

Received February 7, 2021, accepted February 22, 2021, date of publication February 24, 2021, date of current version March 9, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3062205

Deep Neural Networks for Predicting Solar Radiation at Hail Region, Saudi Arabia

SAHBI BOUBAKER^{1,2}, MOHAMED BENGHANEM³, ADEL MELLIT^{4,5}, AYOUB LEFZA⁴,
OMAR KAHOULI¹, AND LIOUA KOLSI⁶

¹Department of Electronics Engineering, Community College, University of Ha'il, Ha'il 2240, Saudi Arabia

²Department of Computer and Network Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah 23218, Saudi Arabia

³Physics Department, Faculty of Science, Islamic University in Madinah, Medina 42351, Saudi Arabia

⁴Renewable Energy Laboratory, University of Jijel, Jijel 18000, Algeria

⁵International Centre of Theoretical Physics (ICTP), 34151 Trieste, Italy

⁶Mechanical Engineering Department, College of Engineering, University of Ha'il, Ha'il 2240, Saudi Arabia

Corresponding author: Adel Mellit (amellit@ictp.it)

This work was supported by the Deputy of Research & Innovation, Ministry of Education in Saudi Arabia, under project RDO-2004.

ABSTRACT Forecasted global horizontal irradiation (GHI) can help for designing, sizing and performances analysis of photovoltaic (PV) systems including water PV pumping systems used for irrigation applications. In this paper, various deep neural networks (DNN) models for one day-ahead prediction of GHI at Hail city (Saudi Arabia) are developed and investigated. The considered DNN models include long-short-term memory (LSTM), bidirectional-LSTM (BiLSTM), gated recurrent unit (GRU), bidirectional-GRU (Bi-GRU), one-dimensional convolutional neural network (CNN_{1D}) and other hybrid configurations such as CNN-LSTM and CNN-BiLSTM. A dataset of daily GHI recordings collected during January 1, 2000 to June 30, 2020 from National Aeronautics and Space Administration (NASA) at an arid location (Hail, Saudi Arabia) is used to develop and compare the above DNN-based models. The parameters affecting the accuracy of the models have been also deeply analyzed. Only historical values of daily GHI have been used to build the DNN-based models whereas additional weather parameters such as air temperature, wind speed, wind direction, atmospheric pressure and relative humidity are not considered in this work. Keras library and Python language have been used to develop and compare the GHI forecasting models. The evaluation metrics such as correlation coefficient (r), Mean Absolute Percent Error (MAPE), Mean Absolute Error (MAE), cumulative distribution function (CDF) and standard deviation (σ) are opted to evaluate the performance of the prediction models. The obtained results showed that the DNN models have provided globally good performances with a maximum reached value of $r = 96\%$, for daily GHI forecasting.

INDEX TERMS Global horizontal irradiation, prediction, deep learning, recurrent neural networks, LSTM, GRU, BiLSTM, BiGRU, CNN, CNN-LSTM, CNN-BiLSTM.


I. INTRODUCTION

Accurate forecasting of solar irradiation including global, direct, diffuse and normal is a crucial task for designing optimal solar energy systems (photovoltaic and thermal) [1].

During the last three decades, numerous studies used machine learning (ML) techniques such as artificial neural networks (ANNs) and support vector machine (SVM) for forecasting global solar irradiation (GSR) [2]. The classical methods including stochastic approaches and shallow ANNs, recurrent neural networks (RNNs) have limited capability in

capturing long-term dependency and the scalability issues with large datasets [3] (e.g. the vanishing gradients problem in RNNs). Recently, with the availability of a huge amount of collected data across the globe and advances in computing technologies, researchers working in this area are more and more attracted towards deep learning (DL) techniques for developing prediction models. The DL belong to broader family of ML algorithms based on ANNs with representation learning [4].

The DL based prediction models are playing an important role in numerous areas such as optimal scheduling of merchant-owned energy storage systems [5] and optimization of networked distributed energy resources [6]. For instance,

The associate editor coordinating the review of this manuscript and approving it for publication was Orazio Gambino .

most works related to solar irradiation prediction are mainly using offline models including ML and deep neural networks (DNNs).

Supervised learning is the most frequently used in solar irradiation forecasting. Reinforcement learning could also contribute to improve the forecasting accuracy in this area since these learning techniques have proven their capability in wind power forecast uncertainty [7] and frequency forecasting in distrusted systems [8].

Cyber-attacks are other issues having an important impact on the power systems operation [9] including distributed generation [10], photovoltaic (PV) systems and wind-turbine systems. Generally, in solar irradiation applications, cyber-attacks should be considered in order to secure the stored data on which the operation of the networks is based.

Currently, DNNs including long short-term memory (LSTM), gated recurrent unit (GRU), one-dimension convolutional neural networks (CNN_{1D}) and other hybrid configurations such as CNN_{1D} -LSTM are considered to be the powerful tools in time series forecasting [11]. For instance, a powerful deep learning NN, named Deep-Energy was introduced for energy load forecasting [12]. In [13], the authors developed a new model named 'SEPNet' for the prediction of hourly electricity prices. Dividing input data into different seasons has contributed significantly to improve the prediction accuracy. However, to the best of our knowledge, the application of DL in solar irradiation forecasting is found to be very limited compared to other fields of research. In fact, few attempts have been carried out over the last couple of years. For example, an LSTM is implemented to predict short-term hourly solar irradiance [14]. The results revealed that LSTM outperformed the well-known feed-forward neural network (FFNN) and SVM. The method has been reported to be easily implemented using Keras deep learning library [15].

Aslam *et al.* [16] confirmed the superiority of LSTM and GRU algorithms in forecasting of GHI in two regions compared with conventional ANNs (FFNN, RNN and SVM). However, additional inputs are required (e.g. Clear-Sky) to implement the model and improve its accuracy. A hybrid model (C-LSTM) based on CNN and LSTM was proposed in [17] for half-hour GHI forecasting. The proposed model used CNN to extract GHI data features and LSTM to encapsulate the features to generate a low latency-based time series GHI prediction model. Results showed that the GHI prediction model outperform the other examined algorithms including CNN, LSTM, DNN, decision tree (DT) and multi-layer perceptron (MLP). Despite the good results, this hybrid configuration (CLSTM) is relatively complex compared to other DNNs (e.g. LSTM and GRU).

Sequence to sequence (S2S) DL models were used to forecast 24-h ahead of GHI. The results revealed superiority of LSTM over FFNN and gradient boosted regression trees algorithms [18]. The method can generally be implemented easily as it requires only historical GHI data. An LSTM

algorithm is developed for multi-step ahead forecasting of GHI (36-h and 24-h) [19]. The models yielded acceptable results in the investigated regions. However, additional meteorological data such as air temperature (T), wind speed (W_S), dew-point (D_p), relative humidity (R_H) and pressure (P_R) are needed in addition to the solar radiation (SR) components (direct DHI and normal DNI). Therefore, the method became relatively complex. Qing and Niu [20] used weather forecast and an LSTM to predict hourly-day ahead of GHI. Although, the model used is simple in implementation, its accuracy is reported to depend mainly on the provided forecast data from weather service organizations (WSO). Moreover, data from WSO are not always accessible at different regions.

An LSTM was implemented to forecast one-step ahead (hour and day in advance) of GHI for complicated weather conditions [21] at three regions in USA: Atlanta, New York, and Hawaii. Results indicated that for cloudy days the LSTM performs better than conventional RNN. The clearness index (K_t) is used to classify days. Husein and Chung [22] developed a method based on LSTM for day-ahead forecasting of GHI in four locations: Germany, U.S.A, Switzerland, and South Korea. It was reported that the LSTM outperforms FFNN of all investigated locations. The method is simple and only historical values of GHI are needed.

