Occlusion-Perturbed Deep Learning for Probabilistic Solar Forecasting via Sky Images

Preprint

Cong Feng, Wenqi Zhang, Bri-Mathias Hodge, and Yingchen Zhang

National Renewable Energy Laboratory

Presented at 2022 IEEE Power & Energy Society General Meeting
Denver, Colorado
July 17-21, 2022

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-83163
September 2022

Occlusion-Perturbed Deep Learning for Probabilistic Solar Forecasting via Sky Images

Preprint

Cong Feng, Wenqi Zhang, Bri-Mathias Hodge, and Yingchen Zhang

National Renewable Energy Laboratory

Suggested Citation
Feng, Cong, Wenqi Zhang, Bri-Mathias Hodge, and Yingchen Zhang. 2022. Occlusion-Perturbed Deep Learning for Probabilistic Solar Forecasting via Sky Images: Preprint. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5D00-83163. https://www.nrel.gov/docs/fy22osti/83163.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5D00-5D00-83163
September 2022

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. This work was supported by the Laboratory Directed Research and Development (LDRD) Program at NREL. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

Occlusion-perturbed Deep Learning for Probabilistic Solar Forecasting via Sky Images

Cong Feng, Wenqi Zhang, Bri-Mathias Hodge, Yingchen Zhang
National Renewable Energy Laboratory
{Cong.Feng, Flora.Zhang, Bri.Mathias.Hodge, Yingchen.Zhang}@nrel.gov

***Abstract*—Solar forecasting is shifting to the probabilistic paradigm due to the inherent uncertainty within the solar resource. Input uncertainty quantification is one of the widely-used and best-performing ways to model solar uncertainty. However, compared to other sources of inputs, such as numerical weather prediction models, pure sky image-based probabilistic solar forecasting lags behind. In this research, an occlusion-perturbed convolutional neural network, named the PSolarNet, is developed. The PSolarNet provides very short-term deterministic forecasts, forecast scenarios, and probabilistic forecasts of the global horizontal irradiance from sky image sequences. Case studies based on 6 years of open-source data show that the developed PSolarNet is able to generate accurate 10-minute ahead deterministic forecasts with a 5.62% normalized root mean square error, realistic and diverse forecast scenarios with a 0.966 average correlation with the actual time series, and reliable and sharp probabilistic forecasts with a 2.77% normalized continuous ranked probability score.**

***Index Terms*—Deep learning, solar forecasting, sky image processing, Bayesian model averaging**

I. Introduction

With the increasing penetration of solar photovoltaics (PV) in power systems, there is a growing need for accurate and reliable solar forecasts. Due to the inherent uncertainty of solar power (e.g., cloud motion, meteorological variation, climate change), any solar forecast is erroneous. Therefore, research interest in the solar forecasting domain is shifting from deterministic forecasts toward probabilistic forecasts represented by predictive distributions, intervals, quantiles, etc. Compared to deterministic solar forecasts, probabilistic forecasts are able to quantify forecast uncertainty, which can be utilized in various power system problems, such as probabilistic power flow, probabilistic optimal power flow, and stochastic optimization [1]. However, probabilistic solar forecasting is still at an early stage compared to the development of deterministic forecasting techniques.

Probabilistic solar forecasting techniques can be categorized by different aspects. The most commonly-used categorization is by the existence of a pre-assumed distribution. One group of methods that assume a predictive distribution beforehand are called parametric methods. For example, a normal distribution was used to fit the density function around deterministic forecasts in Ref. [2]. An unavoidable drawback of this category of method is the lack of a universally optimal distribution that could quantify the solar uncertainty. Therefore, the second group of methods, called the nonparametric methods, dominate the research, and include methods such as quantile regression, kernel density estimation, bootstrapping, gradient boosting, and various analog ensemble methods [3].

Most nonparametric methods seek to represent uncertainties by a (sub)set of the inputs, the models, or the output, which is another interesting way to classify probabilistic solar forecasting methods–by the uncertainty source. The most popular way to represent the forecasting probability is by modeling the input uncertainty. For example, most numerical weather prediction systems provide ensemble forecasts, which add perturbations to the initial boundary conditions [4], which can be post-processed by various methods to generate probabilistic forecasts [5]. Other input uncertainty modeling includes meteorological time series generation using statistical methods. For example, weather parameter correlations were modelled by a Copula joint distribution [6]. To the best of our knowledge, among various input data, solar forecast scenario generation using sky images has not yet been investigated before.

