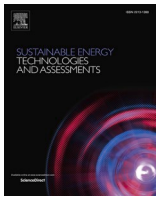
Sustainable Energy Technologies and Assessments 50 (2022) 101864

Contents lists available at ScienceDirect

Sustainable Energy Technologies and Assessments

journal homepage: www.elsevier.com/locate/seta



Sliding window approach with first-order differencing for very short-term solar irradiance forecasting using deep learning models

Ankit Bhatt [a], Weerakorn Ongsakul [a,*], Nimal Madhu M. [b], Jai Govind Singh [a]

[a] Department of Energy, Environment and Climate Change, School of Environment, Resources and Development, Asian Institute of Technology, Pathumthani 12120, Thailand

[b] Center for Computational Engineering and Networking, Amrita School of Engineering, Coimbatore, 641112 Amrita Vishwa Vidyapeetham, India

ARTICLE INFO

Keywords:
Sliding window technique
First-order differencing
Deep neural networks
Multi-step solar irradiance prediction

ABSTRACT

The intermittent behavior of solar is usually responsible for creating uncertainty while generating power. By implementing suitable forecasting techniques for solar irradiance (SI), we can overcome this intermittency which can be helpful in the economic load dispatch as well as to control, manage and optimize the power generation in the microgrid. This paper presents three deep learning (DL) models to forecast SI from 1 step (15-minute) to 6 steps (1 h 30 min) ahead. By implementing the sliding window technique, the input variables are converted into 12 steps lag datasets to train the model whereas outputs are transformed by first-order differencing. A total dataset of 18,277 from average 15-min interval global SI collected during January 1, 2016, to January 6, 2017, from the Asian Institute of Technology (AIT) Metrological station is used to train and evaluate the performance of the DL models. Based on the obtained results from different evaluation parameters such as maximum absolute error, confidence interval (CI), linear regression plot, mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), and R squared, we found that the deep hybrid model consists of convolutional neural network-long short term memory (CNN-LSTM) can outperform during multistep forecasting. The findings of the present work suggest that the proposed deep hybrid LSTM–CNN model is a reliable alternative for very short-term SI prediction due to its high predictive accuracy.

Introduction

The economic and technological growth of a country, substantially dependent upon electricity [1]. Fossil fuels as the primary source of electricity generation, are depleting continuously with rapid growth and also the main cause of global warming due to the greenhouse gas emissions which is a serious threat to the environment [2]. Therefore, in recent years, more inclination is going towards renewable energy (RE) for electricity production [3]. Among several renewable sources, solar energy is the most abundant energy source on Earth which can majorly contribute to fulfilling the electricity demand of the world population [4].

At present due to the several policy supports, large-scale programming of smart grid, and drastic cost reductions, the number of photovoltaic (PV) installations are growing rapidly. In 2019, the global installed capacity from solar reached 586 GW [5]. According to [6], China has a plan to increase solar capacity up to 1300 GW by 2050. Therefore, the PV integration into microgrids is also increasing due to the higher PV adoption. For optimal operation of PV-based microgrids, there is a need for an accurate prediction of supply and demand so that the plant authorities can decide that how much energy is required to purchase. Since higher forecasted error implies higher energy cost in the energy market, therefore, accurate prediction of solar power or energy is highly required for the energy industry due to the intermittent nature of solar energy [4]. Solar irradiance (SI) is an important factor for solar power generation therefore in several microgrid applications such as; demand-supply balancing, economic load dispatching, peak shaving, energy trading in the market, and real-time energy management, SI prediction plays a significant role [7,8].

Depending upon several meteorological parameters of a location such as ambient temperature (T), pressure, wind speed, relative

Abbreviations: AIT, Asian Institute of Technology; ANN, Artificial Neural Network; ARIMA, Auto-regressive Integrated Moving Average; DNN, Deep Neural Network; ELM, Extrement Learning Machine; GRU, Gated Recurrent Unit; BPNN, Back Propagation Neural Network; WPD, Wavelet Packet Decomposition; WMI, Wrapper Mutual Information.

* Corresponding author.
E-mail address: ongsakul@ait.ac.th (W. Ongsakul).

https://doi.org/10.1016/j.seta.2021.101864
Received 9 July 2021; Received in revised form 8 November 2021; Accepted 5 December 2021
Available online 20 December 2021

Nomenclature

CNN_{1D}	One-dimensional convolutional neural network
n	Total number of predicted samples
n RMSE	Normalized root mean square error
r MAE	Relative mean absolute error
r RMSE	Relative root mean square error
X	Actual values of solar irradiance
x_-	Average of the samples (predicted values)
Y	Forecasted values of solar irradiance
z	Confidence level value
σ	Standard deviation of the predicted values
$\sigma/\sqrt{n}$	Standard deviation error

humidity (RH), etc., the value of SI changes therefore by utilizing these meteorological parameters SI can be forecasted. To improve the accuracy of SI forecasting, there are two broad categorizations of forecasting, based on the forecasting horizons and the types of model. In terms of the forecast horizon, forecasting can be categorized as very short-term (few seconds to minutes ahead) utilized in real-time optimization for energy management, short-term (up to 72 h ahead) for unit commitment and intraday market participation, medium-term (one week ahead) for maintenance scheduling, and long term (months or years ahead) for installation and operation of a solar plant [9–12].

Based on the types of model, solar forecasting methods can be classified into several categories [13–15] which are shown in Fig. 1. The persistence model is mostly used as a baseline or reference model to compare and evaluate the performance improvement of other models. Here the SI value at present time 't' is considered as the future (predicting) value at the time '$t+1$' [13]. In [16] it is also highlighted that the forecasted model should be compared with the persistence model. In physical models, the analysis is based on the equivalent circuit models which utilize the numerical weather data for deriving the output power [17]. Since these models consist of complex calculations and need several circuit parameters therefore these models are less robust for PV forecasting applications.

Statistical models are dependent upon the historical time series values of SI which help to construct a mathematical relationship. The forecasting horizon of these models is from 5 min to 6 h [18]. These models mainly consist of linear regression models such as auto-regression (AR), autoregressive integrated moving average (ARIMA) [19], autoregressive moving average (ARMA) [20], exponential smoothing model [21], etc. These linear regression models show poor robustness when the correlation between the meteorological parameters with the irradiance is not strong and if there are missing or incomplete datasets [22,23]. Due to the non-stationary behavior of SI time-series data, these models are not able to capture the non-linearity in data accurately, hence showing low prediction performance [24].

Nowadays machine learning (ML) models are very popular in SI prediction where the forecasted model learns patterns from the historical data, therefore it is also called a data-driven technique. These models can extract the complex and non-linear characteristics from the data. Here, the selection of a proper training dataset plays a significant role in the model accuracy. On the other side, hybrid models utilize the characteristics of multiple models together to improve predictions [25].

Literature work on ML-based solar irradiance forecasting

In the last few years, ML-based forecasting models such as artificial neural network (ANN) [26], support vector machine (SVM) [27], random forest (RF) [28], multi-layer perceptron (MLP) [29], feed-forward neural network (FNN) [30], back-propagation neural network (BPNN) [31], extreme learning machine (ELM) [18] have gained more attention because of their good learning and modeling capability with the complex nonlinear relationships of SI and meteorological data. In [32] three ML techniques (neural network (NN), Gaussian process, and SVM) were compared with linear AR and persistence model for 1-hour ahead SI prediction. They found the minimum relative Root Mean Square Error (r RMSE) as 19.65% from the NN and relative Mean Absolute Error (r MAE) as 12.36% from SVM. Similarly, for short-term forecasting an adaptive feedforward BPNN was proposed in [33].

ML-based methods mostly follow the concept of shallow architecture which has one or no hidden layer, for the prediction of SI. These shallow architectures have the following drawbacks [4,34]:

- Detailed information of the datasets related to the SI is required for better learning of the model which makes feature selection a difficult

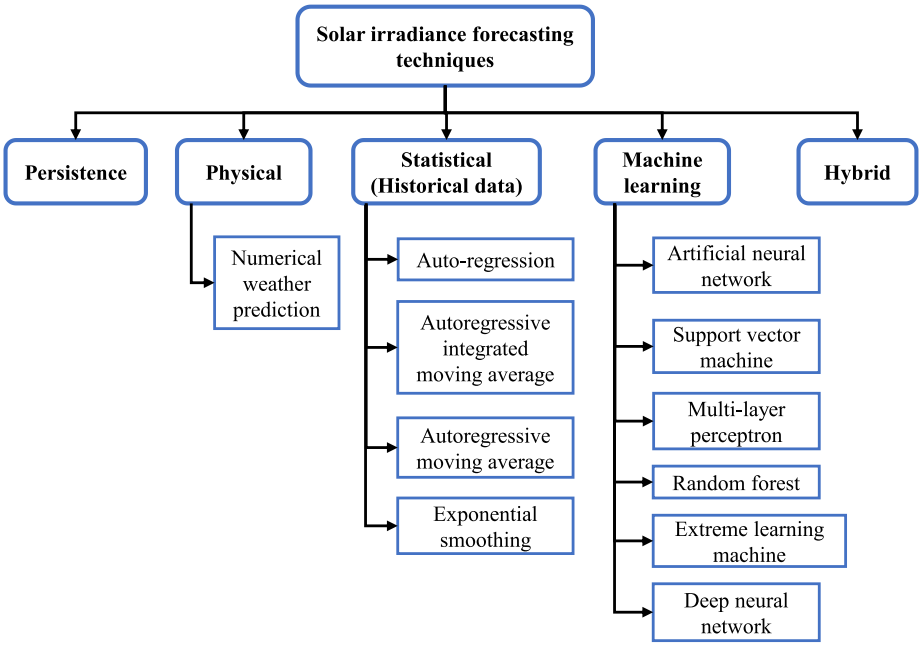


Fig. 1. Different solar irradiance forecasting methods.

process. Therefore, an appropriate input feature selection requires high expert skills which makes these models unreliable for SI prediction.
- Due to less generalization ability, it is difficult for shallow architectures to learn the complex pattern of SI datasets. Therefore, a problem of over-fitting, gradient vanishing, and network training explosion arises in the shallow models.
- For relatively small datasets, shallow models show promising results. On the other side, while handling the huge amount of datasets, shallow models may become unstable and the convergence of parameters may be slow.

Deep neural network (DNN)

To overcome the issues of ML-based models, researchers are inspired towards a more advanced and promising approach to deep learning (DL). DL as a part of ML technology can learn the representation of data by using multiple processing layers at multiple levels [35]. Due to the intermittent nature and the dependency over several factors such as wind speed, T, pressure, and RH, solar data has high-level and non-linear complex characteristics. These characteristics of solar data consist of spatial and temporal information which are extracted by these DL models. Presently research shows that DL models can give superior performance than ML models for time series, classification, and regression-based forecasting problems [36]. For time series forecasting, some DNN models such as long short-term memory (LSTM), one-dimensional convolutional neural network (CNN_{1D}), gated recurrent unit (GRU), and hybrid models are found to be the most powerful and accurate tools [11].