A reliable approach for solar irradiance forecasting based on Choquet Integral and LSTM was developed in [23]. Six datasets collected from different regions in Finland were used to assess the accuracy of the approach. Results indicate that the proposed approach showed lower forecasting errors than the other methods, such as autoregressive integrated moving average (ARIMA). The approach is relatively complex compared to other structures.

De Araujo [24] made a comparison between LSTM algorithm and weather research and forecasting (WRF) model at Gifu, Japan for radiation forecasting. Results demonstrated that WRF model performed better with a lower prediction error compared to the LSTM algorithm. The designed model is simple and requires only GHI data to be implemented. A Gramian Angular Field (time series used in image transformation) and CovLSTM were developed [25] for one day ahead forecasting of solar radiation, at Taiwan. The method outperforms ARIMA, LSTM-FC, and CNN-LSTM. Nevertheless, the method is found to be relatively complex, since raw data are converted into images by using the Gramian Angular Field. GRU algorithm and weather forecasting data are used to predict day-ahead of GHI [26]. The results showed that the GRU performs well compared to other investigated models. However, the accuracy depends mainly on the accuracy of the weather forecast variables.

The Prediction of accuracy of GHI depends on several factors including the geographic location, climate and also on the methods used as well as on the size and quality of the available data.

Most papers claimed that LSTM, GRU, CNN and Cov-CNN algorithms perform better, in terms of accuracy, than other classical ML algorithms including conventional

RNN, FFNN, SVM, and stochastic approaches such as autoregressive (AR), AR moving average (ARMA) and ARIMA. Thus, in this study we focus mainly on the development of DNN algorithms including LSTM, BiLSTM, CNN, GRU, BiGRU and a hybrid CNN-LSTM. This work is not intended to compare the results with conventional approaches. The novelties and contributions of this paper two fold.

Firstly, development and evaluation of different DNN structures (LSTM, BiLSTM, GRU, BiGRU, CNN, CNN-LSTM, and CNN-BiLSTM) for one-day-ahead forecasting of GHI at Hail region, Saudi Arabia, which is the first application in this region. Some DL algorithms were not previously investigated with daily solar irradiation forecasting (e.g. Bi-GRU, CNN-BiLSTM). Secondly, deep analysis of the developed DNN-based models by varying different parameters that can affect the forecasting accuracy (e.g. number of inputs, number of units within the hidden layer, number of layers, batch size, epoch and size of dataset). This issue has not been investigated in the previous studies.

To assess the performances of the DNN-based models, different error metrics and statistical tests (r , RMSE, MAPE, MAE, CDF and σ) have been used [27].

The paper is organized as follows: Section II presents the site location and the dataset used for developing the DNN forecasters. A brief introduction on deep learning, including the used deep neural networks (LSTM, BiLSTM, GRU, Bi-GRU, CNN, CNN-LSTM and CNN-BiLSTM), is given in Section III. Development of forecasting models is also reported in the same Section. Results and discussion are provided in Section IV.

II. SITE LOCATION AND DATASET

A. SITE LOCATION

The considered region (Hail, Saudi Arabia) is located at (27°23'11", 41°38'49") geographic coordinates. Hail has a hot desert climate (Köppen climate classification) with hot summers and cool winters. It has a somewhat milder climate than other Saudi cities due to its higher altitude, temperature, humidity and insolation [28]. Moreover, Hail is known to be among the most active region in agriculture for which using solar energy for pumping water from deep wells can be economically efficient. Note here that this study is the first step in a research project aiming to implement a pilot photovoltaic water pumping system (PVWPS) at Hail region.

B. DATASET REPRESENTATION

The dataset used for DNN-based models' development is collected from the National Aeronautics and Space Administration (NASA) [29] during January 1, 2000 to June 30, 2020 (20 years and 6 months). Figure 1 depicts the distribution of monthly GHI for the year 2019. Seasonal behavior as well as a remarkable variance of GHI distribution are observed. These patterns have been found to be the same for the other years.

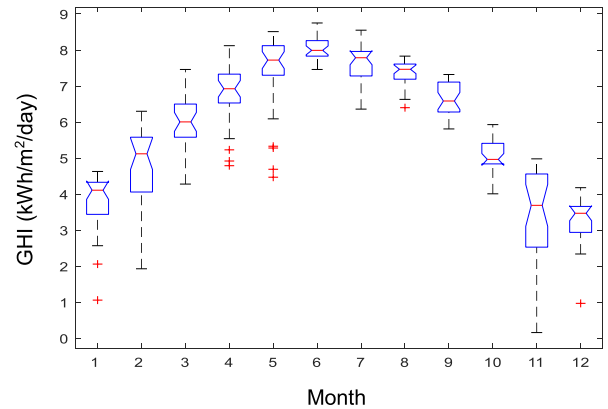


FIGURE 1. Distribution of daily GHI for each month of the year 2019 at Hail location, Saudi Arabia.

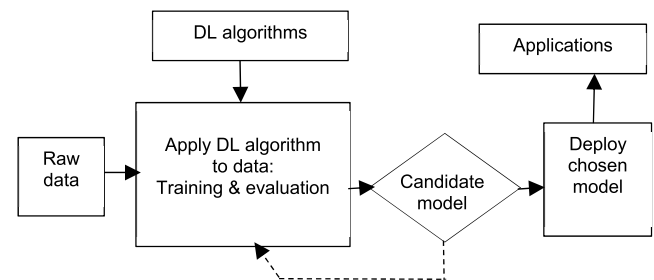


FIGURE 2. Deep learning process for developing prediction model.

III. METHODOLOGY

A. DEEP LEARNING NEURAL NETWORKS

DL is a sub-field of ML where ANNs make up the backbone of DL algorithms. Figure 2 shows a simplified process of a DL algorithm implementation. As can be seen, the feature extraction step is intentionally removed because DL algorithms are able to learn and automatically extract features from raw input data [4].

DL is a relatively new advancement in NNs and represents a way to train DNNs since traditional NN-based methods might be affected by problems such as over-fitting and diminishing gradients [30]. A short description of the commonly used DNNs for GHI forecasting is detailed in the following sub-sections.

1) LONG SHORT-TERM MEMORY

LSTM is a kind of RNN with some modification including cell, input gate, output gate and forget gate [31]. An LSTM layer can learn long-term dependencies. It is mainly used for time series prediction. A simple configuration consists of some LSTM cells and dense output layer. Figure 3 shows the basic LSTM cell. It is composed of three main gates: input gate, forget gate and output gate. The following equations show the relationship between the three gates, input and output [31].

$$f_t = \alpha(w_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

$$i_t = \alpha(w_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

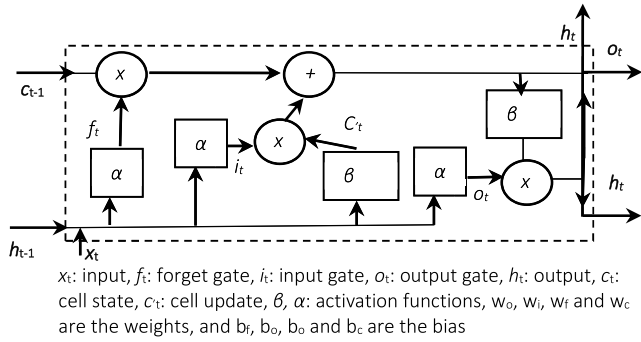


FIGURE 3. The basic structure of an LSTM cell.

$$o_t = \alpha(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{3}$$

$$C_t = f_t C_{t-1} + i_t \cdot \text{th}(w_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

$$h_t = O_t \cdot \beta(C_t) \tag{5}$$

2) BIDIRECTIONAL-LSTM

Bi-LSTM is a modified version of LSTM [32]. It consists of two separate hidden layers; firstly, it computes the forward hidden sequence, then it calculates the backward hidden sequence, and finally, it combines both to calculate the output.