To bridge the aforementioned gaps in probabilistic solar forecasting, we propose a deep learning probabilistic solar forecasting network, called the Probabilistic SolarNet (PSolarNet). The PSolarNet provides very short-term deterministic forecasts, forecast scenarios, and probabilistic forecasts of the global horizontal irradiance (GHI) from sky image sequences. The contributions of this paper include: (i) developing an accurate convolutional neural network (CNN) for deterministic solar forecasting via sky image input, (ii) generating realistic and diverse forecast scenarios from sky image perturbation, (iii) providing reliable and sharp probabilistic forecasts by modeling the sky image input uncertainty using Bayesian model averaging (BMA).

The remainder of this paper is organized as follows. Section II introduces three components of the developed PSolarNe. The case study dataset and benchmark models are described in Section III. Section IV discusses the results. Conclusions are made in Section V.

This work was authored in part by the National Renewable Energy Laboratory (NREL), operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract NO. DE-AC36-08GO28308. The funding is provided through an NREL's LDRD program called the Director's Postdoctoral Fellowships.

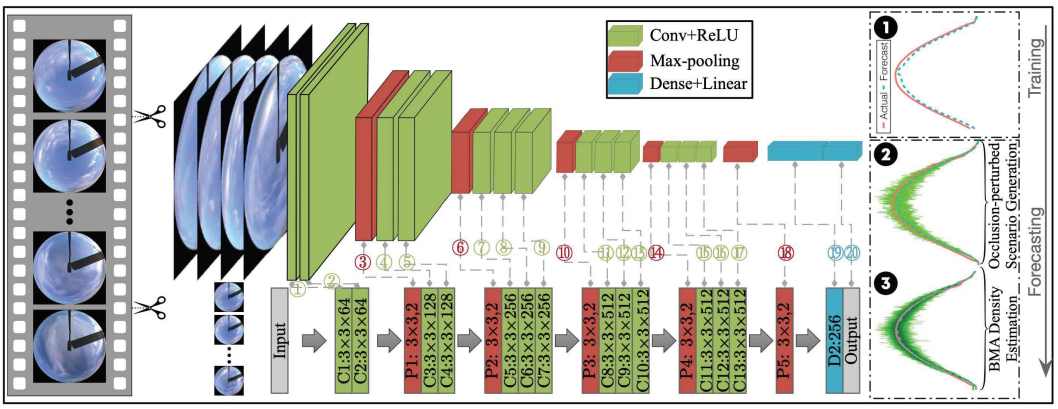


Figure 1. The overall structure of the Probabilistic SolarNet (PSolarNet). Colors indicate different layer types.

II. The Probabilistic SolarNet

In this research, we seek to provide probabilistic 10-minute ahead GHI forecasts using the method we have named PSolarNet. The forecasting problem has a 10-minute lead time with a 10-minute resolution and a 10-minute update rate [7]. The overall framework of the PSolarNet is shown in Fig. 1, where the core forecasting engine is a deterministic SolarNet. There are mainly three stages (indicated by numbers) to generate probabilistic forecasts by PSolarNet: (i) forecasting model training, which trains a SolarNet for deterministic GHI forecasting; (ii) occlusion-perturbed scenario generation, which generates a set of deterministic GHI forecasting scenarios by sky image occlusion perturbations; (iii) BMA probability density estimation, which estimates probability density functions (PDFs) from deterministic GHI forecast scenarios. In the rest of this section, we will further describe these three stages.

A. SolarNet

A SolarNet is composed of 5 feature learning blocks, as shown in the bottom part of Fig. 1. Each block stacks 2 or 3 convolutional layers and a max-pooling layer. The SolarNet contains small filters to extract more receptive fields with local spatial correlations from the input tensors. The filter amount of the next block is doubled to extract more abstract and informative features. The SolarNet forecasts the GHI, $\mathbf{y} \in \mathbb{R}^{N\times 1}$, from the stacked sky image sequence, $\mathbf{X} \in \mathbb{R}^{N\times W\times H\times D}$, without numerical measurements and feature engineering [8].

