SolarNet: A sky image-based deep convolutional neural network for intra-hour solar forecasting

Cong Feng, Jie Zhang*

Department of Mechanical Engineering, The University of Texas at Dallas, Richardson, TX 75080, USA

ABSTRACT

The exponential growth of solar energy poses challenges to power systems, mostly due to its uncertain and variable characteristics. Hence, solar forecasting, such as very short-term solar forecasting (VSTSF), has been widely adopted to assist power system operations. The VSTSF takes inputs from various sources, among which sky image-based VSTSF is not yet well-studied compared to its counterparts. In this paper, a deep convolutional neural network (CNN) model, called the SolarNet, is developed to forecast the operational 1-h-ahead global horizontal irradiance (GHI) by only using sky images without numerical measurements and extra feature engineering. The SolarNet has a set of models that generate fixed-step GHI in parallel. Each model is composed of 20 convolutional, max-pooling, and fully-connected layers, which learns latent patterns between sky images and GHI in an end-to-end manner. Numerical results based on six years data show that the developed SolarNet outperforms the benchmarking persistence of cloudiness model and machine learning models with an 8.85% normalized root mean square error and a 25.14% forecasting skill score. The SolarNet shows superiority under various weather conditions.

1. Introduction

Solar energy, especially solar photovoltaics (PV), has been witnessed an aggressive increase in the last decade. A total of 94.2 GW PV was installed in 2018 worldwide and this exponential growth is foreseen and projected in the future decades. Nevertheless, integrating large amounts of PV into power systems is challenging, mostly due to the uncertain characteristics of PV generation. To deal with this issue, the importance of solar forecasting has been emphasized in power system operations. On one hand, balancing authorities rely on solar power forecasts in their market operations and management. For instance, the Electric Reliability Council of Texas requires solar forecasts with various forecasting time resolutions and horizons for its different markets and procedures, including day-ahead reliability unict commitment, hour-ahead (HA) reliability unit commitment, transmission security analysis, and real-time security-constrained economic dispatch (ERCOT, 2018). On the other hand, PV plant output forecasts are also used to help plant operators bid in energy markets. For example, accurate solar forecasts are able to reduce the economic penalization for output imbalances (Law et al., 2016; Kaur et al., 2016). Therefore, improving solar forecasting accuracy is crucial and has been paid efforts in the literature.

A large collection of methods have been developed for short-term and very short-term solar forecasting (VSTSF), where machine learning (ML) methods are one of the most popular categories. For example, eleven statistical and ML models were analyzed under various meteorological conditions, among which regression tree-based models outperformed others in most cases (Fouilloy et al., 2018). Moreover, short-term PV output power forecasts of three grid-connected PV systems were generated by the extreme learning machine model, which showed better accuracy than support vector regression models and artificial neural networks (ANNs) (Hossain et al., 2017). Lauret et al. (2015) developed a set of benchmarking ML models, including Gaussian processes and support vector regression (SVR) models, for short-term irradiance forecasting. The ensemble learning with ANNs, SVR models, gradient boosting machines (GBMs), and a random forest (RF) model was used to provide very short-term solar forecasts (VSTSFs) with similar hour or similar weather downsampling strategies (Feng et al., 2017; Feng and Zhang, 2018). A detailed ML review can be found in recent review papers (Voyant et al., 2017; Yang et al., 2018).

Compared to ML algorithms, deep learning (DL) approaches have shown potentials to further enhance the VSTSF accuracy. However, investigation of the deep learning-based VSTSF is far from comprehensive. The most popular DL network in solar forecasting is the long short-term memory network (LSTM) due to its powerful capability of learning time series patterns. For example, an LSTM that considered mutual dependencies among weather features was used for VSTSF and

* Corresponding author.
E-mail address: jiezhang@utdallas.edu (J. Zhang).

improved the accuracy by 18.34% over ANN (Qing and Niu, 2018). Another LSTM was combined with the least absolute shrinkage and selection operator to learn nonlinear and linear relationships in the VSTSF, which outperformed single models (Wang et al., 2018). Compared to LSTM, other types of DL methods, e.g., convolutional neural networks (CNNs), are even less researched. This is due to the low feature dimension in VSTSF, which cannot take full advantage of the CNN as local feature extractors (this category of research will be reviewed in the next paragraph).

Heavily relying on the data, ML and DL methods predict VSTSF from various input sources. The most popular inputs are ground numerical measurements, satellite-based information, and sky image features. Among different measuring instruments, total sky imagers (TSIs, also known as sky cameras) show uniqueness in VSTSF over others. For example, sky images provide sky and cloud conditions, from which unique latent features can be extracted. More importantly, the cost demand and expertise requirement for on-site meteorological measurements limit the availability of weather data for distributed PV forecasting. Actually, not all the rooftop PV systems have the meteorological sensors. Therefore, cost-effective TSIs provide alternative opportunities to accurate forecast very short-term PV power for both utility-scale and distributed PV systems (Richardson et al., 2017).

However, the existing TSI-based VSTSF has nonnegligible drawbacks. For example, some VSTSF used human crafted features from TSIs, such as pixel-based features (Feng et al., 2018; Chu et al., 2013; Marquez and Coimbra, 2013), whose optimality and impact are suspicious. DL encoding techniques were introduced to address this problem, but most of them extracted features based on classification models. For example, Zhao et al. (2019) developed a CNN feature extractor to provide features generated by a clear sky index classifier to an ANN model and a linear autoregressive model. Similarly, the forecasting process is conducted after the DL classification by CNNs (Pothineni et al., 2018). Contrarily, limited work has examined to predict VSTSFs directly from TSIs using DL methods.

In this paper, a DL-based network, named the SolarNet, is developed to bridge the mentioned gaps by automatically learning latent sky image features for VSTSF. In such a way, we seek to identify the effectiveness of using a single TSI in intra-hour solar forecasting without the help of other measurements. The reasons we only use one single sky image as the input without endogenous parameter, other exogenous data, and image streams are that (i) the focus of this paper is to verify the effectiveness of automatic feature learning, which replaces the human-defined image features in VSTSF, (ii) the one sky image-based VSTSF could be an alternative approach that benefits distributed PV systems, (iii) modeling interactions among one sky image with endogenous/exogenous variables and sky image streams is significantly different from this developed architecture, which is out of the scope of this paper, and (iv) the consideration of meteorological parameters and image streams is included in the discuss section and will be explored in separate papers.