Deep CNN as a part of FNN, shows very promising results during image, audio, and video processing whereas deep LSTM as an advanced version of recurrent neural network (RNN), shows excellent results with text, speech, and time-series where data is in sequential form. In several applications such as image recognition [37], speech recognition [38], brain circuit reconstruction [39], etc. DL has outperformed then ML techniques. Although the performance of DL models improves due to multiple hidden layers the deeper architecture increases the number of parameters to be optimized which further increases the training time of the model. Some recent works of literature highlight the potential of DL in the application of forecasting.

For time-series forecasting, [40] proposed LSTM for the hourly day-ahead SI forecasting and compared it with the persistence algorithm, linear least square regression, and multilayered FNN using the BPNN algorithm. Results showed that by increasing the training data from 2 years to 10 years, LSTM was able to reduce RMSE from 18.34% to 42.9% as compared to BPNN. In [41] a deep LSTM model was applied on satellite image data of 21 locations for irradiance prediction. Here, LSTM outperformed a large number of alternating methods along with 52.2% of prediction skills over the persistence model.

To overcome the gradient vanishing problem caused by simple RNN, the LSTM model was introduced by [42] for short-term SI prediction. Based on hourly forecasted results they concluded that the LSTM model has better performance than RNN based on different error indicating parameters such as RMSE, MAE, and MAPE. To improve the prediction accuracy, one dimensional (1 D) CNN model was introduced by [43] to solve the time-series forecasting problems which also reduced the time complexity.

Among DL models, LSTMs are capable of handling temporal dependencies in data which makes them popular for SI forecasting [40] whereas CNNs are capable to extract spatial features (correlation) from meteorological parameters [44]. Nowadays researchers are looking towards the hybridization of models so that the performance of DL models can be enhanced. Therefore, many works of literature are presenting the performance of hybrid models, a combination of LSTM and CNN which can extract temporal and spatial characteristics of data for time-series prediction analysis.

A hybrid CNN-LSTM model was proposed by [45] to predict short-term SI by collecting datasets with an interval of 1 h from 34 locations. The hybrid model was found to be better than the individual CNN and LSTM models in terms of generalization and performance. For multi-step ahead short-term forecasting of PV power different DNN models were presented by [46]. For performance evaluation, they used datasets with different time intervals from 1-min to 60-min time horizons. Prediction results with a 1-min time interval were found to be more accurate in comparison with other time interval datasets.

Wavelet packet decomposition (WPD) technique was introduced in [47] with hybrid deep models for 1-hr ahead SI prediction with multivariable inputs. WPD based CNN-LSTM-MLP model was found to be the best in comparison with other DL models. [18] proposed a method with the combination of wrapper mutual information (WMI) and ELM for short-term prediction of solar radiation with 5-min and 60-min interval datasets. They performed their model to forecast from 1 step to 6 steps ahead. DL model-based SI forecasting for multi-region with the 1-hr interval of datasets was reported in [48]. They implemented an ensemble empirical mode decomposition technique with the CNN-LSTM model for 1 h ahead of SI prediction. Datasets from 4 different regions were collected for model evaluation.

A comparative analysis of the hybrid DL model (CNN-LSTM) with different models such as RNN, GRU, MLP, etc. for half-hourly global solar radiation prediction, was shown in [49]. They found that CNN can extract input characteristics from predictive variables very nicely whereas LSTM absorbs them for prediction. For model evaluation, they used RMSE, MAE, and MAPE. Similarly, [50] proposed a hybrid DL model (LSTM-CNN) for long-term solar power prediction with the historical datasets from 1990 to 2014 with a 1-hr time interval in Australian locations. The proposed model was compared with RF, statistical analysis, and ANN with RMSE, normalized RMSE (nRMSE), MAPE, and R^2-value.

Apart from SI forecasting, hybrid CNN-LSTM DL models were also introduced in other time-series applications such as wind speed, wind power, energy consumption prediction, etc. In [51] a CNN-LSTM model along with WPD was proposed for wind speed prediction. From 1-step to 3-step wind speed prediction, the proposed model was found to be best. Similarly, for wind power prediction a hybrid (CNN-LSTM) model was developed by [52]. For the prediction of residential energy consumption in France, again a hybrid model with the combination of CNN and LSTM was proposed by [53]. As compared to the conventional DL models, the proposed hybrid model showed improved results.

Though many researchers have applied a hybrid (CNN-LSTM) model for SI forecasting especially for short-term prediction whereas the literature available for a very short-term prediction (15-minute interval in the present study) is very limited. In most of the literature prediction dataset intervals used were 1-hour or 30-minutes whereas it was also found that by reducing the prediction dataset interval further, the forecasting accuracy can also be enhanced. Since, in very short-term forecasting, the forecasting horizon lies from a few seconds to several minutes ahead which may help in electricity pricing, bidding, and real-time energy management in the microgrid. Additionally, in a real-time microgrid energy management system, the load fluctuations are more where instead of 30-minutes or 1-hour ahead prediction of SI, an accurate prediction of SI for the shorter interval (15-minute) ahead can be more effective. Therefore, addressing these gaps is the motivation of the present work. Further, the present work highlights the following contributions to fulfill the above gaps:

- To address very short-term forecasting, 15-minute average datasets are used for SI prediction which is derived by taking the average of actual recorded 5-minute interval datasets from a meteorological station.
- To convert a time-series problem into a supervised learning problem for ML-based models, a sliding window technique with a window size of 12 is introduced here.

- Due to the non-stationary characteristics of time-series data, the mean and variance of the datasets become unstable which may affect the prediction accuracy. Therefore, to make time-series data stationary for better prediction results, the present work introduces first-order differencing during the data processing stage.
- Finally, the performance of the deep CNN-LSTM model is evaluated with a baseline (Persistence) model along with 2 standalone models (deep LSTM and deep CNN) in terms of different indicators. The forecasting accuracy of the proposed model is evaluated from 1-step (15-minutes) to 6-step (1 h 30 min) ahead.

The remainder of the paper is structured as follows: Section 2 presents an architectural framework for the present work. This section shows the importance of the sliding window technique to convert the time-series data into supervised learning data for a better understanding of DL networks during the training phase. Extensive experimental results based on DNNs are presented and discussed in Section 3, followed by a conclusion in Section 4.

Architectural framework for solar irradiance prediction

The basic idea of this study is to implement the application of artificial intelligence through DL for a very short-term prediction (15-minute interval) of global SI. The output power of PV is highly intermittent and fluctuating in nature due to several influential parameters such as SI, T, RH, and other climate factors. Therefore, accurate forecasting of SI can provide better planning and scheduling, energy management, operational cost minimization, safe operation, quality, and balance between supply and demand for microgrid applications. The architectural framework of the present study is shown in Fig. 2.

Data collection and pre-processing

For the collection of datasets, the meteorological station of the Asian Institute of Technology (AIT) situated in the Pathumthani province, Thailand, is considered. The latitude and longitude of the study area are 14.0785^0N and 100.6140^0 E. In the present work, we are only considering the meteorological data from 1 January 2016 to 6 January 2017 with a recording interval of 5 min. Three parameters, the average global SI at 15^0 tilted surface facing south (W/m^2), the average temperature of ambient air (^{0}C), and the average RH (%), are selected for this study. The processing of the recorded data from 5-minute to the average 15-minute interval is also done for the present study. The overall characteristics of the average global SI, average T, and average RH of the collected datasets are presented in Fig. 3 where the x-axis is representing the total datasets and the y-axis is representing the range of actual values for different parameters.

The data categorization for training and testing samples is shown in Table 1. It shows that the total number of training and testing samples consist of 18,228 datasets with an average 15-minute interval time from which 17,934 datasets are used for training from the year 2016 whereas 294 datasets are considered for testing (forecasting) from the first 6 days of the year 2017. Further, we are splitting training datasets into 80% training and 20% validation datasets to train and validate our model. The time duration for each day starts from hour 00 to 23 h, but for the forecasting of SI, we are only considering the time from 6 AM to 6 PM because for the rest of the time SI is not available for generating power.

The inputs for the DL networks are the data collected from the AIT's meteorological station. The summary of the input and output parameters for the models is illustrated in Table 2 whereas the statistical parameters of the collected dataset are shown in Table 3. Table 2 clearly shows that the output of the forecasting model will be 6 steps ahead of SI with a 15-minute time interval whereas the time-series dataset of the previous 3 h with 12 steps lag, is considered as the input dataset. After categorizing the datasets into the input and output variables, an appropriate normalization process is required to minimize the undesirable effect of raw data to make the ML model robust and efficient. Therefore, the normalization of the input and output datasets is done by using Eq. (1) for the range of 0 to 1.

$$x(normalized) = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

where, x_{max} and x_{min} are defined as the maximum and minimum values of input and output vector ' x ' respectively.

Sliding window technique

SI forecasting is a time series analysis that needs to be framed into a

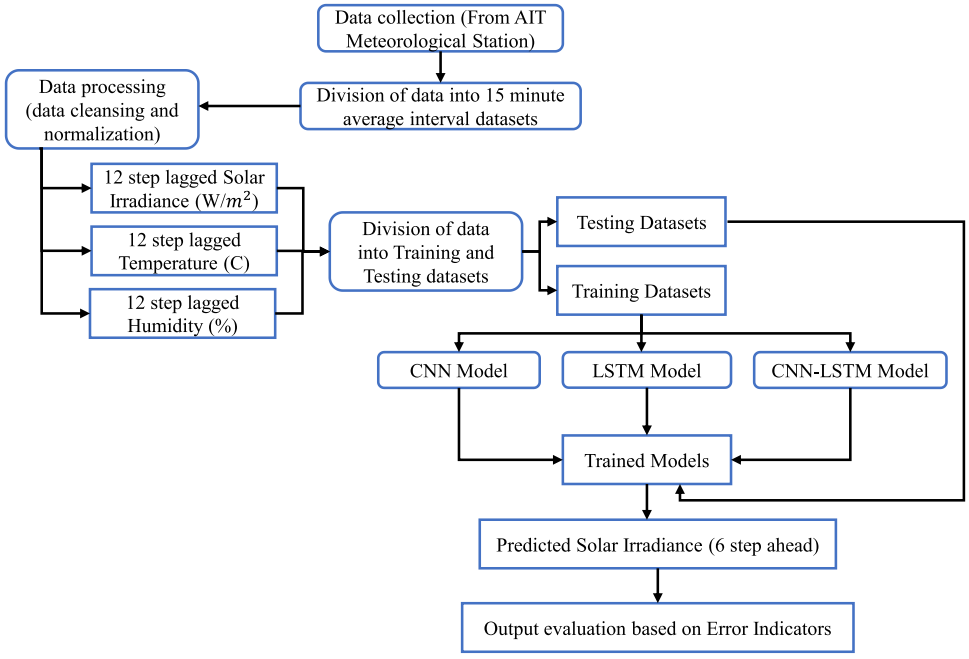


Fig. 2. An architectural framework of the present work.