At the training step, actual GHI and the corresponding sky images are used to guide the model parameter optimization, $\mathbf{y} = F(\mathbf{X}, \mathbf{W})$, where N, W, H, D are the sample size, image width, image height, and image channel amount $\times$ the image sequence length, respectively; $\mathbf{y}$ is the actual GHI; $\mathbf{X}$ is the image input; $\mathbf{W}$ is the trainable parameters in the SolarNet; F is the SolarNet function. To perform GHI forecasting, optimal parameters should be obtained first at the training stage. Since forecasting is a regression problem, we use L1 loss (i.e., the mean absolute error) to respect the overall performance of the model [8]: $J(\mathbf{W}) = -\frac{1}{N}\sum_{n=1}^{N}|\hat{y}_n - y_n|$, where n is the sample index; $\hat{y}$ and y are forecast and actual GHI values at a certain time, respectively. Then, the objective in the training process is to optimize the parameters by minimizing the loss function in an end-to-end manner (i.e., from images to GHI values). Once the optimal model parameters, $\mathbf{W}^*$, are obtained, the well-trained deep learning models are used to generate forecasts in the forecasting/testing stage: $\hat{\mathbf{y}} = F^*(\mathbf{X}, \mathbf{W}^*)$, where $\hat{\mathbf{y}}$ is the GHI forecast vector; F^* indicates the well-trained deep learning model.

B. Occlusion-perturbed Forecast Scenario Generation

Occlusion analysis is widely used in the computer vision field to understand the importance and contribution of image pixels. In this research, we borrow this technique to generate GHI forecast scenarios, which provides direct inputs to stochastic modeling in power systems. As shown in Fig. 2, the occlusion perturbation is achieved by sliding a black window across the input sky image stack. The newly constructed sky image stacks, $\mathbf{X}'$, are input to the trained PSolarNet to generate one forecast scenario, $\hat{\mathbf{y}}'$: $\hat{\mathbf{y}}'_k = F^*(\mathbf{X}'_k, \mathbf{W}^*)$, where k is the scenario index, whose maximum, K, is determined by two parameters: the window size s and the sliding stride l. Both parameters are set to be 16 in this study to balance the number of scenarios and the computational efficiency.

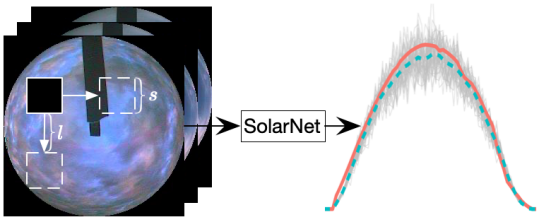


Figure 2. An example of occlusion-perturbed forecast scenario generation. The red and green lines are actual GHI series and forecast GHI series. The grey lines are forecast scenarios with the newly constructed perturbed sky image sequences.

C. BMA Probabilistic Forecasting

BMA is a kernel-dressing method, which was first applied in probabilistic forecasting for precipitation, wind speed, and

was recently introduced in day-ahead probabilistic solar forecasting [5]. In BMA for scenario ensemble forecasting, each forecast scenario $\hat{\mathbf{y}}'_k$ is associated with a conditional PDF $p_k(\mathbf{y}|\hat{\mathbf{y}}'_k)$, which assumes the PDF of the GHI $\mathbf{y}$ conditional on $\hat{\mathbf{y}}'_k$ being the best forecast in all the scenarios. Then, the BMA predictive PDF is [9]:

$$p(\mathbf{y}|\hat{\mathbf{y}}'_1, \ldots, \hat{\mathbf{y}}'_K) = \sum_{k=1}^{K} w_k p_k(\mathbf{y}|\hat{\mathbf{y}}'_k) \quad (1)$$

where w_k is the posterior probability of forecast scenario k being the best one. $p_k(\cdot)$ is assumed to be a truncated normal distribution between 0 and maximal GHI (GHI_{max}) [5]:

$$p_k(\mathbf{y}, \mu_k, \sigma_k) = \frac{\phi(\frac{\mathbf{y}-\mu_k}{\sigma_k})}{\sigma_k(\Phi(\frac{GHI_{max}-\mu_k}{\sigma_k}) - \Phi(\frac{0-\mu_k}{\sigma_k}))} \quad (2)$$