The contributions of this paper include: (i) extracting TSI-based features for VSTSF through an end-to-end representation learning, (ii) utilizing the state-of-the-art DL configuration, (iii) providing operational 1HA solar forecasting for the real-time market, and (iv) achieving significant improvement over benchmark models. The remainder of this paper is organized as follows. Section 2 formulates the TSI-based operational 1HA solar forecasting problem and introduces the developed SolarNet. The case study dataset, DL training details, and benchmark models are described in Section 3. Section 4 discusses the results. Conclusions are made in Section 5.

2. The SolarNet

In this section, we first formulate the TSI-based operational 1HA solar forecasting problem. Then, different types of layers (i.e., basic blocks of the SolarNet) in the SolarNet framework are mathematically described. At last, the overall framework of a SolarNet model is introduced.
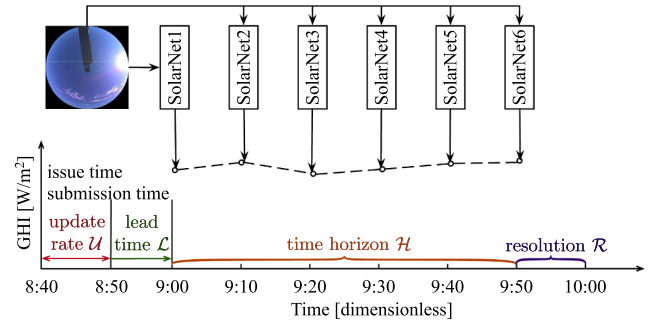

Fig. 1. The operational intra-hour solar forecasting with the SolarNet.

2.1. Problem formulation

In this paper, we seek to forecast the operational 1HA global horizontal irradiance (GHI). The importance of providing operational forecasting based on power system operational time scales has been recognized and recommended by the solar forecasting society (Yang, 2019; Yang et al., 2019). Different from fixed-step forecasting, the operational forecasting is defined by four time-related parameters, as shown in Fig. 1 (Yang et al., 2019). The first term is *lead time*, $\mathcal{L}$, which indicates the difference between the forecasting submission time and the next operating time. The second term is *time horizon*, $\mathcal{H}$, which is the span of forecasts generated at each issue time. The third term is *resolution*, $\mathcal{R}$, which is the interval of forecasts. The last one is *update rate*, $\mathcal{R}$, which is the interval between two issue time. Having different operation procedures, independent system operators (ISOs)/regional transmission operators have different VSTSF requirements. For example, California ISO (CAISO) runs its real-time economic dispatch every 5-min, which needs 65-min-ahead (65MA) solar forecasts with a 5-min resolution (CAISO, 2019). The forecasts should be updated every 15-min and submitted 10-min before the next operating time.[1]

The SolarNet framework provides the intra-hour forecasts with multiple models in parallel, which is flexible based on the dataset time resolution, as shown in Fig. 1. Each SolarNet model forecasts the GHI, $\mathbf{y} \in \mathbb{R}^{N \times 1}$, with a fixed time step. The SolarNet only takes the latest TSI, $\mathbf{X} \in \mathbb{R}^{N \times W \times H \times D}$, as input without numerical measurements and feature engineering:

$$\hat{\mathbf{y}}_{\Delta t} = F(\mathbf{X}, \mathbf{W}) \tag{1}$$

where N, W, H, D are sample size, image width, image height, and image channel amount, respectively; $\mathbf{y}$ and $\hat{\mathbf{y}}$ are the actual and forecast GHI, respectively; $\mathbf{X}$ is the image input; Δt is the forecasting step in terms of minute; $\mathbf{W}$ is the trainable parameters in the developed CNN model. To perform VSTSF, optimal parameters should be obtained first in the training stage. Since VSTSF is a regression problem, we use mean absolute error (MAE), which is a popular used forecasting error metric, as the loss function:

$$J(\mathbf{W}) = \frac{1}{N} \sum_{n=1}^{N} |\hat{y}_{\Delta t,n} - y_{\Delta t,n}| \tag{2}$$

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, $\mathbf{W}$, by minimizing the loss function in an end-to-end manner (i.e., from images to GHI values).

Based on the dataset we used for this work (the dataset will be described in Section 3.1), the four time-related parameters for the intra-hour operational solar forecasting are: $\mathcal{L} = 0$ (assume the real real-time), $\mathcal{H} = 60$-min, $\mathcal{R} = 10$-min (the resolution of sky image data is 10-

---

[1] In the literature, the CAISO lead time requirement is 7.5mins, however, 10mins is required to consider communication lags during real-time operations in CAISO.

min), and $\mathcal{U} = 10$-min. Even though the settings are still different from the real operational forecasting, similar assumptions were made in wind forecasting and the results are able to reflect the operational forecasting accuracy (Xie et al., 2013). To extend the point forecasting formulated in Eq. 1 to sequential forecasting, the parallel approach is selected due to its popularity (Feng et al., 2019a). Thus, 6 SolarNet models, namely, SolarNet1–SolarNet6, are built to provide 10MA–60MA GHI forecasts, as shown in Fig. 1.

2.2. Deep learning layers

To build the SolarNet, different types of layers are required. The layers in the SolarNet include convolutional layers, pooling layers, and fully-connected layers (also known as dense layers).

2.2.1. Convolutional layers

CNN has achieved successes and brought breakthroughs in processing images, videos, and speech, due to its powerful feature learning ability and efficient weight sharing strategy (LeCun et al., 2015). Convolutional layers are the key components in CNN models. Generally, convolution operations are performed to the input tensor by filters in the filter bank (also known as kernels), which are expressed as $\mathbf{Z}^l = \mathbf{W}^l * \mathbf{X}^l + \mathbf{b}^l$, and detailed as:

$$z'^{l}_{w^{(l+1)},h^{(l+1)},d^{(l+1)}} = \omega^{l}_{w^l,h^l,d^l,d^{(l+1)}} \times x^{l}_{w^l,h^l,d^l} + b^{l}_{w^{(l+1)},h^{(l+1)},d^{(l+1)}} \quad (3)$$

where $\mathbf{X}$, $\mathbf{W}$, $\mathbf{b}$, and $\mathbf{Z}$ are the input tensor, parameter tensor, bias tensor, and output tensor in the lth layer, respectively; $*$ is a convolution operator; z, ω, x, and b are elements in the output, parameter, input, and bias tensors, respectively; w, h, and d are element location indices. Please note that the channel number, D^{l+1}, of the next layer equals to the filter number of the previous layer. After the convolution operations, the feature maps are activated by an element-wise rectified linear unit (ReLU) to add nonlinearity, as:

$$z^{l}_{w^{(l+1)},h^{(l+1)},d^{(l+1)}} = \max(0, z'^{l}_{w^{(l+1)},h^{(l+1)},d^{(l+1)}}) \quad (4)$$