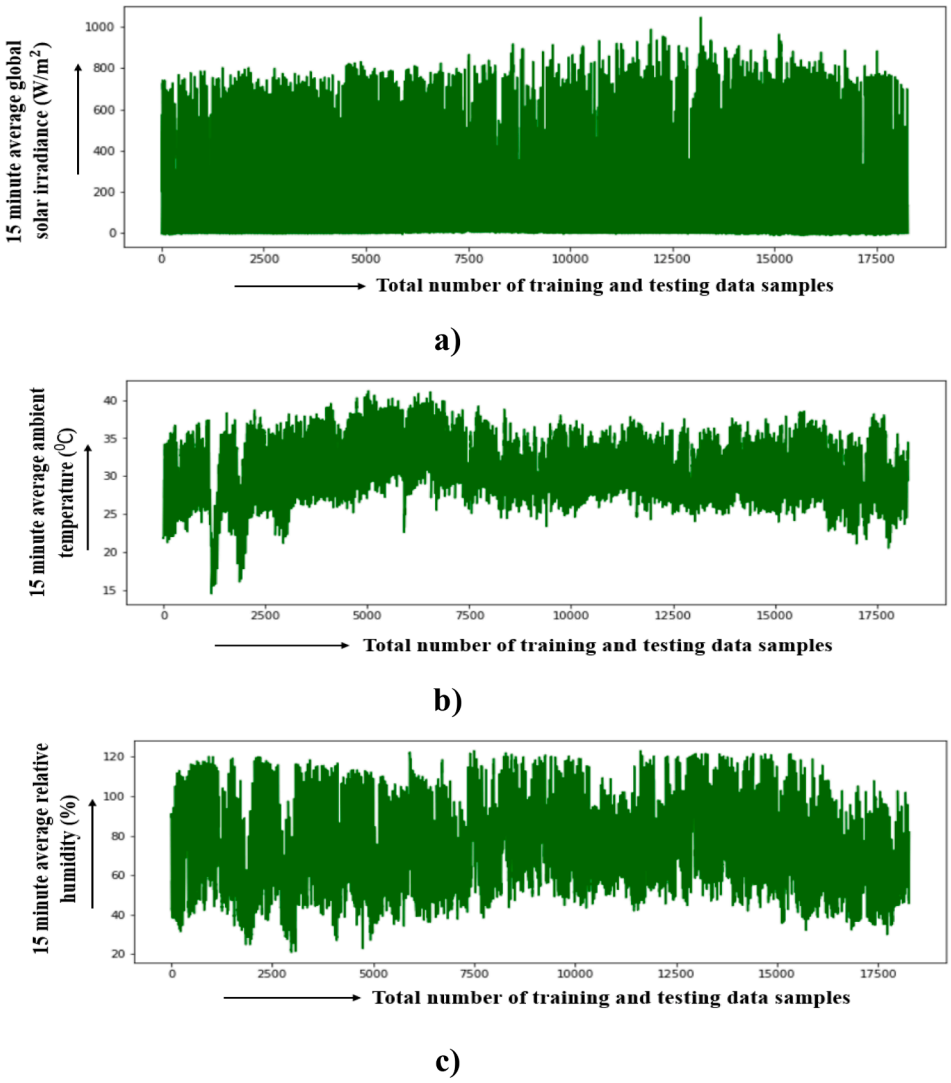


Fig. 3. Characteristics of **a)** average global SI **b)** average T **c)** average RH.

Table 1
Data categorization.

Year	Interval	Training and Testing Datasets	Training Datasets		Testing (Forecasting) Datasets	Testing (Forecasting) Datasets Period
			Training (80%)	Validation (20%)		
2016–17	Average 15 min	18,228	14,347	3587	294	First 6 days of January 2017

Table 2
Inputs and output for the proposed models.

Input Variables	Parameter	Output Variable	Parameter
X_1	Lagged SI (12 steps)	X_4	SI (6 steps ahead)
X_2	Lagged Temperature (12 steps)		
X_3	Lagged Humidity (12 steps)		

supervised learning problem for ML models. By using the sliding (moving) window technique, time-series datasets can be framed into supervised learning for multivariate data and multi-step forecasting. Fig. 4 shows the process of the sliding window for SI forecasting where we are using 3 input variables, SI (average solar irradiance), T (average ambient temperature), and RH (average relative humidity) with window size 12 (12 step lag) and trying to predict SI for 6 steps ahead. After dividing whole datasets using this method, each window is utilized to train and update the model. After completing each computation, the "window" shifts to a new position by one step. [54] implemented a sliding window approach for a comprehensive evaluation of the accuracy and efficiency during SI forecasting using RNN and ANN. [55] also used this method with window size 7 for predicting solar radiation using a DNN. For the present study since we are using 12 step lag input datasets and 6 step ahead output datasets, therefore, the size of the sliding window for input and output is 12 and 6 respectively.

Since time series datasets usually contain trends that are responsible for a varying mean and seasonality which is responsible for a changing

Table 3
Statistical parameters of the collected datasets from the AIT meteorology station.

Parameters		Datasets (Average 15 min interval)
Average global SI at 15^0 tilted surface facing south (W/m^2)	Min	−11.42
	Max	1046.00
	Mean	328.95
	Standard Deviation	244.24
Average temperature of ambient air (^0C)	Min	14.50
	Max	41.24
	Mean	31.69
	Standard Deviation	3.96
Average RH (%)	Min	20.83
	Max	123.10
	Mean	67.28
	Standard Deviation	22.33

variance over time, therefore time-series datasets are called non-stationary. By making datasets to be stationary, mean and variance can be stable which helps the model for better prediction. By introducing differencing for data transformation, we can make time-series data stationary. Therefore, in the present work, we are introducing first-order differencing for output datasets during the training phase where the previous observation will be subtracted from the current observation while during the testing (prediction) phase we are using the inverting process for converting back these difference datasets into its original scale.

Deep learning (DL) model selection

DL is defined as a sub-field of ML. It can learn and feature extraction automatically from the available raw input datasets [55]. DL is the advanced architecture of NN which can overcome the over-fitting and gradient diminishing problem faced by traditional NN [56]. Due to the availability of a large number of datasets, many works of literature have shown their interest in various DL applications to improve the accuracy of the prediction. They have shown that LSTM, CNN_{1D}, and other hybrid structures including CNN-LSTM with deep structures can forecast time series problems more effectively [57]. Present work focuses on the designing of three DL models: Deep LSTM, Deep CNN, and Deep CNN-LSTM, which are compared by the Persistence model as a baseline model for comparative analysis during a very short-term SI prediction.

Deep LSTM model

A deep LSTM model with 2 hidden layers along with dropouts and 2 dense layers is introduced here to predict the SI for 6 steps ahead. The proposed model consists of 8 layers including the input and output layers. Here, LSTM Layer 1 & LSTM Layer 2 are the fully connected layers that help to learn the sequence of the data whereas the dropout helps to overcome the over-fitting issue and dense layers help to linearize the output. At the initial stage, the three variables (average SI, average T, and average RH) are directly fed to the input layers with a lagged sequence of the previous 12 steps (3 h) and directly fed to the first hidden LSTM Layer 1. Before feeding the input data to the hidden layers, the shape of the data must be of 3 dimensional (sample, time steps, and features) format. With the dropout of 20% for both of the hidden LSTM layers, some random neurons are dropped during the training phase which helps to overcome the overfitting issue of the data. The dense layer performs linear operation before it finally approaches the output layer. For hidden LSTM layers we are using default activation functions of LSTM architecture which are tanh and sigmoid functions whereas for output we are using linear activation functions. Here, "MSE" is considered for the error evaluation while the number of epochs and batch size is selected as 500 and 128 respectively during the entire training process based on trial and error procedure. Fig. 5 shows the architecture of the proposed deep LSTM model for the present study.

Deep CNN model

A deep CNN model that has been introduced for the experiment in the present study, consists of two hidden layers (Conv. 1D layer) along with two max-pooling layers, three dropout layers, and three dense layers. To reduce the possibility of over-fitting, here also we are using dropout with a value of 10% for both of the hidden layers (Conv. 1D layer) and after 2 dense layers. For all hidden and dense layers, we are using rectified linear unit (ReLU) as an activation function whereas, for output, a linear activation function is used. Based on "MSE" we are measuring the error during the training while the number of epochs and batch size is selected as 500 and 512 respectively during the entire training process based on trial and error procedure. Fig. 6 shows the architecture of the proposed deep CNN model.

Deep CNN-LSTM hybrid model

This model consists of CNN and LSTM architecture together. The CNN model works with the same principle as described in the previous section 2.3.2. However, for the LSTM model, the inputs are provided from previous dense layers which are getting inputs from CNN pooling layers. Here, the 1D convolutional layer of CNN is utilized for classification and pattern extraction from the input datasets while the LSTM

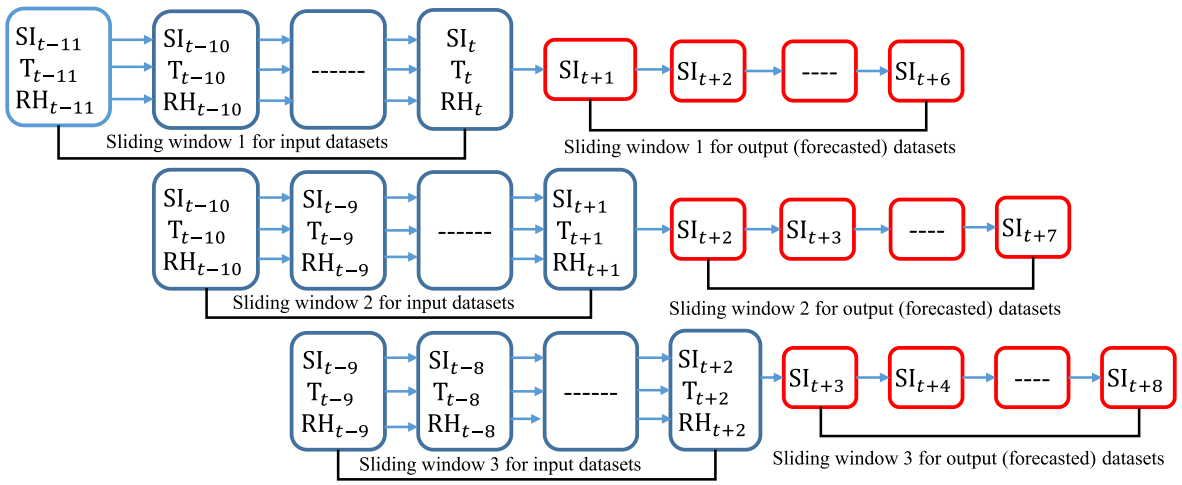


Fig. 4. A sequence of input and output (forecasted value) based on the sliding window technique.