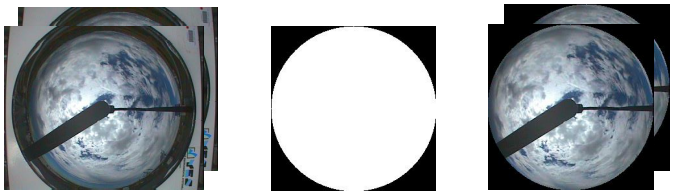
where $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and cumulative distribution function (CDF) of the standard normal distribution with mean μ_k and standard deviation σ_k. With respect to our estimation procedure, μ_k is estimated using a simple linear regression of y on each PDF of forecast scenario y_k for bias correction [5]. Then, w_k and σ_k are estimated using maximum likelihood based on the training data. With the PDF and CDF calculated by the BMA, other forms of probabilistic forecasts, such as quantiles, percentiles, and intervals, could also be readily generated.

III. Experimental Setup

A. Dataset Description

Following the suggestion of using open-source datasets to practice solar forecasting, this research conducts all the experiments on a publicly available dataset–the National Renewable Energy Laboratory (NREL) solar radiation research laboratory (SRRL) dataset. The SRRL dataset is one of the largest publicly available datasets with both sky images and meteorological measurements. The data has been collected since 1981 at the South Table Mountain Campus of NREL (longitude: 105.18° W, latitude 39.74° N, elevation 1,828.2 m). More information about the dataset can be found in Stoffel and Andreas [10]. The data can be accessed through the `OpenSolar` package [11].

Two types of data, namely, sky images and numerical meteorological measurements, are downloaded from the NREL SRRL database. The sky images taken by a Yankee Total Sky Imager every 10-min are selected for the case studies. Six years of sky images are downloaded and pre-processed. Every image contains 352×288 pixels. The developed PSolarNet requires minimal preprocessing due to its end-to-end learning capability. The only preprocessing is using a binary mask to circularly crop the images to avoid the presence of hazy sky and obstacles. Every image with a region-of-interest has 256×256 pixels. Then, the resolution is decreased to 128×128 to reduce the PSolarNet training complexity. Based on our study, stacking 2 images yields the best results, which is adopted in this research [12]. Figure 3 demonstrates the only image pre-processing applied in this research. Numerical measurements are averaged from 1-minute to 10-minute to keep consistency with the sky images. To ensure a successful training process and avoid over-fitting, six years (i.e., 155,644 data points in the daytime) of data are used for case studies, where the first three years of data (i.e., from 2012-01-01 to 2014-12-31) are used for training, the following one year of data (i.e., from 2015-01-01 to 2015-12-31) are used for validation, and the last two years of data (i.e., from 2016-01-01 to 2017-12-31) are used for testing.



(a) An original sky image stack

(b) A binary region-of-interest mask

(c) A processed sky image stack

Figure 3. Sky image pre-processing.

B. Benchmarks

Two sets of benchmarks are used to compare with the PSolarNet deterministic forecasts and probabilistic forecasts, respectively. The deterministic benchmarks include 6 machine learning methods and the persistence of cloudiness (PoC) method. The machine learning methods are selected from a larger volume of methods based on their performance in the empirical study. The selected machine learning benchmarks include two ANNs, three gradient boosting machine methods (GBMs), and a random forest (RF) method. These machine learning represent the state-of-the-art techniques in solar forecasting, which were extensively used in recent research [8], [13], [14]. The hyperparameters are optimized through cross-validation. The inputs to the machine learning models include GHI, direct normal irradiance, diffuse horizontal irradiance, clear sky GHI, clear sky direct normal irradiance, clear sky diffuse horizontal irradiance, dry bulb temperature, wind chill temperature, relative humidity, wind speed, peak wind speed, and atmospheric pressure, all at the current time point and at the previous 10 minute reading. The same data partition is applied to the benchmark modeling.

The second set of benchmarks include four widely-used very short-term probabilistic solar forecasting methods, which are the climatology method (CLI), the complete-history persistence ensemble (Ch-PeEn), the persistence ensemble (PeEn), and the Gaussian error distribution (GAU). The detailed methodology description, model setups, and code can be found in Ref. [7].