2.2.2. Pooling layers

To achieve more translation invariance during the spatial representation learning, pooling layers are used to perform max-pooling over a 2×2 pixel window with stride 2. In such a way, the convoluted feature maps are sub-sampled by the non-overlapping pooling windows by a factor of 2, which can be expressed as:

$$z^{l}_{w^{(l+1)},h^{(l+1)},d^l} = \max\{\mathbf{x}^l | \mathbf{x}^l := x^{l}_{w^l:w^l+\frac{W^l}{W^{l+1}}, h^l:h^l+\frac{H^l}{H^{l+1}}, d^l}\} \quad (5)$$

where $\mathbf{x}$ is a subset of a feature map.

2.2.3. Fully-connected layers

The convolutional and max-pooling layers are powerful in extracting informative features from the TSIs, but cannot directly solve the regression problem. Therefore, the fully-connected layers are introduced in the model to learn the latent patterns between the TSI feature maps and the GHI values. All the inputs are transmitted to the output in fully-connected layers, as:

$$\mathbf{Z}^l = \mathbf{W}^l \cdot \mathbf{X}^l + \mathbf{b}^l \quad (6)$$

2.3. The overall SolarNet model

As shown in Fig. 2, the developed SolarNet is composed of 20 layers, including 13 convolutional layers, 5 max-pooling layers, and 2 fully-connected layers. The configuration of the SolarNet is inspired by the well-known very deep convolutional network developed by the Visual Geometry Group (VGG16) (Simonyan and Zisserman, 2014). The
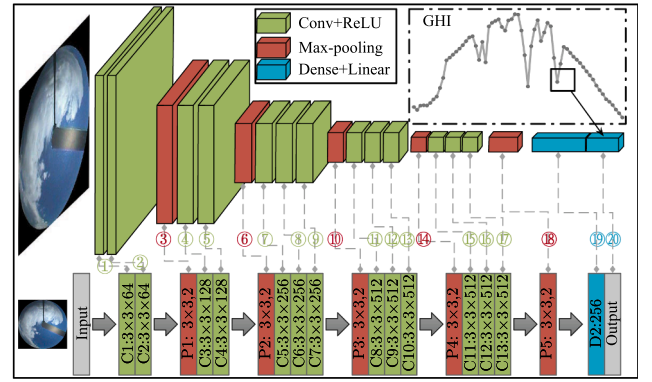

Fig. 2. The overall structure of the SolarNet. Colors indicate different layer types. The upper part shows the SolarNet layers, while the lower part shows the detailed hyperparameters of each layer.
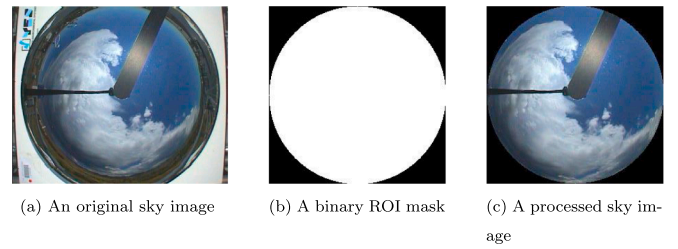

(a) An original sky image

(b) A binary ROI mask

(c) A processed sky image

Fig. 3. Sky image pre-processing.

SolarNet contains 5 feature learning blocks (FLBs), each of which stacks 2 or 3 convolutional layers and a max-pooling layer, which is shown in the upper part of Fig. 2. The SolarNet utilizes small filters with a size of 3×3 to extract more receptive fields from the input tensors. The filter number of the next FLB is doubled to extract more abstract and informative features. The reason of selecting the VGG16 as the sky image feature extractor is that: (i) the network is less complicated than other popular network, such as the residul nets (i.e., ResNet50),[2] which reduces the risk of overfitting, and (ii) the accuracy of the VGG16-based SolarNet is more accurate in the empirical tests (by a smaller dataset) than other popular networks. However, the VGG16 is designed for classification problems. To solve a deep regression problem, the last three dense layers of the VGG16 with ReLU functions was replaced with two dense layers with linear activation functions. Also, the neurons of the two dense layers are 256 and 1. Please note the dense layer topology and hyperparameter selection are determined by trial-and-error. In addition, 20% of neurons are randomly dropped to avoid over-fitting in the dense layers.

The SolarNet forecasts GHI directly from TSIs without any feature engineering and any numerical inputs. The reason of using only one sky image as input is that: (i) the focus of this paper is to verify the effectiveness of sky images for VSTSF without meteorological parameters, (ii) the time dimension of the image is set as 1 to simplify the model and reduce the computational requirement, (iii) VSTSF with one sky image has potential to improve the accuracy and hasn't been explored before.

3. Experimental setup

This section briefly introduces the dataset for case studies, the SolarNet training strategies, and the benchmark models.

3.1. Dataset description

In this paper, the largest publicly available dataset with both TSIs and

---

[2] The ResNet is eight times deeper than VGG nets (He et al., 2016).

Table 1
Machine learning benchmark models.