3) GATED RECURRENT UNIT

The GRU is similar to LSTM with fewer parameters than LSTM [33]. The parameters are learned through the gating mechanism of GRU. GRU is computationally more efficient considering fewer parameters and need less data to generalize. It can also learn long-term dependencies.

4) BIDIRECTIONAL-GRU

Bi-GRU is an improved version of GRU [32] having the same structure as BiLSTM with fewer parameters. It consists basically of two separate hidden layers. The process is relatively similar to BiLSTM.

5) CONVOLUTIONAL NEURAL NETWORK

CNN is a regularized version of the well-known Feed-forward NNs. CNN was firstly developed for 2D problems. It consists of some layers, Conv2D, Max Pooling, Flatten and Fully connected layer [30]. It could be also used for solving one-dimensional problems (CNN_{1D}) such as time series classification and prediction.

6) CNN-LSTM

CNN-LSTM combines CNN with LSTM. Its principle consists of putting in cascade both configurations to get the hybrid configuration (CNN-LSTM) [34], [35]. A simplified schematic illustration of a one-dimensional CNN-LSTM is shown in figure 4.

The CNN-LSTM includes one convolutional layer, one max pooling layer, a flatten layer, an LSTM layer and a fully dense layer (fully connected layer with one output). Combining CNN with LSTM means that CNN (1-Dimensional) is used to reduce the features of raw data, $z_t = \text{CNN}_{1D}(x_t)$

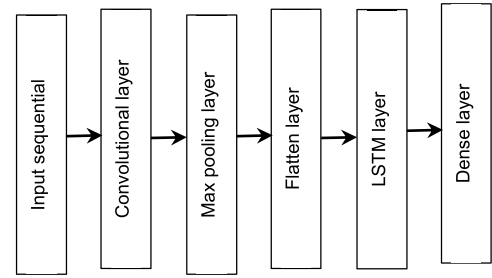


FIGURE 4. One-dimensional CNN-LSTM structure used for daily GHI forecasting.

using conventional filters. Where x_t are the raw data, z_t is the output sequence obtained by the calculation of the convolution matrix in the first layer, max-pooling in the second layer which permits to reduce the matrix size, and flatten layer to convert matrix to vector (flattening operation). Therefore, this output (z_t) is then used to feed the LSTM layer. An example of the developed CNN-LSTM using Python and Keras library [15] is given in the Appendix (Python code of the developed CNN-LSTM).

7) CNN-BI-LSTM

CNN-Bi-LSTM has the same structure as CNN-LSTM except that the LSTM layer (See figure 4) is replaced by a Bi-LSTM.

B. FORECASTING OF DAILY GLOBAL SOLAR IRRADIATION APPROACHES

Broadly forecasting of GHI based artificial intelligence (AI) techniques including ML and DL can be achieved by three ways [36], [37]. The first one uses only historical GHI records (ground measurements):

$$(y_{t+1}, \dots, y_{t+k}) = f(y_{t-n}, y_{t-n+1}, \dots, y_t) \tag{6}$$

where variable y_t is the actual GHI, y_{t-n} is the previous value of GHI, y_{t+k} is the forecasted value of GHI at step k , and f is a functional dependency between past and future samples. $t \in t \in \{1, \dots, n\}$, n is the size of the sequence.

The second one relies upon forecasted meteorological parameters by numerical weather prediction models (NWP) such as T, R_H, W_S, P_R and cloud index (C_I).

$$(y_{t+1}, \dots, y_{t+k}) = f(x_{0t}, x_{1t}, x_{2t}, x_{3t}, x_{4t}, \dots) \tag{7}$$

where variables $x_{0t}, x_{1t}, x_{2t}, x_{3t}$ and x_{4t} correspond to T, R_H, W_S, P_R and C_I respectively.

The third one combines the use of historical GHI data records with meteorological parameters:

$$(y_{t+1}, \dots, y_{t+k}) = f(y_{t+n}, y_{t+n-1}, \dots, x_{0t}, x_{1t}, x_{2t}, x_{3t}, x_{4t}, \dots) \tag{8}$$

These approaches could be used for one-step ($k = 1$) and multistep ($k > 1$) ahead forecasting.

In this paper, the first approach is considered, because other meteorological variables recorded in-situ or forecasted based on numerical weather prediction models are not always

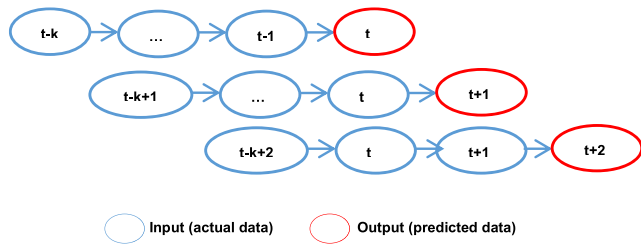


FIGURE 5. Example of input (actual) and output (predicted) sequences.

available in many locations including our case study (Hail city, Saudi Arabia). In addition, we can cite the simplicity of the forecasting models based only on historical values of GHI. Thus, the described DNNs are used to forecast the next value of daily GHI based on some past measurements of GHI. In our case $k = 1$, so,

$$(y_{t+1}, \dots, y_{t+k}) = f(y_{t-n}, y_{t-n+1}, \dots, y_t) \quad (9)$$

Figure 5 shows an example of the adopted process for solar radiation prediction [35].

The dataset is divided into three parts: a subset of 70% is used for training the different DNNs, 15% for validation and remaining 15% for testing the models. The dataset is firstly normalized using the following MinMax equation [37]:

$$y_N = (y - y_{min}) (y_{max} - y_{min})^{-1} \quad (10)$$

where y_N is the normalized value of GHI, y_{min} and y_{max} are minimum and maximum values of GHI, respectively. An example of inputs and output samples is shown in the appendix (*Case of 7 inputs and one output*). Figure 6 shows the flowchart of the forecasting method based on DNN.

To evaluate the performance of the investigated DNNs, four performance measures have been selected: the r , the RMSE and the MAPE in addition to other statistical tests such as CDF and σ . To prevent the overfitting problem, a dropout layer is added in addition to the decision of stopping training process early when the convergence is reached. As consequence, it is not needed to train the network for all epochs. The investigated models have been developed using the Adam optimizer [38]. The updated equations are also shown in the Appendix (*Adam optimization*).

IV. RESULTS AND DISCUSSION

Different DNN models have been evaluated for the prediction of one-day-ahead of GHI in Hail region, Saudi Arabia. The experiments have been conducted on a laptop i7-4600U CPU @2.10GHz, 2 Cores(s), 4 Logical Processor(s), 8 GB of DDR3 RAM. The codes have been written using Python programming language under Ubuntu operating system.

To show the performance of the studied DNN-based models, CDF is firstly plotted in figure 7. As can be seen, the majority of models have good correlation between measured and forecasted daily GHI. However, in the case of GHI lower than 4 kWh/m²/day, the models did not perform well (small variation is observed) as can be observed in the zoom plot of figure 7.