IV. Results and Discussion

Case studies are conducted on high performance computing GPU nodes and CPU nodes in the Eagle system at NREL. Each GPU node contains dual Intel Xeon Gold Skylake 6154

CPUs and Dual NVIDIA Tesla V100 PCIe 16 GB GPUs. Each CPU node has the same configuration as the GPU nodes, except for the GPUs. The PSolarNet and benchmark models are implemented using the `Keras` library with `Tensorflow` backend in `Python 3.7.0`, the `caret` package, and the `solarbenchmarks` in `R`. It took around 18.42 millisecond to generate forecasts for one timestamp, which indicates the high applicability even considering the communication time in realword.

A. Deterministic Forecasts

Three evaluation metrics, the normalized root mean square error (nRMSE), the RMSE-based forecast skill score (FSS) compared to the PoC, and the normalized mean bias error (nMBE), are used to evaluate the overall performance of deterministic forecasts. The definitions of the metrics can be found in Refs. [8], [13]. Table I lists forecast errors and skill scores of the SolarNet (i.e., the deterministic part of PSolarNet) and the seven benchmarks. Smaller nRMSEs and larger FSSs indicate better deterministic forecasts. Therefore, the PSolarNet with two sky images outperforms the PoC baseline and machine learning models using numerical measurements.

Table I. Deterministic forecast nRMSE [%], nMBE [%], and FSS [%]

Metric	SolarNet	PoC	ANN1	ANN2	GBM1	GBM2	GBM3	RF
nRMSE	5.62	7.17	6.74	7.01	6.94	6.93	7.07	6.61
nMBE	0.10	0.33	0.53	-0.49	-0.07	-0.09	-0.19	0.00
FSS	21.65	0.00	6.03	2.32	3.28	3.43	1.40	7.84

Figure 4 shows joint distributions and scatter plots by hexagon bins of the actual and forecast GHI of the SolarNet and deterministic benchmarks. The plots evaluate the quality of the forecasts in a time-independent manner by the dispersion of scatter points along the diagonal. The SolarNet is concentrated along the diagonal in a denser fashion, which means that the model has fewer large biases than the benchmarks. Compared to GBM models that have truncated distributions, the SolarNet covers the full GHI range equally. Compared to ANN models, the SolarNet distributions have more balanced probabilities on the upper and lower sides of the diagonal. The findings are consistent with what is revealed by the overall metrics in Table I. Hence, it is concluded that the SolarNet provides more accurate point forecasts.

B. Forecast Scenarios

A total of 64 scenarios are generated by the occlusion perturbation using a 16×16 sliding window. After removing the scenarios identical to the deterministic forecast time series (due to the overlap of the occlusion window and the mask black margin), 58 forecast scenarios are left. Figure 5a shows the scatter plot of error metrics of the deterministic forecasts and the forecast scenarios, which indicates accurate but distinct forecast scenarios. Statistical properties of the generated scenarios are evaluated by the autocorrelation, CDF, and correlation, which are shown in Figs. 5b–5d. Autocorrelation represents the temporal correlation within a GHI time series.

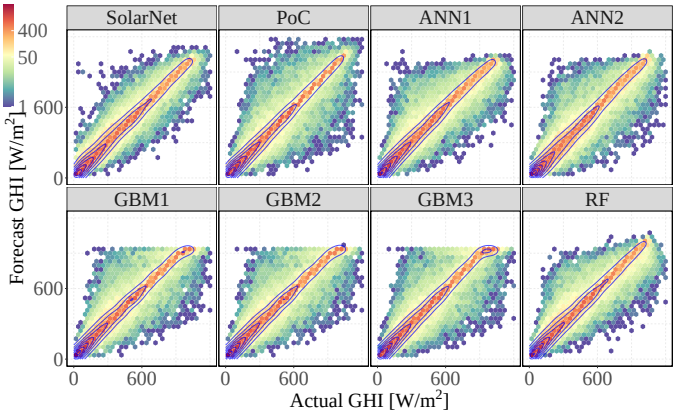
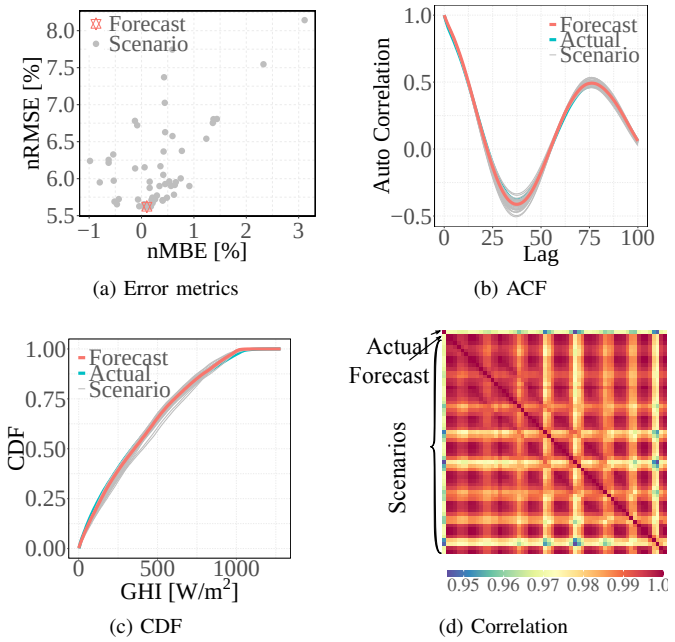