Model	Function/Algorithm	Hyperparameter
ANN1	Momentum BP	`lr = 0.01, max_epoch = 1,000, momentum = 0.9`
ANN2	Standard BP	`lr = 0.01, max_epoch = 1,000`
GBM1	Squared loss	`lr = 0.01, ntrees = 1,000, max_depth = 20, bag_frac = 0.5`
GBM2	Laplace loss	`lr = 0.01, ntrees = 1,000, max_depth = 20, bag_frac = 0.5`
GBM3	T-distribution loss	`lr = 0.01, ntrees = 1,000, max_depth = 20, bag_frac = 0.5, DF = 4`
RF	CART aggregation	`ntrees = 1,000, mtry = 5`
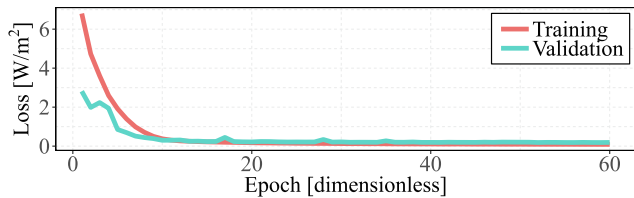
Fig. 4. The SolarNet1 training and validation losses over the course of training.

numerical measurements, namely, the national renewable energy laboratory (NREL) solar radiation research laboratory (SRRL) dataset, is used for case studies. The SRRL has been collecting continuous solar measurements at NREL's South Table Mountain Campus (longitude: 105.18° W, latitude 39.74° N, elevation 1,828.2 m), Golden, Colorado since 1981. In addition to numerical measurements with a 1-min resolution, the SRRL dataset has two sets of total TSIs taken by a Yankee Total Sky Imager (TSI-800) and an EKO All Sky Imager with a 10-min resolution. More information of the SRRL dataset can be found in a recent solar energy data artical (Feng et al., 2019b).

The numerical meteorological features, including GHI, direct normal irradiance, direct horizontal irradiance, temperature, relative humidity, wind speed, and atmosphere pressure, are used (only GHI is used in the SolarNet as labels, while others are used in benchmarks). The calendar features, i.e., month of the year and hour of the day, and the Ineichen and Perez clear sky GHI (denoted as CGHI, obtained from `pvlib` package in Python (Stein et al., 2016)) are also included. For the TSIs, the TSI-800 images are used in this study due to its longer coverage period. To match the resolution with TSIs, numerical data are averaged every 10 min. The original TSIs have 352×288 pixels, whose regions of interest (ROI) are extracted to avoid the hazy sky and obstacle presence. Therefore, a 256×256

binary mask is applied to circularly cropping the TSIs by centering at their center pixels, as shown in Fig. 3. Both the TSIs and numerical data are normalized by the maximum values before being fed into the model (GHI is normalized by the CGHI). To ensure a successful training process without over-fitting, six years (i.e., from 2012-01-01 to 2017-12-31) numerical and image data are downloaded and processed, where the first three years (i.e., from 2012-01-01 to 2014-12-31) data are used for training, the following one year data (i.e., from 2015-01-01 to 2015-12-31) are used for validation, and the last two years data are used for testing (i.e., from 2016-01-01 to 2017-12-31). The SRRL dataset is maintained with the SERI-QC quality control methodology, where the special "twilight" one component min-max empirical limits tests are run on data for which the corresponding solar zenith angle is greater than 80° but less than 90°. We do not perform additional quality control, except discarding meaningless negative and zero GHIs in the early morning and late night in the dataset. All the other data in large zenith angle periods are kept in the dataset. More details about the SERI-QC can be found in (Stoffel and Andreas, 1981; Anderberg and Sengupta, 2014). At last, the sample sizes in the training, validation, and testing sets are 76,708, 26,271, and 52,665, respectively.

3.2. Training strategies

Having 14,846,273 parameters distributed in 20 layers, the SolarNet is trained with mini-batch by an optimizer. Several optimizers were tested, including the stochastic gradient descent (SGD), adaptive moment optimization (Adam), and the AdaDelta, as suggested by Lathuilière et al. (2019). The Adam algorithm outperformed others based on the validation data, therefore, is selected as the optimizer in this paper. The Adam optimizer is a combination of the root mean square propagation and SGD optimizers. The
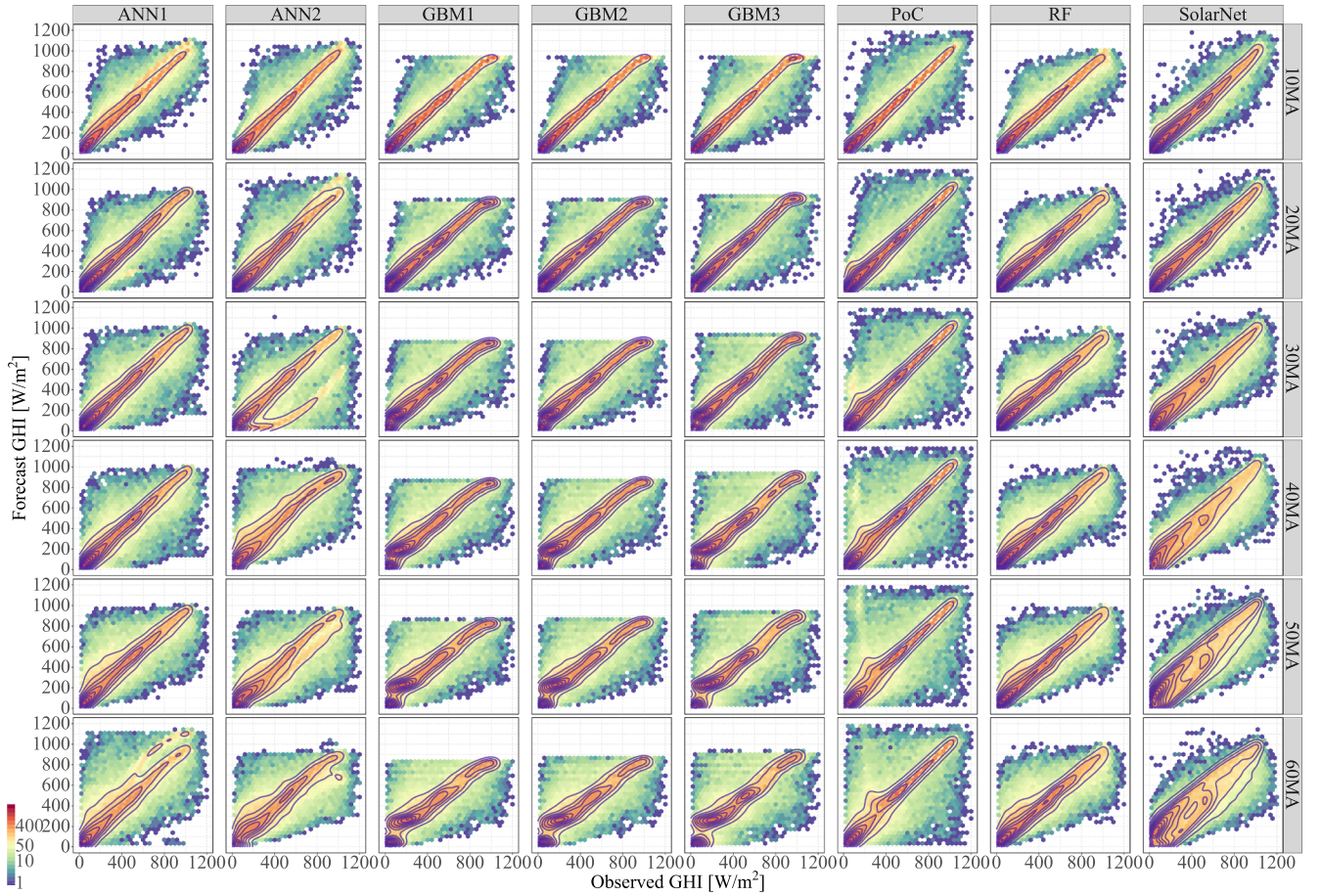
Table 2
GHI forecasting errors.