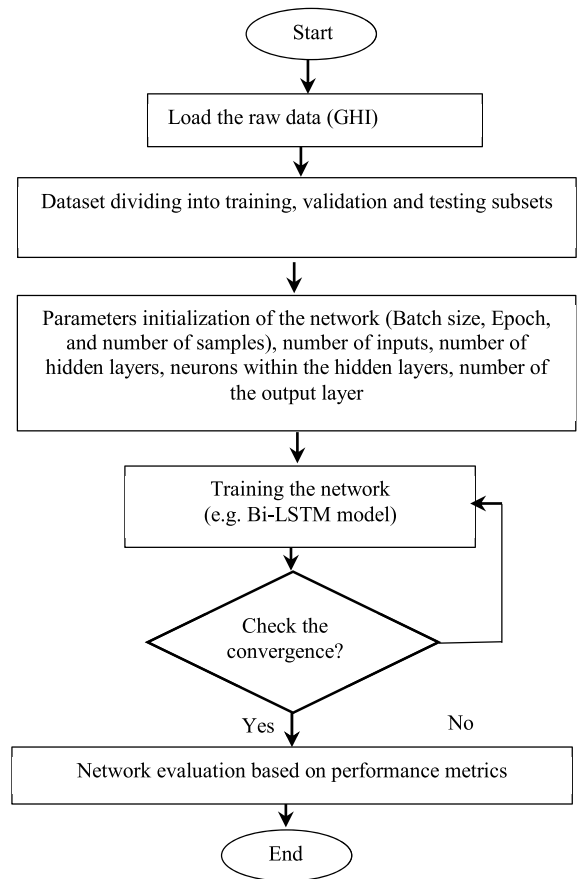


FIGURE 6. Flowchart of the forecasting method based DNNs.

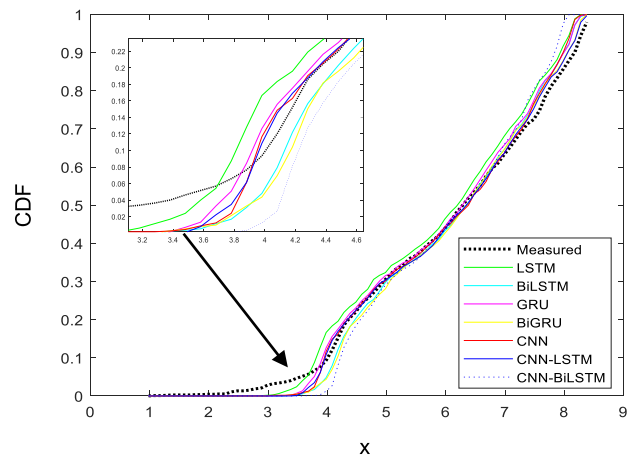


FIGURE 7. Variation of the cumulative distributed function (CDF) of the investigated DNN based-models: LSTM, BiLSTM, GRU, Bi-GRU, CNN, CNN-LSTM and CNN-BiLSTM.

To check the performance variation of the developed Bi-LSTM, the μ and σ variation between measured and forecasted daily GHI during the test period (1 July 2017 to 30 June 2020) are depicted in figure 8.

As can be seen from figure 8 (shading plot), there is no big variation of the standard deviation (gray color) from the μ values which indicates the good accuracy of the Bi-LSTM forecaster.

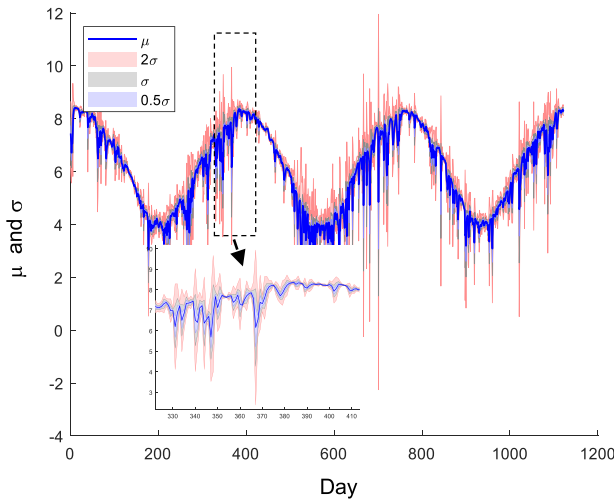


FIGURE 8. Variation of the μ and σ values in the case of Bi-LSTM: σ is the standard deviation and μ is the mean between measured and forecasted daily GHI (1 July 2017 to 30 June 2020).

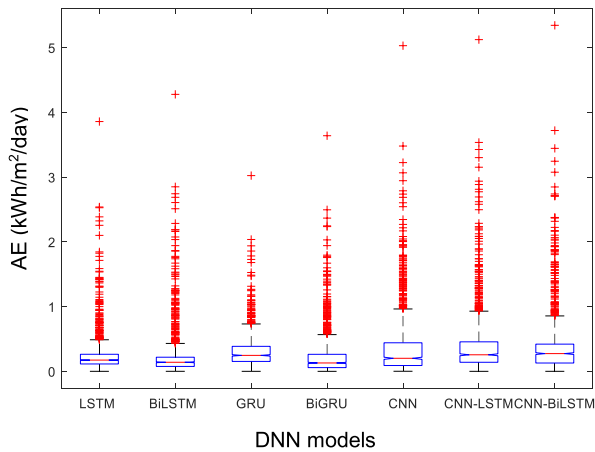


FIGURE 9. The absolute error (AE) of different investigated DNN models (red color indicates outliers which are greater than the maximum value. The center of the box is the medium value and the blue box is the variance).

The boxplot in figure 9 displays the calculated absolute error (AE) for the investigated DNN models. Training accuracy is relatively the same with the test accuracy, which indicates that there is no overfitting problem in our models. We can observe from figure 9 that LSTM, GRU, Bi-LSTM and Bi-GRU models present lower variance values. Nevertheless, for all models, outliers are similar (red color). CNN, CNN-LSTM and CNN-BiLSTM have relatively larger variance values.

After several experiments, the optimal structures of the investigated DNN models are listed in Table 1 including the calculated error metrics (r , RMSE, and MAPE). The batch size (BS = 64), the number of inputs (NI = 7), epoch = 100, number of units in the hidden layer (HL) of mostly investigated models is NU = 100.

Table 1 shows that the r ranges between 92% and 96%. For all investigated models the RMSE do not exceed 1 kWh/m²/day. The lowest MAPE is obtained by

TABLE 1. Error metrics between measured and forecasted Daily GHI for the investigated DNN-Based models (LSTM, Bi-LSTM, GRU, Bi-GRU, CNN, CNN-LSTM and CNN-Bi-LSTM).

Model: Structure	RMSE (kWh/m ² /day)	MAPE (%)	r (%)
LSTM:(7x100x1)	0.55	6.34	96.0
Bi-LSTM:(7x100x1)	0.62	5.24	95.9
GRU:(7x50x1)	0.60	6.76	95.5
Bi-GRU:(7x50x1)	0.54	6.61	95.7
CNN:(7x100x1)	0.75	8.93	92.1
CNN-LSTM:(7x64x100x1)	0.93	10.45	93.8
CNN-Bi-LSTM:(7x64x100x1)	0.97	11.34	93.5

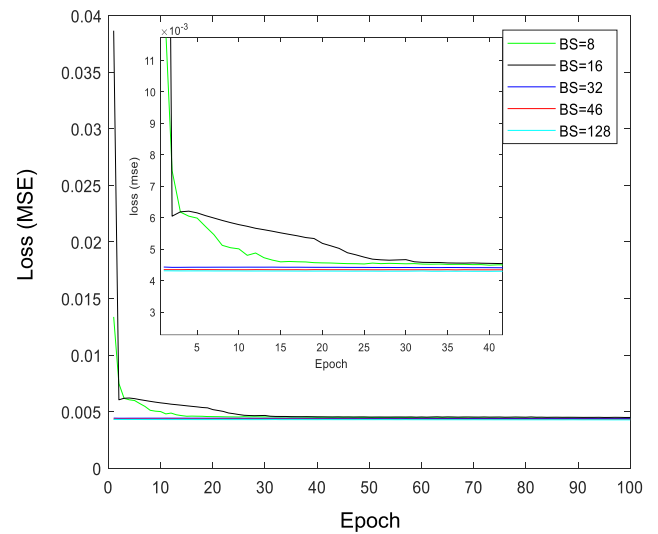


FIGURE 10. Loss (MSE) evolution based on the variation of the batch-size values (8, 16, 32, 64 and 128).