Figure 4. Joint distributions of the actual and forecast GHI.

It is found from Fig. 5b that the generated GHI forecast scenarios maintain the temporal correlation in the actual GHI series. The CDFs in Fig. 5c show that the forecast scenarios are statistically similar to the actual GHI series and the deterministic forecast series, which can be shown from another perspective by the correlations in Fig. 5d. The generated scenarios have a 0.966 average correlation with the actual GHI series. Therefore, we can conclude that the PSolarNet generates diverse but realistic forecast scenarios.



(a) Error metrics (b) ACF

(c) CDF (d) Correlation

Figure 5. Characteristic comparisons among the 58 GHI forecast scenarios, deterministic forecast time series, and actual time series.

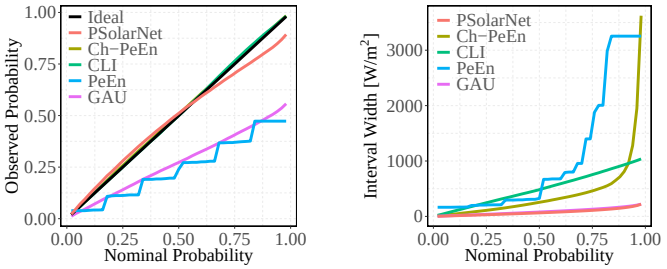
C. Probabilistic Forecasts

The overall probabilistic forecast accuracy is evaluated by two metrics, the normalized continuous ranked probability

score (nCRPS) and the normalized pinball loss (nPL) [7]. In both metrics, smaller values indicate better probabilistic forecasts. Table II lists the overall performance evaluation metrics of PSolarNet and its four benchmarks, which shows that the PSolarNet outperforms the other methods. Reliability and sharpness are the two main characteristics assessing the quality of a probabilistic forecast. Figure 6 compares the reliability and sharpness of the probabilistic forecasting models. A reliability curve that is closer to the diagonal means a more reliable forecasted probability, while a sharpness curve with smaller values indicates a lower level of uncertainty in the forecast. It is not surprising to observe that some forecasting models sacrifice sharpness to improve reliability, such as the Ch-PeEn model and the CLI model. By considering both the reliability and sharpness of probabilistic forecasts, the PSolarNet is the best model.

Table II. Probabilistic forecast nCRPS [%] and nPL [%]

	PSolarNet	Ch-PeEn	CLI	PeEn	GAU
nCRPS	2.77	32.47	14.04	38.72	5.37
nPL	1.39	4.02	6.92	11.49	2.59



(a) Reliability diagram

(b) Sharpness diagram (interval width)

Figure 6. Probabilistic quantile reliability and sharpness.

Figure 7 shows a one-day example of the actual GHI time series, GHI forecast time series, the 58 generated forecast scenarios, quantiles, and the associated image sequences. This specific day (2017-12-23) contains sunny conditions and partially-cloudy conditions. It is observed that the PSolarNet handles large uncertainty by generating various scenarios during solar ramps introduced by cloud development, therefore generating quantiles with wider intervals.