Metrics	$\mathcal{H}/\Delta t$	ANN1	ANN2	GBM1	GBM2	GBM3	RF	PoC	SolarNet
RMSE [W/m^2]	1HA	129.77	146.22	128.84	128.51	131.83	121.01	160.15	116.82
	10MA	94.76	89.15	87.47	87.39	89.12	85.61	90.94	81.03
	20MA	112.32	116.49	114.47	114.55	117.33	109.33	123.76	92.24
	30MA	125.06	210.83	128.79	128.71	131.94	122.55	146.55	108.98
	40MA	135.01	147.22	139.22	138.91	142.71	130.00	170.90	132.60
	50MA	142.78	154.75	147.80	147.14	151.22	136.59	198.67	134.29
	60MA	168.69	158.87	155.30	154.33	158.65	141.95	230.09	151.80
MBE [W/m^2]	1HA	7.94	−16.63	−2.81	−1.23	2.01	−1.72	23.07	−4.56
	10MA	23.72	−5.43	−0.43	−0.69	−2.00	−1.34	4.31	−4.87
	20MA	−7.62	9.69	−0.84	−0.76	0.06	−1.35	10.34	1.65
	30MA	−3.49	−80.08	−1.84	−0.44	2.42	−2.97	17.69	−0.44
	40MA	−19.12	28.27	−2.87	−0.49	4.00	−0.63	26.09	−4.97
	50MA	16.65	−32.96	−4.40	−1.71	4.15	−1.71	35.20	−3.84
	60MA	37.48	−19.25	−6.46	−3.26	3.40	−2.34	44.76	−14.86
FSS [%]	1HA	15.92	5.15	16.69	16.85	14.77	21.23	0	25.14
	10MA	−4.20	1.97	3.82	3.90	2.00	5.87	0	10.91
	20MA	9.24	5.87	7.51	7.44	5.20	11.66	0	25.47
	30MA	14.67	−43.86	12.12	12.17	9.97	16.38	0	25.64
	40MA	21.00	13.85	18.54	18.72	16.49	23.93	0	22.41
	50MA	28.13	22.11	25.61	25.94	23.89	31.25	0	32.41
	60MA	26.69	30.96	32.51	32.93	31.05	38.31	0	34.02

Note: The maximum GHI in the testing dataset is 1,268.393 W/m^2, while the maximum GHI in the entire dataset is 1,319.328 W/m^2. 1HA denotes the operational 1HA solar forecasts, others are results of fixed Δt forecasts. PoC is the reference model for FSS calculation, therefore its FSSs are zero.

Fig. 5. The joint distributions of forecast versus observed GHI. For a higher contrast, the color scheme is based on the logarithm of scater point frequency.

parameters are initialized with the ImageNet pre-trained weights to promote the convergence. During the training process, the Adam minimizes the objective function $J(\mathbf{W})$ in Eq. 2 by independently updating every parameter, ω_i, in the opposite direction of its gradient:

$$m_{i,j} = \beta_1 m_{i,j-1} + (1 - \beta_1)\nabla_{w_i}J(\mathbf{W};\mathbf{B}_j) \tag{7}$$

$$v_{i,j} = \beta_2 v_{i,j-1} + (1 - \beta_2)\nabla_{w_i}^2 J(\mathbf{W};\mathbf{B}_j) \tag{8}$$

$$\hat{m}_{i,j} = \frac{m_{i,j}}{1 - \beta_1^j} \tag{9}$$

$$\hat{v}_{i,j} = \frac{v_{i,j}}{1 - \beta_2^j} \tag{10}$$

$$\omega_{i,j} = \omega_{i,j-1} - \Delta\omega_{i,j} = \omega_{i,j-1} - \eta \frac{\hat{m}_{i,j}}{\sqrt{\hat{v}_{i,j}} + \epsilon} \tag{11}$$

where m and v are exponential moving averages of the gradient and the square gradient along ω_i in the jth mini-batch, $\mathbf{B}_j = [\mathbf{X}_j, \mathbf{y}_j]$, respectively ($m_{i,0} = v_{i,0} = 0$). $\beta_1 = 0.9$ and $\beta_2 = 0.999$ are hyperparameters of the Adam optimizer. $\eta = 0.01$ is the initial learning rate and $\epsilon = 1 \times 10^{-8}$ is an extremely small constant to prevent zero division. In the training process, the complete training dataset is passed forward and backward 60 epochs, during which mini-batches with a batch size of 64 (determined by trial-and-error) are randomly generated to shuffle the data order. In each iteration, gradients are averaged over the mini-batch.[3]

---

[3] Note that the iteration number is determined by the training/validation data size and the corresponding batch size.