Bi-LSTM model. Several experiments have been carried out in order to deeply assess the influence of some parameters (such as BS, NU, NI, epoch, NL, and dataset size) on the forecasting accuracy including the convergence time and the complexity of the DNN models.

A. EXPERIMENT 1

To verify the impact of the batch size (BS) on the accuracy and the loss function, figure 10 displays the loss function (MSE) during the training phase through varying the BS from 8 to 128.

In the zoomed part of figure 10, it has been observed that high values of BS significantly increase the convergence of the Bi-LSTM model during the training process (starting from BS = 32). In all cases (values of BS), after 50 epochs, the model converges to the same point.

B. EXPERIMENT 2

To verify the effect of the input size (number of inputs: historical values of daily GHI) on the convergence time and the accuracy of the examined Bi-LSTM model, several inputs

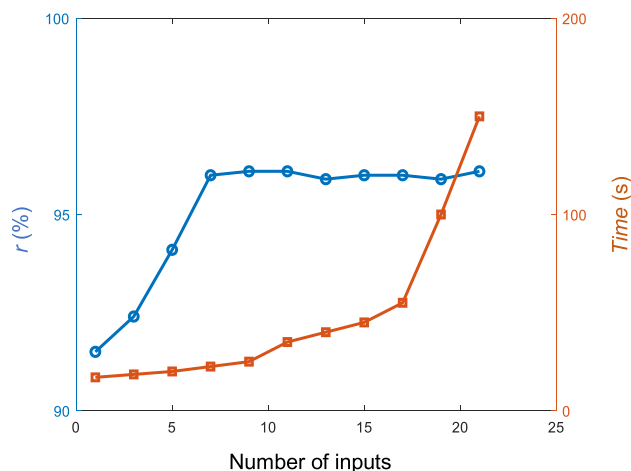


FIGURE 11. Evolution of the calculated r and training time as function of the number of inputs (1 day- 21 days).

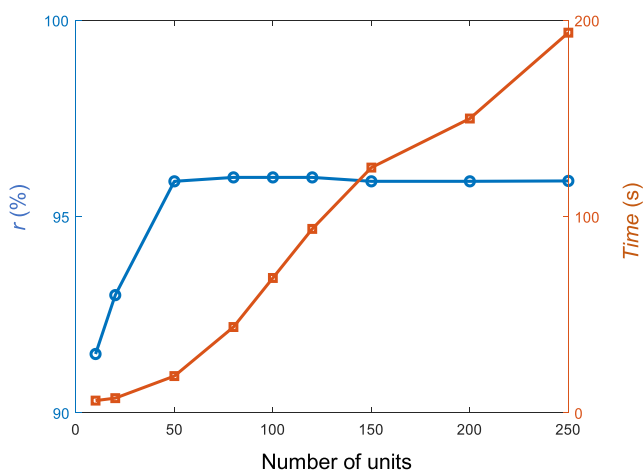


FIGURE 12. Evolution of r and the training time as function of the number of units (Cells in the hidden layer of the Bi-LSTM model).

size varying from 1 day to 21 days have been investigated. Figure 11 displays the evolution of the calculated r and the training time for different inputs sample sizes.

It can be pointed out from figure 11 that the optimal value of r can be reached by using 7 past days as input, though increasing the number of inputs (NI) up to 21 days conducted to a significant increase in the convergence time with a slight variation in r .

C. EXPERIMENT 3

The BS, the epoch and the NI are kept at 64, 100 and 7 respectively and the number of units is varied (from 10 to 250) in order to see its influence on the convergence of the model as well as the forecasting accuracy. The results are depicted in figure 12.

Best results are obtained with 100 units. However, increasing the number of units does not increase significantly the accuracy based on the value of r which is found to globally range between 91% and 96%. It can also be noticed that increasing the number of units led to consuming more time

TABLE 2. Comparison between measured and forecasted daily GHI using different structures of LSTM and GRU.

Model	r (%)	Time (s)
Batch size=64		
Number of inputs =7		
Number of layers=2		
One dense layer		
Activation function (ReLU)		
Optimizer (Adam)		
Epoch =100		
Stacked-LSTM		
1 st configuration (7x100x50x1)	95.0	240
2 nd configuration (7x50x50x1)	95.2	190
Stacked-GRU		
1 st configuration (100x50x1)	95.3	210
2 nd configuration (7x50x50x1)	95.1	170

during the training process without significant improvement of the accuracy.

D. EXPERIMENT 4

In this experiment, the number of hidden layers (HL) have been varied for the cases of LSTM and GRU. For both DNN models, the NU in the first HL is set to 100 units and 50 units in the second HL (1st structure). NU in the first HL is set to 50 units and 50 units in the second HL (2nd structure). The results are reported in table 2. The BS = 64, epoch = 100 and NI = 7.

As shown in table 2, the average value of r is 95%. It can be concluded that the forecasting accuracy is not significantly improved. In addition, much time is required during the training process and the configuration will be complicated. Hence, one hidden layer is enough.

E. EXPERIMENT 5

To verify and confirm the influence of the size of the dataset on the model performance, three sub-datasets have been built: Dataset # 1: (7480 samples) 20 years + 6 months, Dataset # 2 (3740 samples) 10 years + 3 months, and Dataset # 3 (1870 samples) 5 years + 45 days. Results are reported in figure 13a and figure 13b, showing a comparison (superposition and scattered curves) between measured and forecasted GHI using the Bi-LSTM model.

F. COMPARISON AND REMARKS

Table 3 reports a comparison between some recently developed methods for GHI forecasting and our method in terms of the length of the dataset, time horizons and forecasting accuracy. According to table 3, only few papers are related to daily GHI forecasting using deep NNs. The large dataset used is 15 years. LSTM model is the most investigated type of DNNs. Generally, most evaluated DNN-based models provide good results with varying in complexity and accuracy which are the main criteria to select the appropriate model.

With reference to the above experiments, the following points can be highlighted:

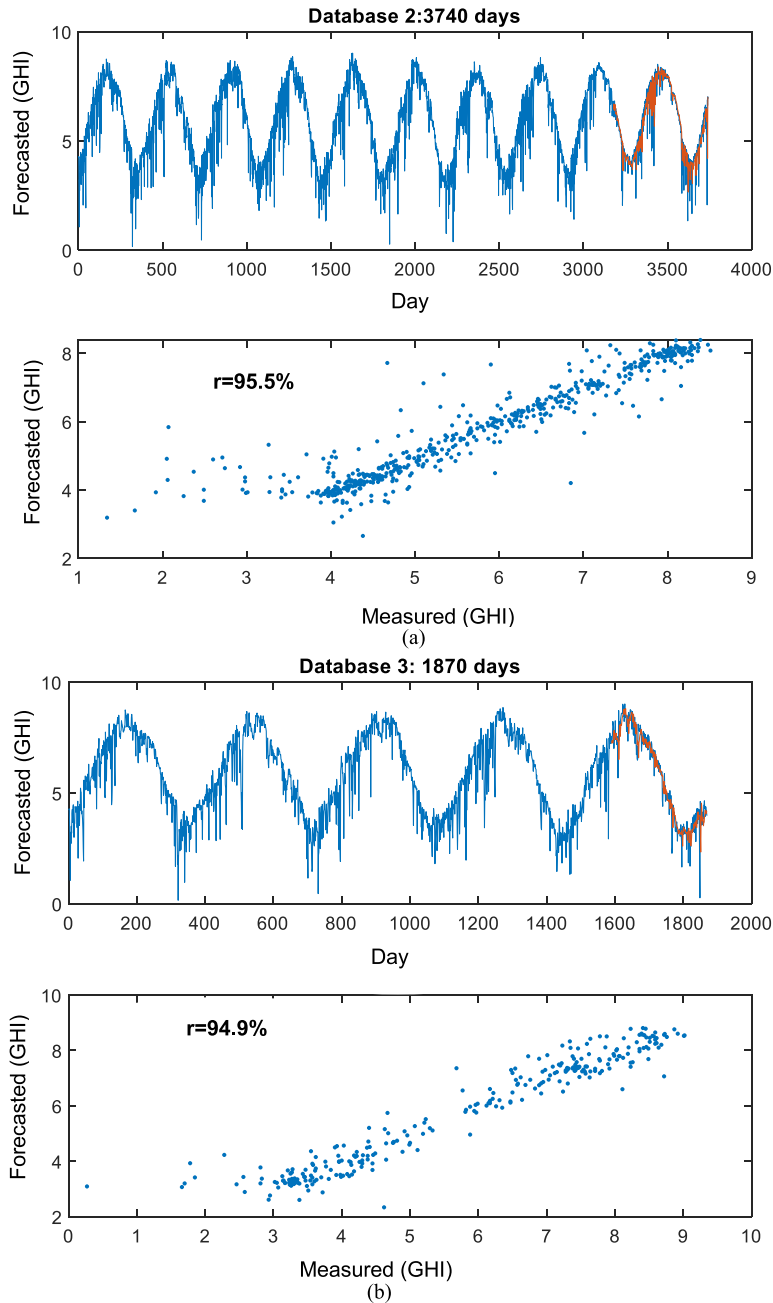


FIGURE 13. a. Measured versus forecasted daily GHI using BiLSTM-based model for the second dataset (superposed and scatter curves): red color indicates the forecasted daily GHI b. Measured versus forecasted daily GHI using BiLSTM-based model for the third dataset (superposed and scatter curves): red color indicates the forecasted daily GHI.

- Using past seven days as input are largely enough to obtain good results. The DNN models become more complicated and take more time in the training process (loss function convergence), when adding more inputs.
- In most examined DNNs, 100 epochs are enough, but in some cases 50 epochs are sufficient to obtain acceptable results.
- Increasing the value of batch-size (e.g. 8, 16, 32, 64 and 128) reduces significantly the convergence time of the

MSE during the training process, but there is no big improvement in the accuracy. A small variation of the r value is observed, therefore in our case $BS = 64$ is the most appropriate.

- Complicated structures of the analyzed DNN-based models, by increasing the number of hidden layers (e.g. stacked LSTM and stacked GRU), do not increase the forecasting accuracy. Therefore, adding more hidden layers is not recommended in this application.

TABLE 3. Comparison between DNNs forecasting methods used in recent selected papers (From 2018 to 2020) and our work.

Ref. & year	Type of DNN	Input	Location	Data collection period	Time horizon	Accuracy
[11], 2018	LSTM	GHI, weather forecast	Santiago, Cape Verde	3 years	Up to 24 h	RMSE=76.24%
[7], 2019	LSTM, GRU	GHI, CS	Seoul, Busa (Korea)	10 years	1-day	RMSE=[5.37%, 5.33%]
[12], 2019	LSTM	GHI, weather data and geographical coordinates	Atlanta, New York, Hawaai	5 years	1-h	MAPE= [8%, 11.1% , 8.1%]
[8], 2019	CNN-LSTM	GSR	Alice Springs (Australia)	13 year	½ -h	MAPE=4.67%
[9], 2019	S2S	GHI	Kharagpur (India)	15 years	Up to 24 h	MAE=30 W/m ²
[13], 2019	LSTM	GHI	Germany, USA, Switzerland, South Korea	7-11 years	1-day	MAE=60.31W/m ²
[10], 2020	LSTM	GHI, DHI, DNI. etc	India	5 years	Up to 24 h	MAPE=6.19%
[14], 2020	LSTM, CI	GHI	Finland	1 year	1-h	RMSE=[30%-34%]
[16], 2020	Cov-LSTM	GHI	Fuhai (Taiwan)	2 years	1-day	RMSE=[0.15%-0.22%]
Our work	LSTM, GRU, BiGRU, BiLSTM, CNN, CNN-BiLSTM,	GHI	Hail (Saudi Arabia)	20 years	1-day	MAPE=[5.24%]

- Increasing the size of the dataset in this area is not needed, this can be explained by the seasonality variation of the data during the year. In contrast to other areas of DL applications, where millions of data are needed (e.g. pattern recognition, natural languages, speech, video, etc.).

V. CONCLUSION

In this paper, various DNN-based models have been developed and evaluated for one-day-ahead forecasting of GHI at Hail region, Saudi Arabia. It has been found that a simple structure such as LSTM or Bi-LSTM can provide good forecasting accuracy ($r = 96\%$).

It has been pointed out that parameters such as batch size, number of units, epochs, and number of hidden layers should be carefully selected. Other parameters such as filters size in CNN, dropout rate and kernel size should be also considered.

It has been verified that with a dataset of 7480 samples, good performances can be achieved by the evaluated DNN models in this specific climate (desert conditions). Globally, the obtained accuracy is enough in our application which concerns the design of PV water pumping systems in this area. To increase the forecasting accuracy, more efforts should be deployed particularly in terms of algorithms' development. Further contributions of the current study can be built around what follows:

- Multi-step ahead forecasting of daily GHI based sequence-to-sequence algorithms (e.g. encoder-decoder LSTM) will be investigated.
- Application of reinforcement learning (e.g. Q-learning) to improve the forecasting accuracy.

- Considering other meteorological variables such as air temperature, wind speed and relative humidity.

APPENDIX

Python code of the developed CNN-LSTM (main)

```
CNN_LSTM = Sequential()
CNN_LSTM.add(TimeDistributed(Conv1D(filters = 64,
kernel_size =2, activation = 'relu'),input_shape = (None,x_train_sub.shape[2], x_train_sub.shape[3])))
CNN_LSTM.add(TimeDistributed(MaxPooling1D(pool_size = 2)))
CNN_LSTM.add(TimeDistributed(Flatten()))
CNN_LSTM.add(LSTM(50, activation = 'relu'))
CNN_LSTM.add(Dropout(0.5))
CNN_LSTM.add(Dense(1))
CNN_LSTM.compile(loss = 'mse', optimizer = 'adam')
CNN_LSTM.summary() ADAM optimization
```

The updated equations for each parameter w^j can be summarized as follows [38]:

$$v_t = \beta_1 * v_{t-1} - (1 - \beta_1) * g_t \quad (a1)$$

$$s = \beta_2 * s_{t-1} - (1 - \beta_2) * g_t * g_t \quad (a2)$$

$$\Delta w_t = -\rho \frac{v_t}{\sqrt{s_t + \epsilon}} * g_t \quad (a3)$$

$$w_{t+1} = w_t + \Delta w_t \quad (a4)$$

ρ : Initial learning rate

g_t : Gradient at time t along w^j .

v_t : Exponential average of gradients along w^j .

s_t : Exponential average of squares of gradients along w^j .