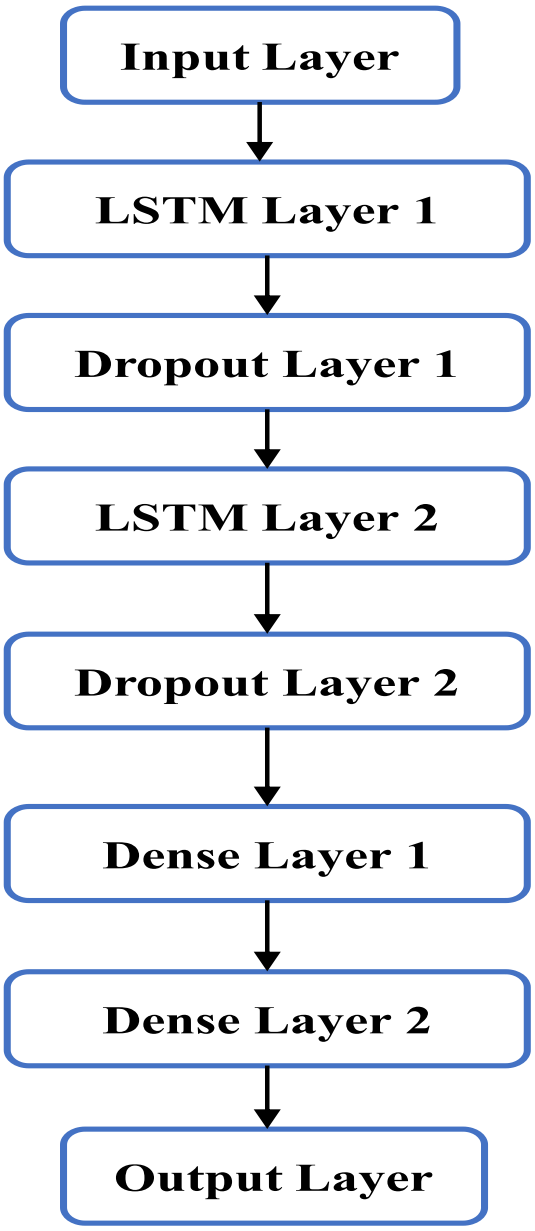


Fig. 5. Deep LSTM model architecture.

helps to predict the input sequence. Dropout values for LSTM layers are selected as 20% whereas for the rest of the hidden layers it is 50% based on trial and error procedure. For Conv. 1D layer the selected activation function is ReLU, whereas for LSTM the default activation functions are used. For output, we are again using the linear activation function. "MSE" is the error measurement criterion while the number of epochs and batch size is selected as 500 and 128 respectively during the entire training process based on trial and error procedure. Fig. 7 displays the architecture of the proposed deep CNN-LSTM model.

The simulation and coding for the chosen DL architectures are done by the Python Jupyter Notebook with the Intel i5-7200U CPU of 2.50 GHz. This software contains many libraries such as NumPy, TensorFlow, Matplotlib, Keras, etc. We are also using Adam optimizer with a learning rate of 0.001. For the performance evaluation of the proposed 3 DL models, we are implementing RMSE, MAE, MAPE, and R-Squared (R^2) as the error indicating parameters.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(Y_j - Y_j^{'}\right)^2} \tag{2}$$

$$\text{MAE} = \frac{1}{n}\sum_{j=1}^{n}\left|Y_j - Y_j^{'}\right| \tag{3}$$

$$\text{MAPE} = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{Y_j - Y_j^{'}}{Y_j}\right| *100 \tag{4}$$

where n denotes the sample size, Y_j is the actual value and $Y_j^{'}$ is the predicted value.

Results and discussion

The obtained results reflect the performance outcome of the proposed deep models. For the performance evaluation of the developed models as discussed in sections 2.3.1, 2.3.2, and 2.3.3, the accuracy measurements are based on different error indicating parameters. During performance evaluation of the SI prediction models, the data of the first 6 days of January-2017 is considered as sample input for testing our models. For each particular day, the time considered for sample data is from 6 AM to 6 PM. Since the forecasting will be done on an average 15-minute time interval basis, therefore the sample datasets per day are 49 and the total sample datasets are 49*6 = 294. Table 4 displays the details of the testing datasets for the present study. Based on the splitting of data into training and validation datasets, we train our model first and then compare the prediction results with the baseline model.

Persistence vs proposed 3 deep models

In the first stage of the experiment, the results obtained by using the

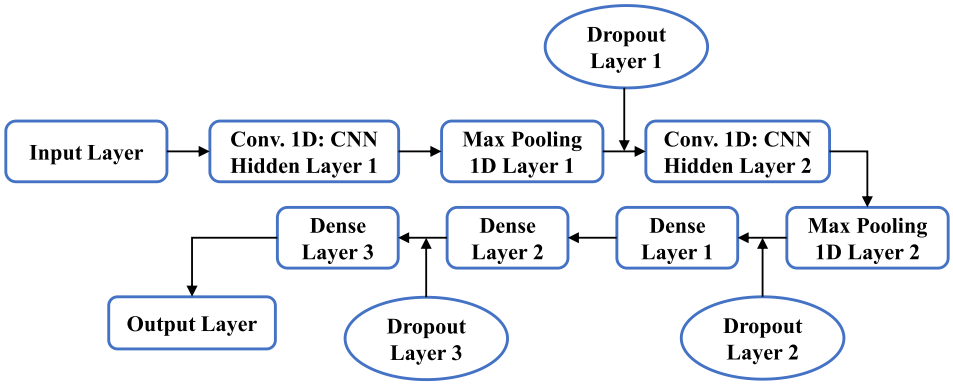


Fig. 6. Deep CNN model architecture.

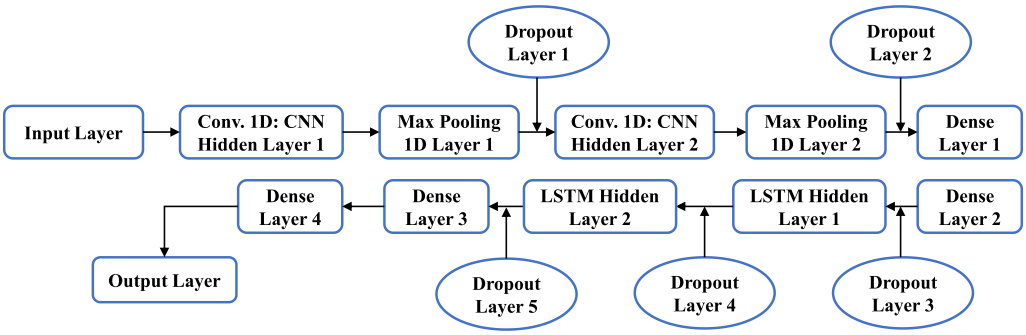


Fig. 7. Deep CNN-LSTM model architecture.

Table 4
Testing sample details of the experiment.

No	Sample No	Date	Time (hour)
1	1–49	January 1, 2017	6:00 – 18:00
2	50–98	January 2, 2017	6:00 – 18:00
3	99–147	January 3, 2017	6:00 – 18:00
4	148–196	January 4, 2017	6:00 – 18:00
5	197–245	January 5, 2017	6:00 – 18:00
6	246–294	January 6, 2017	6:00 – 18:00

three deep models with average 15-minute interval datasets are compared with the baseline model (Persistence) in terms of MAE (W/m^2), RMSE (W/m^2), and MAPE (%). Fig. 8 shows that during 1 step (15-min) ahead prediction all the three proposed DL models can outperform the baseline model in terms of different error parameters. It also shows that out of three deep models, CNN-LSTM (hybrid) model has the minimum value of error parameters which indicates the outperformance of the CNN-LSTM model for prediction accuracy. Further, the performance of the proposed deep models for multistep forecasting is also presented and compared in the below sections.

Performance analysis of proposed 3 deep models for multistep forecasting

Based on the comparison with the baseline model, the deep CNN-LSTM model is found to be the best prediction model. To present the experimental results obtained by using the Deep CNN-LSTM model with average 15-minute interval datasets with various accuracy indexes of the experiment, Fig. 9 presents the forecasted SI against the actual values for the testing period of 6 days with 294 test samples. Results comprise the forecasting horizon from 1 step (15 min) to 6 steps (1 hr 30 min) which shows that 1 step ahead forecasting has better forecasting ability than higher forecasting steps. To show the forecasting results more clearly, Fig. 10 shows the multistep forecasting results for 1 day with 49 test samples. It clearly shows that for longer time horizons, the forecasted values are showing higher deviations from their actual value especially during fluctuations in the actual value of SI.

To show the forecasted error during multistep forecasting of SI using the Deep CNN-LSTM model, Fig. 11 indicates several variations in forecasted error from positive (+) to negative (-) values during different time step horizons. It also states that there is a significant amount of increment in the magnitude of the forecasted error as the number of time steps are increasing. As shown in Table 5, the magnitude of the maximum forecasted error (Max Absolute Error) during 1 step ahead forecasting is 26.25 W/m^2 whereas, for 6 steps ahead, it is 163.02 W/m^2 which is more than 6 times.

To show the linear dependency between the forecasted SI and the actual SI, Fig. 12 presents a linear regression plot during multistep forecasting using the Deep CNN-LSTM model. Table 5 shows the linear dependency of forecasted values (Y) on actual values (X) in terms of different slopes and intercepts during multistep forecasting. It also shows that for 1 step ahead prediction model reflects the higher

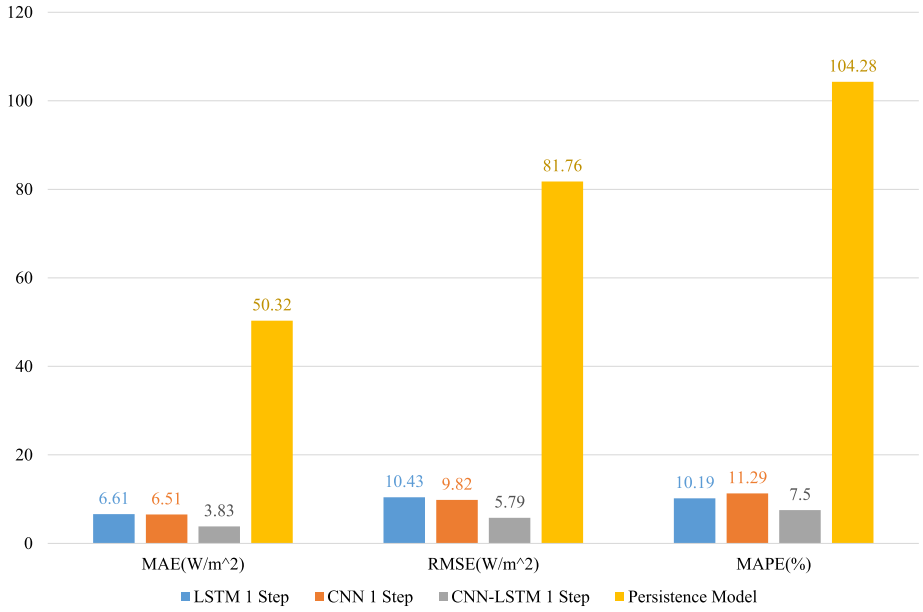


Fig. 8. Comparative analysis of deep models with baseline model for 1 step ahead prediction.