3.3. Benchmark models

To compare with the developed SolarNet, the persistence of cloudiness model (Feng et al., 2017) (denoted as PoC and also known as the smart persistence or cloudiness-index persistence) and a collection of 6 ML models are built based on the numerical data (with the latest meteorological features). The ML models include two ANNs with different back-propagation (BP), three GBMs with different loss functions, and an RF model. Please note that several ANNs, GBMs, and RF models were tested but only the best ones are reported in the paper. The training data and testing data are meteorological measurements with the same training/testing partition. Input features to ML models include GHI, direct normal irradiance (DNI), direct horizontal irradiance (DHI), clear sky GHI, clear sky DNI, clear sky DHI, infrared radiation, try bulb temperature, wind chill temperature, relative humidity, wind speed, peak wind speed, pressure in the past four hours. The hyperparameters of the benchmark models are empirically determined and summarized in Table 1, including the learning rate (`lr`) and the maximum number of epochs (`max_epoch`) in ANN1 and ANN2; the minimum update value (`min_delta`) and the maximum update value (`max_delta`) in ANN1; the momentum (`momentum`) in ANN2; the number of boosting iterations (`ntrees`), maximum tree depth (`max_depth`), learning rate (`lr`), out-of-bag fraction (`bag_frac`) in GBM1–GBM3; the degree of freedom (`DF`) in GBM3; and the number of trees (`ntrees`) and the number of variables randomly sampled as candidates at each split (`mtry`) in RF.

4. Results and discussion

The SolarNet case studies are conducted on a high performance computing (HPC) node in the Texas Advanced Computing Center (TACC)
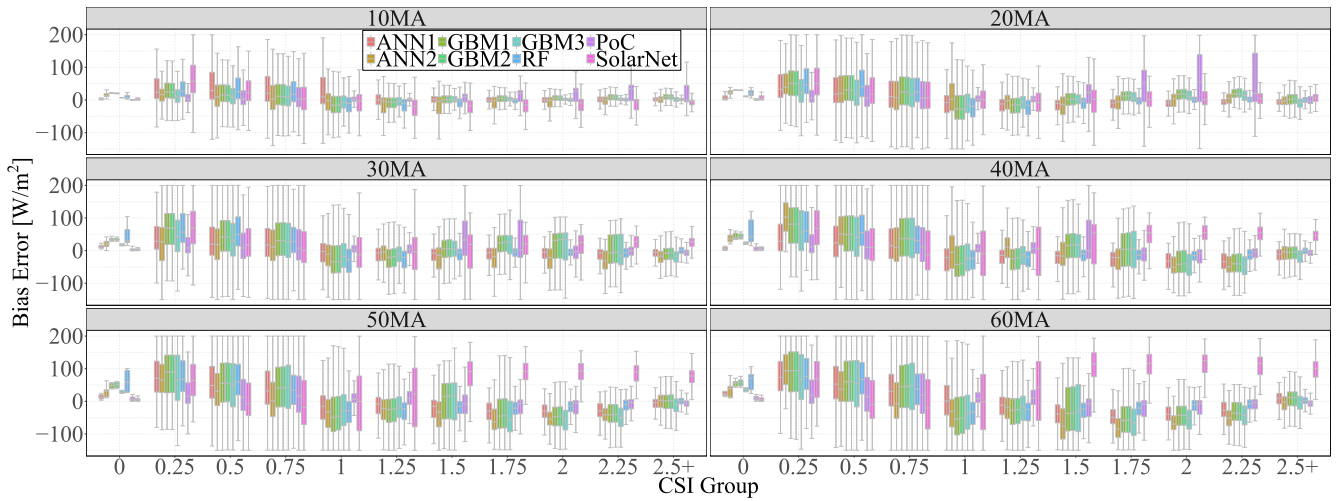
Fig. 6. Forecasting bias error statistics in CSI groups. The interquartile range box represents the middle 50% of bias errors. The upper and lower bounds are maximum and minimum bias error, respectively, excluding the outliers.

Maverick2 system at the University of Texas System. Each node contains 16 Intel(R) Xeon(R) CPUs and four Nvidia GTX 1080-TI GPUs. The benchmark case studies are conducted on a HPC node at the University of Texas at Dallas Ganymede system, which has 40 Intel(R) Xeon(R) CPUs. The SolarNet and benchmark models are implemented using the `Keras` library with `Tensorflow` backend and the `scikit-learn` library in Python version 3.6. It took averagely 429.96 min to train a SolarNet model and 6.31 min to generate forecasts of the entire testing set (around 7.19 ms for a single forecast). Considering the computational cost, the SolarNet is applicable for intra-hour solar forecasting. An example of training loss values are shown in Fig. 4, which reveals a successful training (other SolarNet models have similar training curves).

4.1. Overall accuracy

To quantify the overall performance of the SolarNet and its benchmarks, three error metrics are selected, which are the mean bias error (MBE), root mean square error (RMSE), and forecasting RMSE skill score (FSS):

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i) \tag{12a}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{12b}$$

$$\text{FSS} = (1 - \frac{\text{RMSE}_{\text{model}}}{\text{RMSE}_{\text{reference}}}) \tag{12c}$$

where n is the testing data length. The reference model is PoC, as suggested by multiple papers (Chu et al., 2013; Chu et al., 2015a; Chu et al., 2015b).

Table 2 lists the operational 1HA and fixed-step GHI forecasting errors of the developed SolarNet and 7 benchmarks. It is first observed that the SolarNet forecasting RMSE and MBE are 116.82 W/m^2 and -4.56 W/m^2 (nRMSE = 8.85% and nMBE = 0.35%), which outperforms other models. Specifically, SolarNet generates more accurate forecasts than the PoC in all the time horizons, which is revealed by the positive FSSs. In contrast, the selected well-performing ML benchmarks have a risk to provide worse forecasts compared to the PoC, such as ANN1 and ANN2. Additionally, the SolarNet operational 1HA solar forecasting FSS is 25.14%, which is the highest among all models and is a significant improvement. Moreover, compared to ML benchmarks, the SolarNet shows evident superior performance, especially for horizons shorter than 30-min. In addition, the forecasting RMSE increases with forecast horizon as expected. Considering the MBE, the SolarNet

underforecasts the GHI but is still competitive compared to benchmarks.

To compare the SolarNet with benchmarks in detail, joint distributions of the GHI observations and forecasts are shown as hexagon binning in Fig. 5. The following interesting observations can be made from the figure. The SolarNet improves forecasting accuracy by reducing the number of severely underpredictions (points in the bottom right corner) and overpredictions (points in the top left corner). In addition, the SolarNet has more underpredictions than overpredictions in all 6 time horizon scenarios, which is on the contrary to other models. Comparing to benchmark models, the joint distributions of SolarNet are more concentrated to the diagnal, especially for longer-term time horizons. Additionally, some unexpected patterns in the data are captured by benchmark models. For example, the distribution of the two ANN models may have a furcation pattern, such as 60MA forecasts of ANN1 and 30MA forecasts of ANN2. GBM models fail to forecast GHI larger than 1000 W/m^2, which is possibly due to reletively small samples in that range.