β_1 and β_2 : Hyper-parameters.

Example of input and output samples

Samples (example of 7 inputs and one output)							
Input	Sample #1	Sample #2	Sample #3	Sample #4	Sample #5	Sample #6	Sample #7
	4.23	4.29	3.65	2.57	1.06	3.23	2.72
	4.29	3.65	2.57	1.06	3.23	2.72	3.5
	3.65	2.57	1.06	3.23	2.72	3.5	3.19
	2.57	1.06	3.23	2.72	3.5	3.19	4.08
	1.06	3.23	2.72	3.5	3.19	4.08	4.26
	3.23	2.72	3.5	3.19	4.08	4.26	3.68
	2.72	3.5	3.19	4.08	4.26	3.68	4.37
Output	3.5	3.19	4.08	4.26	3.68	4.37	4.47

NOMENCLATURE

ANN	Artificial neural networks
CNN	Convolutional neural networks
LSTM	Long-Short Term Memory
r	Correlation coefficient
MAPE	Mean absolute percentage error
RMSE	Root mean squared error
GHI	Global horizontal irradiation
S2S	Sequence to sequence
GRU	Gated recurrent unit
RNN	Recurrent neural networks
DL	Deep learning
ML	Machine learning
BiLSTM	Bidirectional LSTM
BiGRU	Bidirectional GRU
ReLU	Rectified linear unit
DT	Decision tree
GSR	Global solar radiation
MLP	Multi-Layer Perceptron
GBRT	Gradient Boosted Regression Trees
K_t	Clearness index
CI	Choquet Integral
T	Air temperature
R_H	Relative humidity
W_S	Wind speed
P_R	Pressure
C_I	Cloud index

ACKNOWLEDGMENT

The authors extend their appreciation to the Deputy of Research & Innovation, Ministry of Education in Saudi Arabia for funding this research work through the project number RDO-2004.

REFERENCES

- [1] A. Mellit and A. M. Pavan, "A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at trieste, italy," *Sol. Energy*, vol. 84, no. 5, pp. 807–821, May 2010, doi: [10.1016/j.solener.2010.02.006](https://doi.org/10.1016/j.solener.2010.02.006).
- [2] A. Graves, F. Santiago, and S. Jürgen, "Multi-dimensional recurrent neural networks," in *Proc. Int. Conf. Artif. Neural Netw.* Berlin, Germany: Springer, 2007, pp. 549–558.
- [3] S. Shamshirband, T. Rabczuk, and K.-W. Chau, "A survey of deep learning techniques: Application in wind and solar energy resources," *IEEE Access*, vol. 7, pp. 164650–164666, 2019, doi: [10.1109/ACCESS.2019.2951750](https://doi.org/10.1109/ACCESS.2019.2951750).
- [4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. New York, NY, USA: MIT Press, 2016.
- [5] C. Opathella, A. Elkasrawy, A. A. Mohamed, and B. Venkatesh, "Optimal scheduling of merchant-owned energy storage systems with multiple ancillary services," *IEEE Open Access J. Power Energy*, vol. 7, pp. 31–40, 2020, doi: [10.1109/OAJPE.2019.2952811](https://doi.org/10.1109/OAJPE.2019.2952811).
- [6] S. Fan, G. He, X. Zhou, and M. Cui, "Online optimization for networked distributed energy resources with time-coupling constraints," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 251–267, Jan. 2021, doi: [10.1109/TSG.2020.3010866](https://doi.org/10.1109/TSG.2020.3010866).
- [7] E. Oh and H. Wang, "Reinforcement-learning-based energy storage system operation strategies to manage wind power forecast uncertainty," *IEEE Access*, vol. 8, pp. 20965–20976, 2020, doi: [10.1109/ACCESS.2020.2968841](https://doi.org/10.1109/ACCESS.2020.2968841).
- [8] C. Chen, M. Cui, F. Li, S. Yin, and X. Wang, "Model-free emergency frequency control based on reinforcement learning," *IEEE Trans. Ind. Informat.*, vol. 17, no. 4, pp. 2336–2346, Apr. 2021, doi: [10.1109/TII.2020.3001095](https://doi.org/10.1109/TII.2020.3001095).
- [9] C. Chen, M. Cui, X. Fang, B. Ren, and Y. Chen, "Load altering attack-tolerant defense strategy for load frequency control system," *Appl. Energy*, vol. 280, Dec. 2020, Art. no. 116015, doi: [10.1016/j.apenergy.2020.116015](https://doi.org/10.1016/j.apenergy.2020.116015).
- [10] M. Cui and J. Wang, "Deeply hidden moving-target-defense for cybersecure unbalanced distribution systems considering voltage stability," *IEEE Trans. Power Syst.*, early access, Oct. 15, 2020, doi: [10.1109/TPWRS.2020.3031256](https://doi.org/10.1109/TPWRS.2020.3031256).
- [11] B. Jason, *Deep Learning for Time Series Forecasting: Predict the Future with MLPs, CNNs and LSTMs in Python*. Vermont, VA, USA: Machine Learning Mastery, 2018.
- [12] P.-H. Kuo and C.-J. Huang, "A high precision artificial neural networks model for short-term energy load forecasting," *Energies*, vol. 11, no. 1, p. 213, Jan. 2018, doi: [10.3390/en11010213](https://doi.org/10.3390/en11010213).
- [13] C. J. Huang, Y. Shen, Y. H. Chen, and H. C. Chen, "A novel hybrid deep neural network model for short-term electricity price forecasting," *Int. J. Energy Res.*, vol. 45, no. 2, pp. 2511–2532, 2020, doi: [10.1002/er.5945](https://doi.org/10.1002/er.5945).
- [14] A. Alzahrani, P. Shamsia, C. Dagli, and M. Ferdowsi, "Solar irradiance forecasting using deep neural networks," in *Proc. IEEE 6th Int. Conf. Renew. Energy Res. Appl. (OCRERA)*, San Diego, CA, USA, Nov. 2017, pp. 5–8.
- [15] *Scikit-Learn*. Accessed: 2011. [Online]. Available: <https://scikit-learn.org/>
- [16] M. Aslam, J.-M. Lee, H.-S. Kim, S.-J. Lee, and S. Hong, "Deep learning models for long-term solar radiation forecasting considering micro-grid installation: A comparative study," *Energies*, vol. 13, no. 1, p. 147, Dec. 2019, doi: [10.3390/en13010147](https://doi.org/10.3390/en13010147).
- [17] S. Ghimire, R. C. Deo, N. Raj, and J. Mi, "Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms," *Appl. Energy*, vol. 253, Nov. 2019, Art. no. 113541, doi: [10.1016/j.apenergy.2019.113541](https://doi.org/10.1016/j.apenergy.2019.113541).
- [18] B. P. Mukhoty, V. Maurya, and S. K. Shukla, "Sequence to sequence deep learning models for solar irradiation forecasting," in *Proc. IEEE Milan PowerTech*, Milan, Italy, Jun. 2019, pp. 1–6.
- [19] D. Chandola, H. Gupta, V. A. Tikkiwal, and M. K. Bohra, "Multi-step ahead forecasting of global solar radiation for arid zones using deep learning," *Procedia Comput. Sci.*, vol. 167, pp. 626–635, Jan. 2020, doi: [10.1016/j.procs.2020.03.329](https://doi.org/10.1016/j.procs.2020.03.329).
- [20] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, Apr. 2018, doi: [10.1016/j.energy.2018.01.177](https://doi.org/10.1016/j.energy.2018.01.177).
- [21] Y. Yu, J. Cao, and J. Zhu, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019, doi: [10.1109/ACCESS.2019.2946057](https://doi.org/10.1109/ACCESS.2019.2946057).
- [22] M. Husein and I.-Y. Chung, "Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach," *Energies*, vol. 12, no. 10, p. 1856, May 2019, doi: [10.3390/en12101856](https://doi.org/10.3390/en12101856).
- [23] M. Abdel-Nasser, K. Mahmoud, and M. Lehtonen, "Reliable solar irradiance forecasting approach based on Choquet integral and deep LSTMs," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1873–1881, Mar. 2021, doi: [10.1109/TII.2020.2996235](https://doi.org/10.1109/TII.2020.2996235).
- [24] J. M. S. de Araujo, "Performance comparison of solar radiation forecasting between WRF and LSTM in Gifu, Japan," *Environ. Res. Commun.*, vol. 2, no. 4, Apr. 2020, Art. no. 045002.
- [25] Y.-Y. Hong, J. J. F. Martinez, and A. C. Fajardo, "Day-ahead solar irradiation forecasting utilizing graminian angular field and convolutional long short-term memory," *IEEE Access*, vol. 8, pp. 18741–18753, 2020, doi: [10.1109/ACCESS.2020.2967900](https://doi.org/10.1109/ACCESS.2020.2967900).