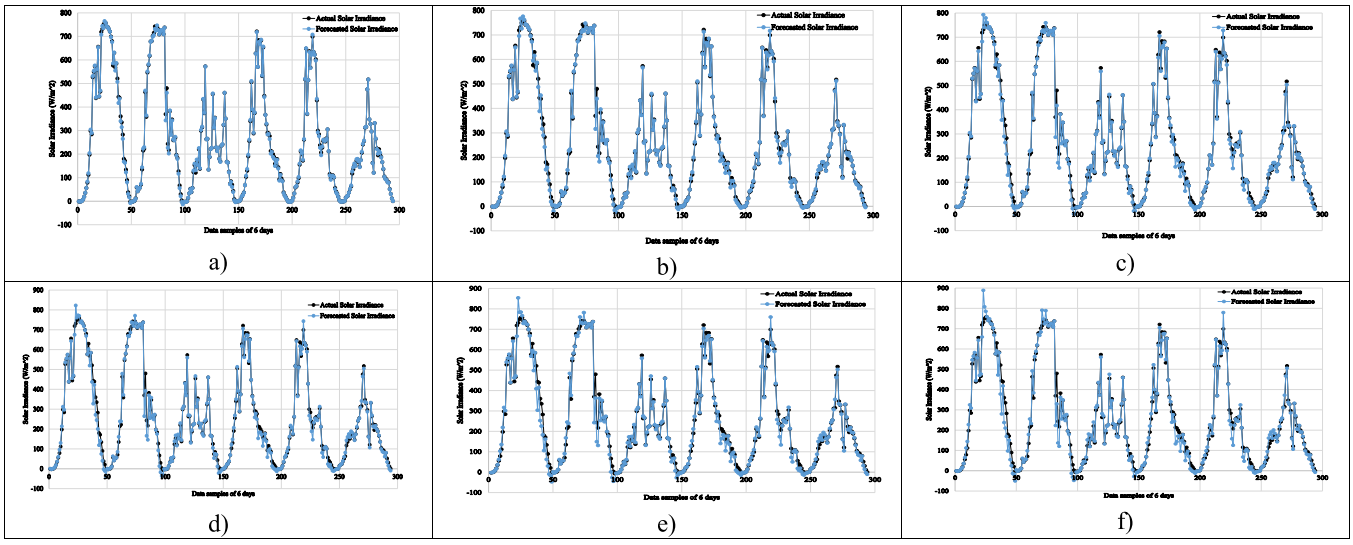


Fig. 9. Forecasting of SI for a) 1 step b) 2 step c) 3 step d) 4 step e) 5 step, and f) 6 step ahead during 6 days using deep CNN-LSTM model.

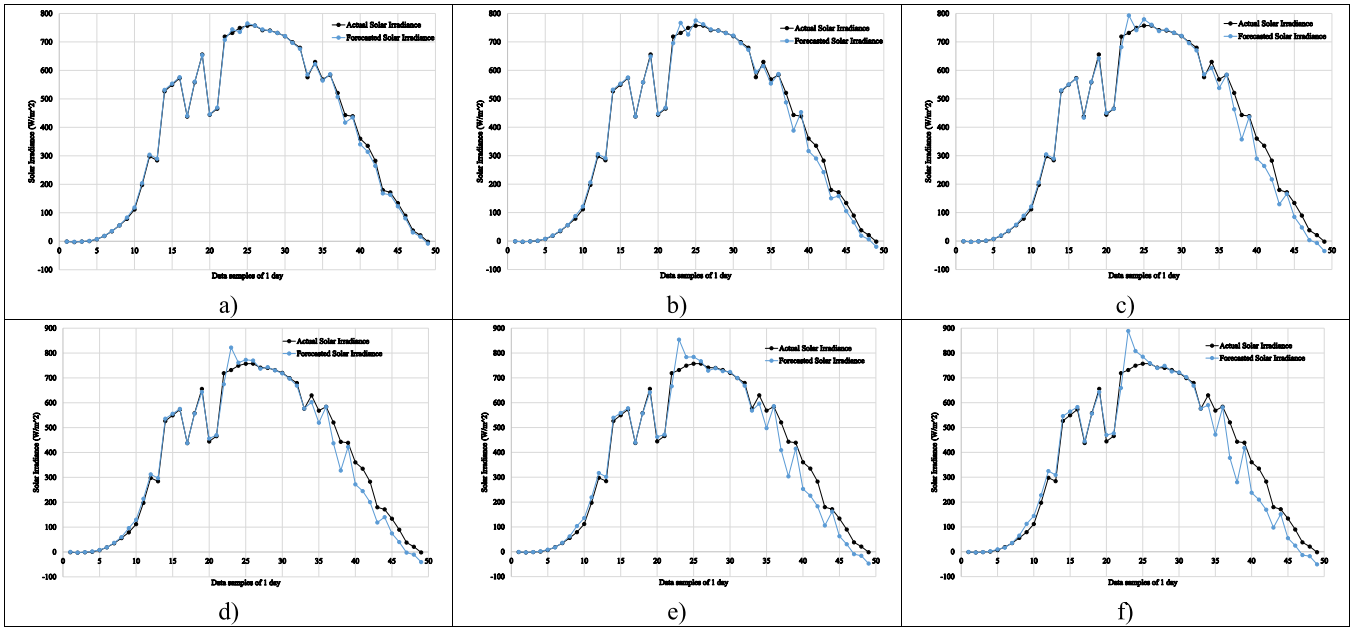


Fig. 10. Forecasting of SI for a) 1 step b) 2 step c) 3 step d) 4 step e) 5 step, and f) 6 step ahead during 1 day using deep CNN-LSTM model.

accuracy whereas for longer time steps, for example, 6 steps ahead, the model reflects the lower accuracy with a high deviation of the sample points from the best fit line. Linear regression equations also explain that based on the increasing number of the prediction steps, the magnitude of the slope and the intercept also increases from [0.999, 0.6159] to [1.0092, 9.2398].

For optimal control and management of PV plants, it is essential to estimate the uncertainties related to SI forecasting. By calculating the confidence interval (CI), we can quantify the uncertainty of an estimate [58]. Smaller CI signifies that the estimate is more precise whereas larger CI signifies a less precise estimate. Here, we are using a 95% CI which means that the mean of the entire population will fall within this range however it does not mean that 95% of observations fall within this range. The lower and upper bounds of CI are calculated by Eqs. (5) and Eq. (6), respectively.

$$\text{Lower Bound} = \underline{x} - z^* \frac{\sigma}{\sqrt{n}} \tag{5}$$

$$\text{Upper Bound} = \underline{x} + z^* \frac{\sigma}{\sqrt{n}} \tag{6}$$

Here, x_{-} is the mean (average) of the samples (predicted values), z is the confidence level value which is equal to 1.96 (for 95% CI), and $\sigma/\sqrt{n}$ is the standard deviation error, where σ is the standard deviation of the predicted values and n is the total number of predicted samples. Table 5 clearly shows that while proceeding to the higher time steps prediction, the CI is continuously increasing from 26.60 W/m^2 to 27.15 W/m^2, which further shows that the forecasting results for 1 step ahead are more precise and accurate than the longer time steps. CI also states that there is 95% confidence that the mean of the predicted values will exist between lower and upper bounds defined by the CI.

Similar experiments are also performed on deep LSTM and deep CNN models for multistep forecasting. Table 6 and Table 7 shows that the magnitude of the forecasted error (Max Absolute Error) during 1 step ahead forecasting is 34 W/m^2 and 29.88 W/m^2 whereas, for 6 steps ahead, it is 212.41 W/m^2 and 138.84 W/m^2 from deep LSTM and deep

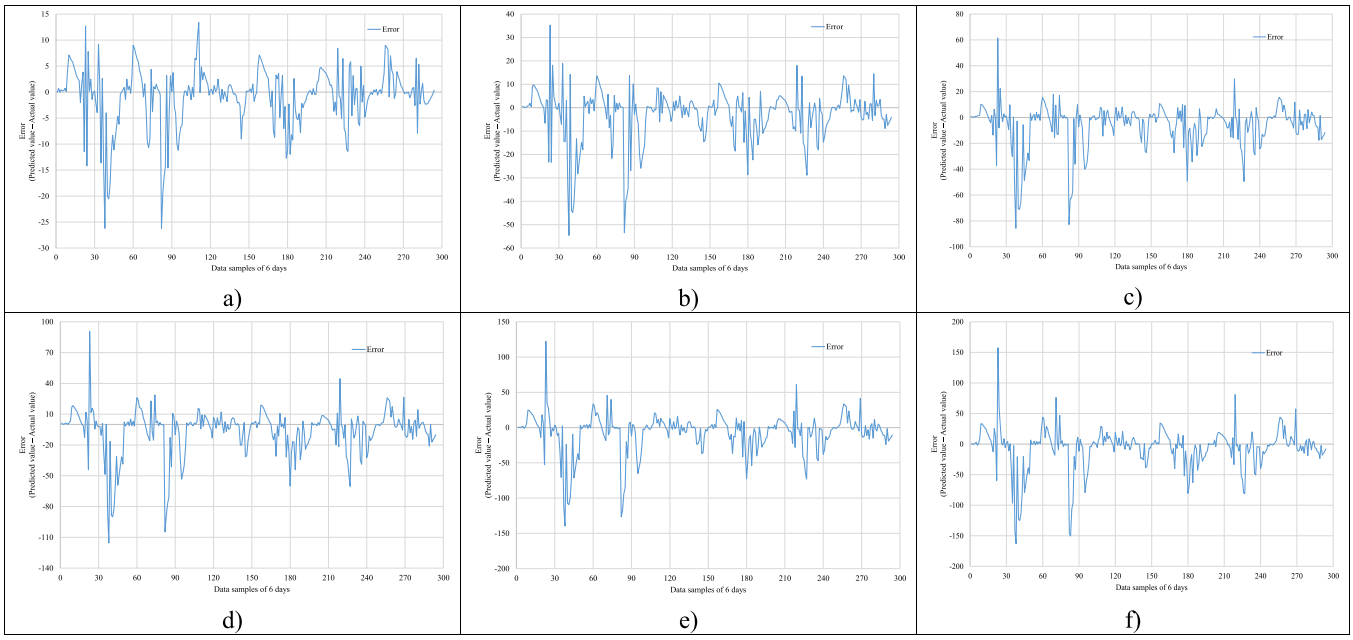


Fig. 11. Error plot for a) 1 step b) 2 step c) 3 step d) 4 step e) 5 step, and f) 6 step ahead forecasting during 6 days using deep CNN-LSTM model.

Table 5
Maximum absolute error, CI, and linear regression equation during multistep forecasting using deep CNN-LSTM model.