4.2. Weather effects

It is observed that the irradiance condition impacts SolarNet forecasts. To dig into the weather effects, forecasts are analyzed based on the clear sky index (CSI) groups. The CSI is an alternative parameter to indicate weather conditions, and was used to replace meteorologically-defined weather types (Gigoni et al., 2017; Lorenz et al., 2009; Feng et al., 2018).[4] Fig. 6 shows the statistics of forecasting bias error of the SolarNet and benchmark models under various weather conditions. Generally, the SolarNet outperforms benchmarks in partially cloudy and clear sky conditions in shorter-term horizons, since it has smaller MBE absolute values when CSI $\geqslant$ 0.75. Nonetheless, the performance of SolarNet is not as good as the best benchmark models in cloudy conditions and longer-term horizons, which is revealed by the large MBE and error variance (e.g., CSI = 0.25 and CSI = 0.5). This gives us directions to improve SolarNet in future work.

The findings are supported by the GHI time series, CSI time series, and their corresponding TSIs. Fig. 7 shows a typical day with clear sky (i.e., 5am–10am), partially cloudy (i.e., 2 pm and 5 pm–7 pm), and cloudy conditions (i.e., 11am–1 pm and 3 pm–4 pm). It is found that the SolarNet forecast errors are small under the clear sky and partially cloudy conditions (CSI $\geqslant$ 0.75). In contrast, forecast errors of the

---

[4] There are more than 33 meteorologically-defined weather types, which may not be proper, reliable, and available for weather effect analysis in solar forecasting (Wang et al., 2015).
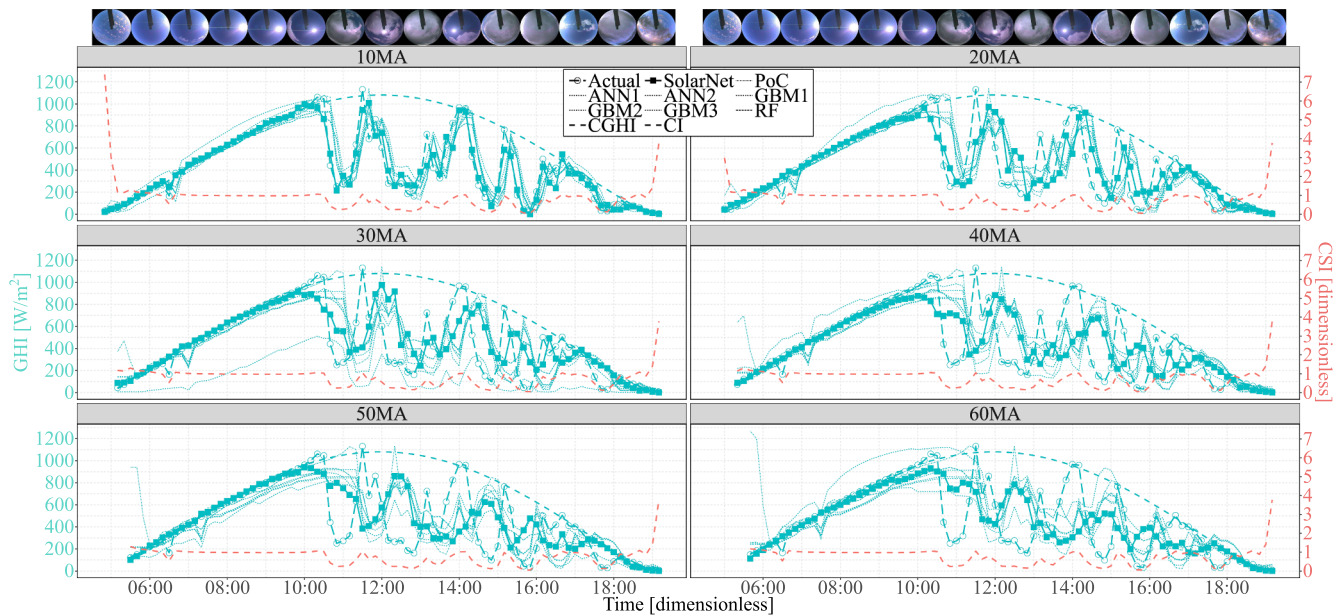
Fig. 7. Observed, forecast, CGHI time series, CSI time series, and TSIs of a typical day (i.e., 2017-05-29). Only hourly images are shown here, due to the space limitation.

SolarNet are large under cloudy conditions (CSI <0.75), which can be observed by the large deviations between the forecast and actual GHI.

4.3. Discussion

This is the first version of SolarNet, which aims to investigate the effectiveness of using only one image to forecast intra-hour GHI. Nonetheless, the network is flexible to (i) incorporate features from multiple sources, such as meteorological measurements, (ii) be extended to 3D CNN to take TSI time series as inputs, and (iii) provide probabilistic forecasts through quantile estimation by multi-output. Additionally, the impact of weather effects on the SolarNet forecasts also inspires us to include post-processing (e.g., solar zenith angle filtering) and weather recognition techniques to mitigate the negative effects under cloudy conditions in the future network.

5. Conclusion

This paper developed a 20-layer deep convolutional neural network (CNN), called the SolarNet, to forecast intra-hour (i.e., 10-min-ahead to 60-min-ahead with a 10-min interval) global horizontal irradiance (GHI). The SolarNet used only one total sky image (TSI) as the input without any numerical measurements and extra feature engineering process. The SolarNet was trained based on an end-to-end manner, during which the latent relations between TSIs and GHI were captured by the five feature learning blocks and two fully-connected layers. Numerical experiments based on six years publicly available data showed that the SolarNet provide multi-step forecasts with an 8.85% normalized root mean square error and a 25.14% forecasting skill score, which is a significant improvement compared with the persistence of cloudiness model and machine learning models. It was also found that the SolarNet can be further improved by considering the weather effects. Future research directions were also discussed in the paper.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Anderberg, M., Sengupta, M., 2014. Comparison of data quality of noaa's isis and surfrad networks to nrel's srrl-bms. Tech. rep., National Renewable Energy Lab. (NREL), Golden, CO (United States).