V. Conclusion

This paper developed a deep probabilistic convolutional neural network, named the PSolarNet, for probabilistic solar forecasting with sky image sequences. The PSolarNet quantified the solar forecasting uncertainty by generating forecast scenarios using a 21-layer deep learning model from occlusion-perturbed sky images. The Bayesian model averaging was used to post-process forecast scenarios and generate probabilistic solar forecasts. Case studies based on 6 years of data showed that the PSolarNet is able to provide: (i) accurate 10-minute ahead deterministic solar forecasts with a 5.62% normalized root mean square error, (ii) realistic and diverse solar forecast scenarios with a 0.966 average correlation with the actual time series, and (iii) reliable and sharp probabilistic forecasts with a 2.77% normalized continuous ranked probability score. In the future, we will extend this research by optimizing the sliding window parameters, further improving the probabilistic forecasts, and verifying multi-step forecasting performance.

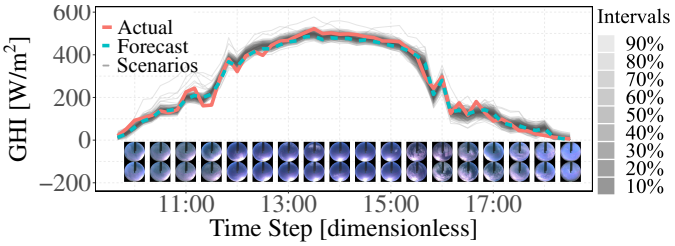


Figure 7. GHI time series of the actuals, forecasts, scenraios, and quantiles (2017-12-23) with the associated image sequence input.

References

[1] B. Li and J. Zhang, "A review on the integration of probabilistic solar forecasting in power systems," *Solar Energy*, vol. 210, pp. 68–86, 2020.

[2] M. David, F. Ramahatana, P.-J. Trombe, and P. Lauret, "Probabilistic forecasting of the solar irradiance with recursive arma and garch models," *Solar Energy*, vol. 133, pp. 55–72, 2016.

[3] D. W. Van der Meer, J. Widén, and J. Munkhammar, "Review on probabilistic forecasting of photovoltaic power production and electricity consumption," *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 1484–1512, 2018.

[4] R. Buizza, "Introduction to the special issue on "25 years of ensemble forecasting"," *Quarterly Journal of the Royal Meteorological Society*, vol. 145, pp. 1–11, 2019.

[5] K. Doubleday, S. Jascourt, W. Kleiber, and B.-M. Hodge, "Probabilistic solar power forecasting using bayesian model averaging," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, pp. 325–337, 2020.

[6] M. Sun, C. Feng, and J. Zhang, "Probabilistic solar power forecasting based on weather scenario generation," *Applied Energy*, vol. 266, p. 114823, 2020.

[7] K. Doubleday, V. V. S. Hernandez, and B.-M. Hodge, "Benchmark probabilistic solar forecasts: Characteristics and recommendations," *Solar Energy*, vol. 206, pp. 52–67, 2020.

[8] C. Feng and J. Zhang, "Solarnet: A sky image-based deep convolutional neural network for intra-hour solar forecasting," *Solar Energy*, vol. 204, pp. 71–78, 2020.

[9] J. M. Sloughter, T. Gneiting, and A. E. Raftery, "Probabilistic wind speed forecasting using ensembles and bayesian model averaging," *Journal of the american statistical association*, vol. 105, no. 489, pp. 25–35, 2010.

[10] T. Stoffel and A. Andreas, "Nrel solar radiation research laboratory (srrl): Baseline measurement system (bms); golden, colorado (data)," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 1981.

[11] C. Feng, D. Yang, B.-M. Hodge, and J. Zhang, "Opensolar: Promoting the openness and accessibility of diverse public solar datasets," *Solar Energy*, 2019.

[12] C. Feng, J. Zhang, W. Zhang, and B.-M. Hodge, "Convolutional neural networks for intra-hour solar forecasting based on sky image sequences," *Applied Energy*, 2022.

[13] C. Feng, M. Cui, B.-M. Hodge, S. Lu, H. Hamann, and J. Zhang, "Unsupervised clustering-based short-term solar forecasting," *IEEE Transactions on Sustainable Energy*, 2018.

[14] I. Jebli, F.-Z. Belouadha, M. I. Kabbaj, and A. Tilioua, "Prediction of solar energy guided by pearson correlation using machine learning," *Energy*, vol. 224, p. 120109, 2021.