Step Ahead	Max Absolute Error (W/m^2)	Mean Predicted (W/m^2)	StandardDeviation Predicted (W/m^2)	CI (95%) (W/m^2)	Lower Bound (W/m^2)	Upper Bound (W/m^2)	Linear Regression Equation
1	26.25	257.92	231.79	26.60	231.43	284.42	Y = 0.999X-0.6159
2	54.58	255.57	232.26	26.66	229.03	282.13	Y = 1.0003X-3.2351
3	85.71	252.51	233.10	26.75	225.86	279.16	Y = 1.0024X-6.8401
4	115.49	252.13	233.87	26.84	225.40	278.87	Y = 1.0037X-7.5673
5	139.52	251.56	234.96	26.97	224.70	278.42	Y = 1.0057X-8.6465
6	163.02	251.87	236.54	27.15	224.84	278.91	Y = 1.0092X-9.2398

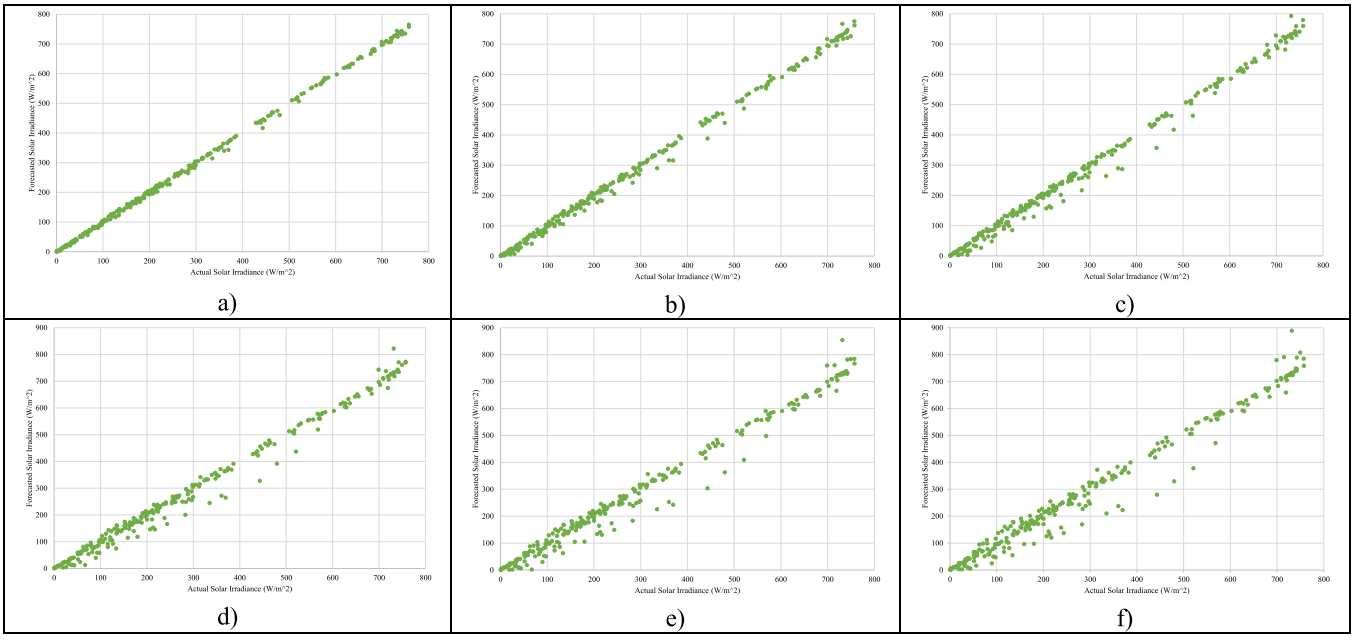


Fig. 12. Linear regression plot for a) 1 step b) 2 step c) 3 step d) 4 step e) 5 step, and f) 6 step ahead forecasting using deep CNN-LSTM model.

Table 6
Maximum absolute error, CI, and linear regression equation during multistep forecasting using deep LSTM model.

Step Ahead	Max Absolute Error (W/m^2)	Mean Predicted (W/m^2)	StandardDeviation Predicted (W/m^2)	CI (95%) (W/m^2)	Lower Bound (W/m^2)	Upper Bound (W/m^2)	Linear Regression Equation
1	34.00	260.04	234.02	26.86	233.28	286.79	Y = 1.0081X-0.7946
2	70.96	262.78	237.45	27.25	235.63	289.92	Y = 1.0204X-1.2368
3	103.51	263.97	240.67	27.62	236.46	291.48	Y = 1.0297X-2.4412
4	135.19	264.30	242.96	27.89	236.53	292.07	Y = 1.0339X-3.1997
5	176.34	266.16	247.63	28.42	237.85	294.47	Y = 1.0449X-4.1817
6	212.41	267.31	251.28	28.84	238.59	296.03	Y = 1.0515X-4.7452

Table 7
Maximum absolute error, CI, and linear regression equation during multistep forecasting using deep CNN model.

Step Ahead	Max Absolute Error (W/m^2)	Mean Predicted (W/m^2)	StandardDeviation Predicted (W/m^2)	CI (95%) (W/m^2)	Lower Bound (W/m^2)	Upper Bound (W/m^2)	Linear Regression Equation
1	29.88	259.08	233.76	26.83	232.36	285.81	Y = 1.0071X-1.4758
2	52.46	259.79	235.52	27.03	232.86	286.71	Y = 1.0126X-2.2097
3	75.60	260.23	237.71	27.28	233.05	287.39	Y = 1.0188X-3.3689
4	96.98	261.08	239.75	27.52	233.68	288.49	Y = 1.0233X-3.6911
5	117.86	261.62	242.25	27.80	233.92	289.31	Y = 1.0283X-4.4314
6	138.84	262.57	244.59	28.07	234.61	290.53	Y = 1.0324X-4.6229

CNN model respectively. Linear regression equations also explain that based on the increasing number of prediction steps, the magnitude of the slopes and the intercepts are also increasing from [1.0081, 0.7946] to [1.0515, 4.7452] for the deep LSTM model whereas for the deep CNN model it ranges from [1.0071, 1.4758] to [1.0324, 4.6229]. Further, these tables also show that while proceeding from 1 step to 6 steps ahead prediction, the CI also increases from 26.86 W/m^2 to 28.42 W/m^2 for the deep LSTM whereas for the deep CNN it varies from 26.83 W/m^2 to 27.80 W/m^2. These results show that the forecasting results for 1 step ahead are more precise and accurate than the longer time steps. While comparing the results of 95% CI for Deep CNN-LSTM with Deep LSTM and Deep CNN it further shows that Deep CNN-LSTM can show better CI for all time steps horizons compared to Deep LSTM and Deep CNN model.

Comparative analysis of proposed 3 deep models based on error indicators

The performance of the developed models with 15-minute time interval datasets by comparing with different accuracy index MAE, RMSE, MAPE, and R^2 is shown in Table 8. The comparison shows that the Deep CNN-LSTM is the best performing model in terms of MAE, RMSE, and R^2 for all 6-time steps horizons whereas in terms of MAPE, there are some fluctuations in the results of Deep LSTM and Deep CNN during longer time horizons. Deep CNN-LSTM shows a gradual increase in the value of MAPE as the number of forecasting steps are increasing. This signifies that the Deep CNN-LSTM is more stable and accurate in comparison to the other two models.

For further analysis and representation, Fig. 13 shows the performance of the 3 models during multistep forecasting with the help of bar charts whereas Fig. 14 shows the error improvement in Deep CNN-LSTM and Deep CNN to the Deep LSTM model because the study shows that Deep LSTM has lower performance ability than the Deep CNN and Deep CNN-LSTM model. MAPE with negative representation shows that Deep LSTM is showing better results in comparison with other models. To compare the maximum error improvements of the models, the analysis is shown in the sections below:

Deep LSTM versus deep CNN

This analysis shows the error improvement of the Deep CNN with the Deep LSTM model. Deep CNN shows around 1.5% improvement in MAE and 5.8% in RMSE whereas for MAPE, LSTM shows around 11% improved results during 1 step ahead prediction. For 2 steps ahead, CNN shows around 4% improvement in MAE, 12% in RMSE, and 28% in MAPE. For 3 steps ahead, CNN shows around 9% improvement in MAE and 15% in RMSE whereas, for MAPE, LSTM shows around 82% improved results. For 4 steps ahead, CNN shows around 10% improvement in MAE and 14% in RMSE whereas, for MAPE, LSTM shows around 30% improved results. For 5 steps ahead, CNN shows around 13% improvement in MAE and 18% in RMSE whereas, for MAPE, LSTM shows around 70% improved results. Similarly, for 6 steps ahead, CNN shows around 12% improvement in MAE and 18% in RMSE whereas for MAPE, LSTM shows around 184% improved results.

Deep LSTM versus deep CNN-LSTM

This analysis shows the error improvement of the Deep CNN-LSTM with the Deep LSTM model. During 1 step ahead forecasting, Deep CNN-LSTM shows around 42% improvement in MAE, 44% in RMSE, and 26% in MAPE. For 2 steps ahead, CNN-LSTM shows around 40% improvement in MAE, 43% in RMSE, and 62% in MAPE. For 3 steps ahead, CNN-LSTM shows around 42% improvement in MAE and 42% in RMSE whereas, for MAPE, LSTM shows around 11% improved results. For 4 steps ahead, CNN-LSTM shows around 43% improvement in MAE, 42% in RMSE, and 29% in MAPE. For 5 steps ahead, CNN-LSTM shows around 45% improvement in MAE and 44% in RMSE whereas, for

Table 8
Comparative analysis based on different error indicators.

Step Ahead	MAE (W/m^2)			RMSE (W/m^2)			MAPE (%)			R^2		
	LSTM	CNN	CNN-LSTM	LSTM	CNN	CNN-LSTM	LSTM	CNN	CNN-LSTM	LSTM	CNN	CNN-LSTM
1	6.61	6.51	3.83	10.43	9.82	5.79	10.19	11.29	7.50	0.998	0.998	0.999
2	12.15	11.69	7.32	20.46	18.09	11.71	51.74	37.45	19.64	0.993	0.994	0.997
3	18.39	16.73	10.61	31.12	26.41	18.18	28.62	52.21	31.87	0.984	0.988	0.994
4	24.16	21.81	13.68	40.33	34.53	23.48	70.51	91.84	50.08	0.974	0.979	0.990
5	31.21	27.26	17.01	52.50	43.11	29.25	30.99	52.79	55.47	0.957	0.969	0.985
6	36.89	32.38	20.07	62.27	50.99	34.89	44.81	127.29	59.28	0.942	0.957	0.979

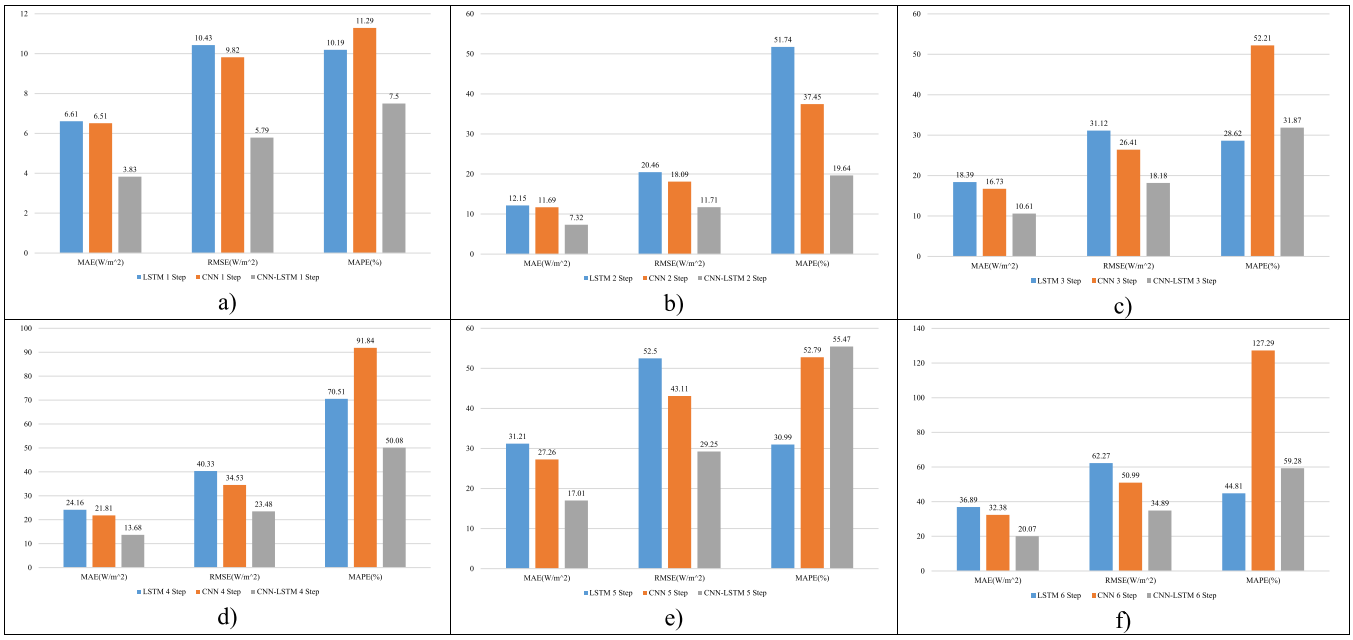


Fig. 13. Performance comparison of 3 proposed deep models for a) 1 step b) 2 step c) 3 step d) 4 step e) 5 step, and f) 6 step ahead forecasting.