CAISO, 2019. Business practice manual for market operations. http://www.caiso.com/rules/Pages/BusinessPracticeManuals/Default.aspx.

Chu, Y., Pedro, H.T., Coimbra, C.F., 2013. Hybrid intra-hour dni forecasts with sky image processing enhanced by stochastic learning. Sol. Energy 98, 592–603.

Chu, Y., Pedro, H.T., Li, M., Coimbra, C.F., 2015a. Real-time forecasting of solar irradiance ramps with smart image processing. Sol. Energy 114, 91–104.

Chu, Y., Urquhart, B., Gohari, S.M., Pedro, H.T., Kleissl, J., Coimbra, C.F., 2015b. Short-term reforecasting of power output from a 48 mwe solar pv plant. Sol. Energy 112, 68–77.

ERCOT, 2018. ERCOT nodal protocols. http://www.ercot.com/content/wcm/libraries/151823/April_11_2018_Nodal_Protocols.pdf.

Feng, C., Cui, M., Hodge, B.-M., Lu, S., Hamann, H., Zhang, J., 2018. Unsupervised clustering-based short-term solar forecasting. IEEE Trans. Sustain. Energy.

Feng, C., Cui, M., Lee, M., Zhang, J., Hodge, B.-M., Lu, S., Hamann, H.F., 2017. Short-term global horizontal forecasting based on sky imaging and pattern recognition. In: 2017 IEEE Power & Energy Society General Meeting. IEEE, pp. 1–5.

Feng, C., Sun, M., Zhang, J., 2019a. Reinforced deterministic and probabilistic load forecasting via q-learning dynamic model selection. IEEE Trans. Smart Grid.

Feng, C., Yang, D., Hodge, B.-M., Zhang, J., 2019b. Opensolar: Promoting the openness and accessibility of diverse public solar datasets. Sol. Energy.

Feng, C., Zhang, J., 2018. Hourly-similarity based solar forecasting using multi-model machine learning blending. In: 2018 IEEE Power & Energy Society General Meeting. IEEE, pp. 1–5.

Fouilloy, A., Voyant, C., Notton, G., Motte, F., Paoli, C., Nivet, M.-L., Guillot, E., Duchaud, J.-L., 2018. Solar irradiation prediction with machine learning: Forecasting models selection method depending on weather variability. Energy 165, 620–629.

Gigoni, L., Betti, A., Crisostomi, E., Franco, A., Tucci, M., Bizzarri, F., Mucci, D., 2017. Day-ahead hourly forecasting of power generation from photovoltaic plants. IEEE Trans. Sustain. Energy 9 (2), 831–842.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Hossain, M., Mekhilef, S., Danesh, M., Olatomiwa, L., Shamshirband, S., 2017. Application of extreme learning machine for short term output power forecasting of three grid-connected pv systems. J. Cleaner Prod. 167, 395–405.

Kaur, A., Nonnenmacher, L., Pedro, H.T., Coimbra, C.F., 2016. Benefits of solar forecasting for energy imbalance markets. Renewable Energy 86, 819–830.

Lathuilière, S., Mesejo, P., Alameda-Pineda, X., Horaud, R., 2019. A comprehensive analysis of deep regression. IEEE Trans. Pattern Anal. Mach. Intell.

Lauret, P., Voyant, C., Soubdhan, T., David, M., Poggi, P., 2015. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol. Energy 112, 446–457.

Law, E.W., Kay, M., Taylor, R.A., 2016. Calculating the financial value of a concentrated solar thermal plant operated using direct normal irradiance forecasts. Sol. Energy 125, 267–281.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436.

Lorenz, E., Hurka, J., Heinemann, D., Beyer, H.G., 2009. Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 2 (1), 2–10.

Marquez, R., Coimbra, C.F., 2013. Intra-hour dni forecasting based on cloud tracking image analysis. Sol. Energy 91, 327–336.

Pothineni, D., Oswald, M.R., Poland, J., Pollefeys, M., 2018. Kloudnet: Deep learning for sky image analysis and irradiance forecasting. In: German Conference on Pattern Recognition. Springer, pp. 535–551.

Qing, X., Niu, Y., 2018. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 148, 461–468.

Richardson, W., Krishnaswami, H., Vega, R., Cervantes, M., 2017. A low cost, edge computing, all-sky imager for cloud tracking and intra-hour irradiance forecasting. Sustainability 9 (4), 482.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Stein, J.S., Holmgren, W.F., Forbess, J., Hansen, C.W., 2016. Pvlib: Open source photovoltaic performance modeling functions for matlab and python. In: 2016 IEEE 43rd photovoltaic specialists conference. IEEE, pp. 3425–3430.

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

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F., Fouilloy, A., 2017. Machine learning methods for solar radiation forecasting: A review. Renewable Energy 105, 569–582.

Wang, F., Zhen, Z., Mi, Z., Sun, H., Su, S., Yang, G., 2015. Solar irradiance feature extraction and support vector machines based weather status pattern recognition model for short-term photovoltaic power forecasting. Energy Build. 86, 427–438.

Wang, Y., Shen, Y., Mao, S., Chen, X., Zou, H., 2018. Lasso and lstm integrated temporal model for short-term solar intensity forecasting. IEEE Internet Things J. 6 (2), 2933–2944.

Xie, L., Gu, Y., Zhu, X., Genton, M.G., 2013. Short-term spatio-temporal wind power forecast in robust look-ahead power system dispatch. IEEE Trans. Smart Grid 5 (1), 511–520.

Yang, D., 2019. A guideline to solar forecasting research practice: reproducible, operational, probabilistic or physically-based, ensemble, and skill (ropes). J. Renewable Sustain Energy 11 (2), 022701.

Yang, D., Kleissl, J., Gueymard, C.A., Pedro, H.T., Coimbra, C.F., 2018. History and trends in solar irradiance and pv power forecasting: A preliminary assessment and review using text mining. Sol. Energy 168, 60–101.

Yang, D., Wu, E., Kleissl, J., 2019. Operational solar forecasting for the real-time market. Int. J. Forecast.

Zhao, X., Wei, H., Wang, H., Zhu, T., Zhang, K., 2019. 3D-CNN-based feature extraction of ground-based cloud images for direct normal irradiance prediction. Sol. Energy 181, 510–518.