrated Energy and Smart Microgrid Technology.

Xunjian Xu received the B.S. and M.S. degrees in electrical engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2005 and 2007, respectively, and the Ph.D. degree in electronic and photonic systems engineering at Kochi University of Technology, Kochi, Japan, in 2010.

He is currently a Professor Level Senior Engineer with the State Key Laboratory of Disaster Prevention and Reduction for Power Grid Transmission and Distribution Equipment, Changsha, China, and also a Postdoc Researcher with the Department of Electrical Engineering, Hunan University. His research interests include power transmission line icing forecasting, PV power forecasting, wind speed and power forecasting.

Kangping Li (Member, IEEE) received the B.S. degree in electrical engineering in 2015 from North China Electric Power University, Baoding, China, where he is currently working toward the Ph.D. degree with the Department of Electrical Engineering.

His research interests include demand response, electricity market and power system optimization.

Miadreza Shafie-Khah (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from Tarbiat Modares University, Tehran, Iran, in 2008 and 2012, respectively.

He was a Postdoc from the University of Beira Interior, Covilha, Portugal, in 2015. He was a Postdoc with the University of Salerno, Salerno, Italy, in 2016. He is currently an Associate Professor with the University of Vaasa, Vaasa, Finland. His research interests include power market simulation, market power monitoring, power system optimization, demand response, electric vehicles, price forecasting and smart grids.

Dr. Shafie-Khah was considered one of the Outstanding Reviewers of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, in 2014 and 2017, one of the Best Reviewers of the IEEE TRANSACTIONS ON SMART GRID, in 2016 and 2017, and one of the Outstanding Reviewers of the IEEE TRANSACTIONS ON POWER SYSTEMS, in 2017 and 2018.

Payman Dehghanian (Senior Member, IEEE) received the B.Sc. degree from the University of Tehran, Tehran, Iran, in 2009, the M.Sc. degree from Sharif University of Technology, Tehran, Iran, in 2011, and the Ph.D. degree from Texas A&M University, College Station, TX, USA in 2017, all in electrical engineering.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering in George Washington University, Washington, DC, USA. His research interests include power system online situational awareness, real-time decision making, power system reliability and resiliency, asset management, and smart electricity grid applications.

Dr. Dehghanian is the recipient of the 2013 IEEE Iran Section Best M.Sc. Thesis Award in Electrical Engineering, the 2014 and 2015 IEEE Region five Outstanding Professional Achievement Awards, and the 2015 IEEE-HKN Outstanding Young Professional Award.

João P. S. Catalão (Senior Member, IEEE) received the M.Sc. degree from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2003, and the Ph.D. degree and Habilitation for Full Professor ("Agregação") from the University of Beira Interior (UBI), Covilha, Portugal, in 2007 and 2013, respectively.

He is currently a Professor with the Faculty of Engineering of the University of Porto (FEUP), Porto, Portugal, and a Research Coordinator at INESC TEC. He was a Visiting Professor with North China Electric Power University, Beijing, China. He was the Primary Coordinator of the EU-funded FP7 project SiNGULAR ("Smart and Sustainable Insular Electricity Grids Under Large-Scale Renewable Integration"), a 5.2-million-euro project involving 11 industry partners. He has authored or coauthored more than 700 publications, including 275 journal papers (more than 80 IEEE Transactions/Journal papers), 372 conference proceedings papers, 5 books, 35 book chapters, and 14 technical reports, with an h-index of 47, an i10-index of 197, and over 8500 citations (according to Google Scholar), having supervised more than 70 post-docs, Ph.D. and M.Sc. students. He is the Editor of the books entitled *Electric Power Systems: Advanced Forecasting Techniques and Optimal Generation Scheduling* and *Smart and Sustainable Power Systems: Operations, Planning and Economics of Insular Electricity Grids* (Boca Raton, FL, USA: CRC Press, 2012 and 2015, respectively). His research interests include power system operations and planning, hydro and thermal scheduling, wind and price forecasting, distributed renewable generation, demand response and smart grids.

Dr. Catalão is an Editor of the IEEE TRANSACTIONS ON SMART GRID, an Editor of the IEEE TRANSACTIONS ON POWER SYSTEMS, an Associate Editor of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS. From 2011 to 2018 (seven years), he was an Editor for the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY and an Associate Editor for the *IET Renewable Power Generation*. He was also a Subject Editor for the *IET Renewable Power Generation*. He was the Guest Editor-in-Chief for the Special Section on "Real-Time Demand Response" of the IEEE TRANSACTIONS ON SMART GRID, published in December 2012, the Guest Editor-in-Chief for the Special Section on "Reserve and Flexibility for Handling Variability and Uncertainty of Renewable Generation" of the IEEE TRANSACTIONS ON SUSTAINABLE ENERGY, published in April 2016, the Corresponding Guest Editor for the Special Section on "Industrial and Commercial Demand Response" of the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, published in November 2018, and the Lead Guest Editor for the Special Issue on "Demand Side Management and Market Design for Renewable Energy Support and Integration" of the *IET Renewable Power Generation*, to be published in March 2019. He was the recipient of the 2011 Scientific Merit Award UBI-FE/Santander Universities, the 2012 Scientific Award UTL/Santander Totta, the 2016 and 2017 FEUP Diplomas of Scientific Recognition, the 2017 Best INESC-ID Researcher Award, and the 2018 Scientific Award ULisboa/Santander Universities, in addition to an Honorable Mention in the 2017 Scientific Award ULisboa/Santander Universities. Moreover, he has won four Best Paper Awards at IEEE Conferences.