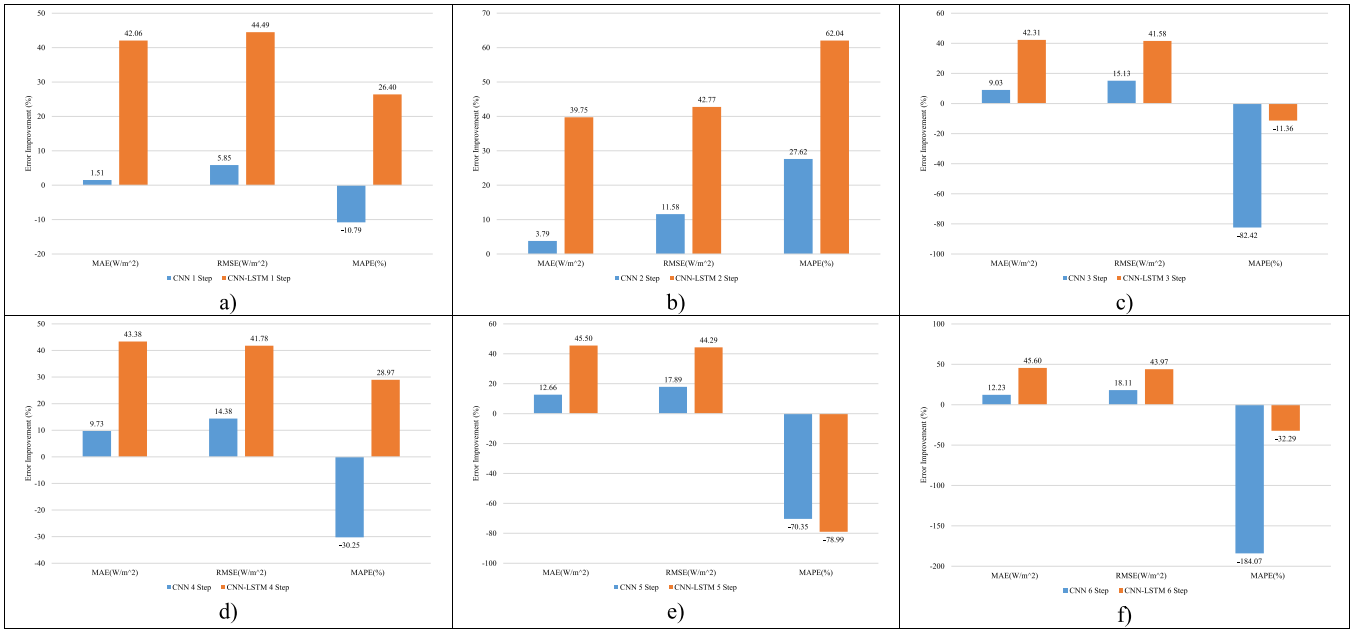


Fig. 14. Error improvement analysis of deep CNN-LSTM and deep CNN model with deep LSTM model for a) 1 step b) 2 step c) 3 step d) 4 step e) 5 step, and e) 6 step ahead forecasting.

MAPE, LSTM shows around 79% improved results. Similarly, for 6 steps ahead, CNN-LSTM shows around 46% improvement in MAE, 44% in RMSE whereas for MAPE, LSTM shows around 32% improved results.

Overall analysis shows that during 1 step, 2 steps, and 4 steps ahead forecasting, Deep CNN-LSTM is outperforming in comparison with Deep LSTM and Deep CNN model in terms of all error indicating parameters whereas, during 3 steps, 5 steps and 6 steps ahead forecasting, Deep CNN-LSTM outperforms only in terms of MAE, RMSE and R^2 values while MAPE values are little bit unsatisfactory.

Comparison between proposed work and existing literature

The present work also compares the prediction outcomes of its best-performing model, Deep CNN-LSTM with some recent literature that has also adopted DNNs for forecasting. Most of the works of literature have proposed CNN-LSTM and LSTM models for their research with a different interval of datasets. Some literature also mentions that by reducing the dataset interval, the model can give better forecasting accuracy. Existing work evaluates the performance of their models with different error indicators (RMSE, MAE, MAPE, R^2). The present work proposes very short-term forecasting from 1 Step ahead (15-min ahead) to 6 steps ahead (1 hr 30 min) based on 3 DNNs and finds that the Deep CNN-LSTM performs better in comparison with some recent existing works which is presented in Table 9.

Table 9
Comparative analysis with some existing literature.

Literature	Types of DNN	Dataset Interval	Evaluation Parameters			
			RMSE (W/m^2)	MAE (W/m^2)	MAPE (%)	R^2 (%)
[45]	CNN-LSTM	1-hr	**48.13**	**25.24**	–	**98.25**
	LSTM	1-hr	53.48	29.14	–	97.84
	CNN	1-hr	60.67	36.22	–	97.25
[18]	ELM	5-min	**14.17**	**–**	**7.4**	**96.7**
[49]	CNN-LSTM	30-min	**8.19**	**6.66**	**4.67**	**100**
	CNN	30-min	9.99	7.53	4.84	99.9
	LSTM	30-min	21.05	18.34	6.48	99.9
[42]	LSTM	1-hr	**22.13**	**18.72**	**5.9**	**99.00**
[50]	CNN-LSTM	1-hr	**3.89**	**–**	**2.83**	**90.00**
Proposed Work	**CNN-LSTM (1 Step Ahead)**	**15-min**	**5.79**	**3.83**	**7.50**	**99.9**
	CNN (1 Step Ahead)	15-min	9.82	6.51	11.29	99.8
	LSTM (1 Step Ahead)	15-min	10.43	6.61	10.19	99.8

Conclusion

This study proposes three different DL-based forecasting models (LSTM, CNN, and CNN –LSTM) capable of forecasting SI up to six steps ahead on a 15-min interval basis. For performing experiments average 15-min interval datasets are categorized into training and validation datasets by utilizing the input variables namely, SI (W/m^2), T (^{0}C), and RH (%) with 12 steps lagged sequence by incorporating sliding window technique whereas the outputs are transformed by first-order differencing technique to convert the non-stationary characteristics of time-series data into stationary characteristics. For the performances of the models, the actual versus forecasted SI plots along with the forecasted error plots for 6 days as well as 1 day show the prediction capability of the network while the linear regression plots show the correlation between actual and forecasted values. With the CI, we can illustrate the probable range of the mean prediction.

The overall study concludes that the developed deep hybrid model (CNN-LSTM) performed better than the standalone models (deep LSTM and deep CNN) in forecasting the SI up to six steps. While moving from 1-step to 6-steps ahead, the error indicators of all 3 deep models are showing increasing nature. The range of MAE and RMSE is found to be 6.61–36.89 W/m^2 and 10.43–62.27 W/m^2 respectively for deep LSTM, 6.51–32.38 W/m^2 and 9.82–50.99 W/m^2 respectively for deep CNN, and 3.83–20.07 W/m^2 and 5.79–34.89 W/m^2 respectively for deep CNN-LSTM. We also found that during higher time steps some model (LSTM and CNN) shows higher and fluctuating nature of MAPE, whereas CNN-LSTM model shows stable and more accurate performance during entire time steps. Based on comparing results with the existing works of literature it is also found that the outputs from the LSTM and CNN model are acceptable up to the 5 steps while comparing results with R^2, whereas for CNN-LSTM it is acceptable up to 6 steps.

Conclusively, the proposed hybrid CNN-LSTM model is highly stable which exhibits excellent performance in very short-term forecasting of SI. The comparative analysis of DL models presented in this work can enrich the future of researchers and engineers to select the DL model in various other domains such as wind speed forecasting, load forecasting, etc. Further, the prediction accuracy achieved with a hybrid deep model for very short-term forecasting of SI can be more effective for real-time microgrid energy management system where load fluctuations are more especially in shorter intervals.

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

[1] Guo Z, Zhou K, Zhang C, Lu X, Chen W, Yang S. Residential electricity consumption behavior: Influencing factors, related theories and intervention strategies. Renew Sustain Energy Rev 2018;81:399–412. https://doi.org/10.1016/J.RSER.2017.07.046.

[2] Kumari P, Toshniwal D. "Impact of lockdown measures during COVID-19 on air quality– A case study of India," *https://doi.org/10.1080/09603123.2020.1778646*, 2020, 10.1080/09603123.2020.1778646.

[3] Wang H, et al. Deep learning-based interval state estimation of ac smart grids against sparse cyber attacks. IEEE Trans Ind Informatics 2018;14(11):4766–78. https://doi.org/10.1109/TII.2018.2804669.

[4] Kumari P, Toshniwal D. Deep learning models for solar irradiance forecasting: A comprehensive review. J Clean Prod 2021;318:128566. https://doi.org/10.1016/J.JCLEPRO.2021.128566.

[5] I. Renewable Energy Agency, "Renewable capacity highlights (31 March 2020)," 2020. Accessed: Apr. 24, 2021. [Online]. Available: www.irena.org/publications.

[6] Yang XJ, Hu H, Tan T, Li J. "China's renewable energy goals by 2050," *Environmental Development*, vol. 20. Elsevier B.V., pp. 83–90, Nov. 01, 2016, 10.1016/j.envdev.2016.10.001.

[7] Ma T, Yang H, Lu L, Peng J. Optimal design of an autonomous solar–wind-pumped storage power supply system. Appl Energy 2015;160:728–36. https://doi.org/10.1016/J.APENERGY.2014.11.026.