- [26] B. Gao, X. Huang, J. Shi, Y. Tai, and R. Xiao, "Predicting day-ahead solar irradiance through gated recurrent unit using weather forecasting data," *J. Renew. Sustain. Energy*, vol. 11, no. 4, Jul. 2019, Art. no. 043705, doi: [10.1063/1.5110223](https://doi.org/10.1063/1.5110223).
- [27] J. Zhang, A. Florita, B.-M. Hodge, S. Lu, H. F. Hamann, V. Banunarayanan, and A. M. Brockway, "A suite of metrics for assessing the performance of solar power forecasting," *Sol. Energy*, vol. 111, pp. 157–175, Jan. 2015, doi: [10.1016/j.solener.2014.10.016](https://doi.org/10.1016/j.solener.2014.10.016).
- [28] S. Boubaker, "Identification of monthly municipal water demand system based on autoregressive integrated moving average model tuned by particle swarm optimization," *J. Hydroinform.*, vol. 19, no. 2, pp. 261–281, Mar. 2017, doi: [10.2166/hydro.2017.035](https://doi.org/10.2166/hydro.2017.035).
- [29] *POWER Data Access Viewer*. Accessed: Feb. 19, 2020. [Online]. Available: <https://power.larc.nasa.gov/data-access-viewer/>
- [30] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- [31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- [32] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, 1997, doi: [10.1109/78.650093](https://doi.org/10.1109/78.650093).
- [33] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: <http://arxiv.org/abs/1406.1078>
- [34] N. Xue, I. Triguero, G. P. Figueredo, and D. Landa-Silva, "Evolving deep CNN-LSTMs for inventory time series prediction," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Wellington, New Zealand, Jun. 2019, pp. 1517–1524.
- [35] C.-J. Huang and P.-H. Kuo, "A deep CNN-LSTM model for particulate matter (PM_{2.5}) forecasting in smart cities," *Sensors*, vol. 18, no. 7, p. 2220, Jul. 2018, doi: [10.3390/s18072220](https://doi.org/10.3390/s18072220).
- [36] A. Mellit, "Artificial Intelligence technique for modelling and forecasting of solar radiation data: A review," *Int. J. Artif. Intell. Soft Comput.*, vol. 1, no. 1, pp. 52–76, 2008, doi: [10.1504/IJAISC.2008.021264](https://doi.org/10.1504/IJAISC.2008.021264).
- [37] A. Mellit, M. Benghaneim, and S.A. Kalogirou, "An adaptive wavelet-network model for forecasting daily total solar-radiation," *Appl. Energy*, vol. 83, no. 7, pp. 705–722, 2006, doi: [10.1016/j.apenergy.2005.06.003](https://doi.org/10.1016/j.apenergy.2005.06.003).
- [38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*. [Online]. Available: <http://arxiv.org/abs/1412.6980>



SAHBI BOUBAKER received the B.E. degree in electrotechnics and industrial electronics and the M.Eng. degree in systems' analysis and digital processing from the National School of Engineering of Tunis, Tunis, Tunisia, in 1997 and 1999, respectively, and the Ph.D. degree in electrical engineering from the National School of Engineering of Monastir, Monastir, Tunisia, in 2010. He was an Associate Professor at the Department of Electronics, University of Ha'il, Community College, Ha'il City, Saudi Arabia. He is currently an Associate Professor with the Department of Computer and Network Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah City, Saudi Arabia. His research interests include solar energy systems design, time-series modeling, analysis and design of hybrid dynamical systems, including robotics and drones.



MOHAMED BENGHANEM received the B.E., M.Eng., and Ph.D. degrees in electrical engineering from the Polytechnic School of Algiers, in 1987, and the M.Eng. and Ph.D. degrees from USTHB University, Algiers, Algeria, in 1991 and 2000, respectively. He was a Professor at Taibah University, Faculty of Science, Medina, Saudi Arabia, and a Senior Associate with the International Centre of Theoretical Physics, ICTP, Italy, since 2004. He is currently a Professor at Islamic University, Faculty of Science, Physics Department, Medina. His research interests are solar instrumentations, data acquisition systems, renewable energy systems, prediction, and modelling of solar radiation data.



ADEL MELLIT received the M.Eng. and Ph.D. degrees in electronics from the University of Sciences and Technology (USTHB), Algeria, in 2002 and 2006, respectively. He is currently a Professor of Electronics and the Head of Renewable Energy Laboratory at Jijel University, Algeria. His research interests include the application of artificial intelligence techniques in photovoltaic systems and micro-grids. He has authored and coauthored more than 150 articles in international peer-reviewed journals (mostly in Elsevier), and papers in conference proceedings (Mostly in IEEE) mainly on photovoltaic systems, several chapter books, and three books. He is an Associate Member at the ICTP Trieste, Italy. He is serving on the editorial board of the *Renewable energy and Energy journals* (Elsevier Ltd).



AYOUB LEFZA was born in Milia, Algeria, in 1998. He received the master's degree in the field of electronic and embedded systems (EES) from the University of Jijel, Algeria, in 2020. He is currently pursuing the Ph.D. degree with the Renewable Energy Laboratory, Jijel University, Algeria. His research interests include the application of machine learning and deep learning in embedded systems.



OMAR KAHOU LI is currently an Assistant Professor with the Electronics Engineering Department of the Community College, University of Hail, Saudi Arabia. He is a member of the "Control and Energy Management Laboratory," National Engineering School of Sfax, University of Sfax, Tunisia. His current research projects focus on interdisciplinary applications of the electrical power system and Renewable Energy. He is an active contributor to his research fields through the different original works published either in high ranking journals or in conference proceedings. His main research interests include renewable energy (wind turbine generation, solar energy), power system control and stability, application of non-conventional techniques for power system stabilizers, reconfiguration and planification of the distribution power networks.



LIOUA KOLSI is currently an Associate Professor with the Mechanical Engineering Department, University of Ha'il, Saudi Arabia, where he is also the Coordinator of the Research Excellence Unit at the Deanship of Scientific Research. He has coauthored numerous (more than 100) highly cited journal publications, conference articles, and book chapters in the aforementioned topics and has received several awards and grants from various funding agencies. His H-index is 24 according to Web of Science and he was awarded by Publons as one of the 1% best reviewers. His current research projects focus on interdisciplinary applications of nanotechnology and Heat and fluid flow. His areas of expertise include nanofluids applications, CFD, heat and mass transfer, flow visualisation, and MHD.

...