[8] Frías-Paredes L, Mallor F, Gastón-Romeo M, León T. Assessing energy forecasting inaccuracy by simultaneously considering temporal and absolute errors. Energy Convers Manag 2017;142:533–46. https://doi.org/10.1016/J.ENCONMAN.2017.03.056.

[9] Sobri S, Koohi-Kamali S, Rahim NA. Solar photovoltaic generation forecasting methods: A review. Energy Convers Manag 2018;156:459–97. https://doi.org/10.1016/J.ENCONMAN.2017.11.019.

[10] Hong YY, Martinez JJF, Fajardo AC. Day-Ahead Solar Irradiation Forecasting Utilizing Gramian Angular Field and Convolutional Long Short-Term Memory. IEEE Access 2020;8:18741–53. https://doi.org/10.1109/ACCESS.2020.2967900.

[11] Kumari P, Toshniwal D. Long short term memory–convolutional neural network based deep hybrid approach for solar irradiance forecasting. Appl Energy 2021;295:117061. https://doi.org/10.1016/J.APENERGY.2021.117061.

[12] Kumari P, Toshniwal D. Extreme gradient boosting and deep neural network based ensemble learning approach to forecast hourly solar irradiance. J Clean Prod 2021;279:123285. https://doi.org/10.1016/J.JCLEPRO.2020.123285.

[13] U. K. Das *et al.*, "Forecasting of photovoltaic power generation and model optimization: A review," *Renewable and Sustainable Energy Reviews*, vol. 81. Elsevier Ltd, pp. 912–928, Jan. 01, 2018, 10.1016/j.rser.2017.08.017.

[14] C. Voyant *et al.*, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105. Elsevier Ltd, pp. 569–582, May 01, 2017, 10.1016/j.renene.2016.12.095.

[15] Yang D, Kleissl J, Gueymard CA, Pedro HTC, Coimbra CFM. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. Sol Energy Jul. 2018;168:60–101. https://doi.org/10.1016/j.solener.2017.11.023.

[16] Perez R, Kivalov S, Schlemmer J, Hemker K, Renné D, Hoff TE. Validation of short and medium term operational solar radiation forecasts in the US. Sol Energy 2010;84(12):2161–72. https://doi.org/10.1016/j.solener.2010.08.014.

[17] Ramsami P, Oree V. A hybrid method for forecasting the energy output of photovoltaic systems. Energy Convers Manag 2015;95:406–13. https://doi.org/10.1016/j.enconman.2015.02.052.

[18] Bouzgou H, Gueymard CA. Fast short-term global solar irradiance forecasting with wrapper mutual information. Renew Energy 2019;133:1055–65. https://doi.org/10.1016/j.renene.2018.10.096.

[19] Yang D, Jirutitijaroen P, Walsh WM. Hourly solar irradiance time series forecasting using cloud cover index. Sol Energy 2012;86(12):3531–43. https://doi.org/10.1016/j.solener.2012.07.029.

[20] Diagne M, David M, Lauret P, Boland J, Schmutz N, "Review of solar irradiance forecasting methods and a proposition for small-scale insular grids," *Renewable and Sustainable Energy Reviews*, vol. 27. Pergamon, pp. 65–76, Nov. 01, 2013, 10.1016/j.rser.2013.06.042.

[21] Dong Z, Yang D, Reindl T, Walsh WM. Short-term solar irradiance forecasting using exponential smoothing state space model. Energy 2013;55:1104–13. https://doi.org/10.1016/J.ENERGY.2013.04.027.

[22] Turrado C, López M, Lasheras F, Gómez B, Rollé J, Juez F. Missing data imputation of solar radiation data under different atmospheric conditions. Sensors 2014;14(11):20382–99. https://doi.org/10.3390/s141120382.

[23] Huang X, Shi J, Gao B, Tai Y, Chen Z, Zhang J. Forecasting hourly solar irradiance using hybrid wavelet transformation and elman model in smart grid. IEEE Access 2019;7:139909–23. https://doi.org/10.1109/ACCESS.2019.2943886.

[24] Sharma V, Yang D, Walsh W, Reindl T. Short term solar irradiance forecasting using a mixed wavelet neural network. Renew Energy 2016;90:481–92. https://doi.org/10.1016/J.RENENE.2016.01.020.
[25] Chen SX, Gooi HB, Wang MQ. Solar radiation forecast based on fuzzy logic and neural networks. Renew Energy 2013;60:195–201. https://doi.org/10.1016/j.renene.2013.05.011.
[26] Aguiar LM, Pereira B, Lauret P, Díaz F, David M. Combining solar irradiance measurements, satellite-derived data and a numerical weather prediction model to improve intra-day solar forecasting. Renew Energy 2016;97:599–610. https://doi.org/10.1016/j.renene.2016.06.018.
[27] Baser F, Demirhan H. A fuzzy regression with support vector machine approach to the estimation of horizontal global solar radiation. Energy 2017;123:229–40. https://doi.org/10.1016/j.energy.2017.02.008.
[28] Nagy GI, Barta G, Kazi S, Borbély G, Simon G. GEFCom2014: Probabilistic solar and wind power forecasting using a generalized additive tree ensemble approach. Int J Forecast 2016;32(3):1087–93. https://doi.org/10.1016/j.ijforecast.2015.11.013.
[29] Alfadda A, Rahman S, Pipattanasomporn M. Solar irradiance forecast using aerosols measurements: A data driven approach. Sol Energy Aug. 2018;170:924–39. https://doi.org/10.1016/J.SOLENER.2018.05.089.
[30] Khosravi A, Koury RNN, Machado L, Pabon JJG. Prediction of hourly solar radiation in Abu Musa Island using machine learning algorithms. J Clean Prod 2018;176:63–75. https://doi.org/10.1016/J.JCLEPRO.2017.12.065.
[31] G. G. Bizzarri F, Bongiorno M, Brambilla A, Gruosso G, "Model of photovoltaic power plants for performance analysis and production forecast," *IEEE Trans Sustain Energy*, vol. 4, pp. 278–85, 2013.
[32] Lauret P, Voyant C, Soubdhan T, David M, Poggi P. A benchmarking of machine learning techniques for solar radiation forecasting in an insular context. Sol Energy 2015;112:446–57. https://doi.org/10.1016/j.solener.2014.12.014.
[33] A. Mellit, A. Massi Pavan, and V. Lughi, "Short-term forecasting of power production in a large-scale photovoltaic plant," *Sol. Energy*, vol. 105, pp. 401–413, Jul. 2014, 10.1016/j.solener.2014.03.018.
[34] Khodayar M, Kaynak O, Khodayar ME. Rough deep neural architecture for short-term wind speed forecasting. IEEE Trans Ind Informatics 2017;13(6):2770–9. https://doi.org/10.1109/TII.2017.2730846.
[35] Khodayar M, Wang J. Spatio-Temporal Graph Deep Neural Network for Short-Term Wind Speed Forecasting. IEEE Trans. Sustain. Energy 2019;10(2):670–81. https://doi.org/10.1109/TSTE.2018.2844102.
[36] Alzahrani A, Shamsi P, Dagli C, Ferdowsi M. Solar Irradiance Forecasting Using Deep Neural Networks. Procedia Comput Sci 2017;114:304–13. https://doi.org/10.1016/j.procs.2017.09.045.
[37] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM Jun. 2017;60(6):84–90. https://doi.org/10.1145/3065386.
[38] Hinton G, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. IEEE Signal Process Mag 2012;29(6):82–97. https://doi.org/10.1109/MSP.2012.2205597.
[39] Helmstaedter M, Briggman KL, Turaga SC, Jain V, Seung HS, Denk W. Connectomic reconstruction of the inner plexiform layer in the mouse retina. Nature Aug. 2013;500(7461):168–74. https://doi.org/10.1038/nature12346.
[40] Qing X, Niu Y. Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy 2018;148:461–8. https://doi.org/10.1016/J.ENERGY.2018.01.177.
[41] Srivastava S, Lessmann S. A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data. Sol Energy 2018;162:232–47. https://doi.org/10.1016/J.SOLENER.2018.01.005.
[42] Yu Y, Cao J, Zhu J. An LSTM Short-Term Solar Irradiance Forecasting under Complicated Weather Conditions. IEEE Access 2019;7:145651–66. https://doi.org/10.1109/ACCESS.2019.2946057.
[43] Hong YY, Rioflorido CLPP. A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. Appl Energy 2019;250:530–9. https://doi.org/10.1016/J.APENERGY.2019.05.044.
[44] Ye L, Cao Z, Xiao Y. DeepCloud: ground-based cloud image categorization using deep convolutional features. IEEE Trans Geosci Remote Sens 2017;55(10):5729–40. https://doi.org/10.1109/TGRS.2017.2712809.
[45] Zang H, Liu L, Sun L, Cheng L, Wei Z, Sun G. Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. Renew. Energy 2020;160:26–41. https://doi.org/10.1016/j.renene.2020.05.150.
[46] Mellit A, Pavan AM, Lughi V. Deep learning neural networks for short-term photovoltaic power forecasting. Renew. Energy 2021;172:276–88. https://doi.org/10.1016/j.renene.2021.02.166.
[47] Huang X, et al. Hybrid deep neural model for hourly solar irradiance forecasting. Renew. Energy 2021;171:1041–60. https://doi.org/10.1016/j.renene.2021.02.161.
[48] Gao B, Huang X, Shi J, Tai Y, Zhang J. Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks. Renew. Energy 2020;162:1665–83. https://doi.org/10.1016/j.renene.2020.09.141.
[49] Ghimire S, Deo RC, Raj N, Mi J. Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. Appl Energy 2019;253:113541. https://doi.org/10.1016/j.apenergy.2019.113541.
[50] Ray B, Shah R, Islam MR, Islam S. A new data driven long-term solar yield analysis model of photovoltaic power plants. IEEE Access 2020;8:136223–33. https://doi.org/10.1109/ACCESS.2020.3011982.
[51] Liu H, Mi X, Li Y. Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network. Energy Convers. Manag. 2018;166:120–31. https://doi.org/10.1016/J.ENCONMAN.2018.04.021.
[52] Qin Y, et al. Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. Appl Energy 2019;236:262–72. https://doi.org/10.1016/J.APENERGY.2018.11.063.
[53] Kim TY, Cho SB. Predicting residential energy consumption using CNN-LSTM neural networks. Energy 2019;182:72–81. https://doi.org/10.1016/J.ENERGY.2019.05.230.
[54] Pang Z, Niu F, O'Neill Z. Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons. Renew. Energy 2020;156:279–89. https://doi.org/10.1016/j.renene.2020.04.042.
[55] Boubaker S, Benghanem M, Mellit A, Lefza A, Kahouli A, Kolsi L. Deep Neural Networks for Predicting Solar Radiation at Hail Region, Saudi Arabia. IEEE Access 2021. https://doi.org/10.1109/ACCESS.2021.3062205.
[56] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, May 27, 2015, 10.1038/nature14539.
[57] Jason B. Deep Learning for Time Series Forecasting: Predict the Future with MLPs. CNNs and LSTMs in Python. 2020.
[58] "Confidence Intervals for Machine Learning." https://machinelearningmastery.com/confidence-intervals-for-machine-learning/ (accessed Apr. 26, 2021).