

Short-Term Solar Irradiance Forecasting Model Based on Bidirectional Long Short-Term Memory Deep Learning

Fahad Radhi Alharbi
Engineering Department
Lancaster University
Lancaster, UK
f.alharbi@lancaster.ac.uk

Denes Csala
Engineering Department
Lancaster University
Lancaster, UK
d.csala@lancaster.ac.uk

Abstract—Recurrent neural networks (RNNs) are the most effective technology to study and analyze the future performance of solar irradiance. Bidirectional RNNs (BRNNs) provide the key benefit of manipulating the data with two different hidden layers in two opposite directions and can feed back to the same layer of output. In this approach, the output layers can simultaneously receive information from the past (backward layers) and the future (forward layers). A bidirectional long short-term memory (BI-LSTM) model was developed and employed to predict solar irradiance values for the next 169 hours based on hourly historical data (01-01-1985 to 16-09-2020) from Tabuk city. The findings specifically demonstrate that in terms of classification and considerations, the BI-LSTM model has promising performance with notable accuracy. In addition, the model is capable of coping with the selected size of sequential data. The prediction accuracy and performance of the BI-LSTM model were highly enhanced when external data such as wind speed and temperature were provided.

Keywords— BI-LSTM, recurrent neural network, backward layers, forward layers, bidirectional, prediction, solar irradiances.

I. INTRODUCTION

Solar energy is perceived to be a great solution to coping with climate crisis and the power generation issues, making it the most important option along with the increasing interest in renewable energy [1]. Studies devoted to wind and solar energy mainly focus on the intermittent, fluctuation, and non-dispatchable issues of these kinds of energy and the stability of conventional electricity systems. Solar energy forecasting has been commonly used as a sufficient technique to alleviate multiple issues caused by intermittent issues. Several approaches have been proposed to build effective models that can be applied to predict the pursuit and behaviour of solar energy [1-11]. Zhao et al. [12] mentioned that time series prediction models that focus on historical data or historical observations, such as artificial neural networks (ANNs), are appropriate for ultra-short-term minutes to hours. In addition, Yadav and Chandel [13] thought that most of these approaches are based on complicated statistics, artificial intelligence (AI) techniques and vast volumes of topographical and meteorological historical data. However, these approaches can

minimize the probability of solar energy system failure by predicting future scenarios. According to Alsharif et al. [1], prediction models contain three main categories: (i) Qualitative approaches, (ii) Quantitative approaches, and (iii) ANNs. (i) Qualitative approaches are based on experts' experience, judgment and/or resource evaluation [1]. (ii) Quantitative approaches are based on mathematical models and can be categorized as time series or causal forecasting techniques that determine the relationships between the variables as dependent and independent [1]. However, the accuracy of these approaches is questionable due to uncertain and high fluctuation factors of solar energy. (iii) ANN techniques are developed to demonstrate the method of how the human brain's nervous system analyzes and controls more complicated data such as nonlinear data, optimization and pattern recognition [5]. ANNs are the cornerstone of AI which address problems that would be impossible to solve using human or computational criteria. ANN techniques have learning abilities which allowing them to achieve the best possible results as more data become available and can connect to large numbers of neurons to process any data and respond to it automatically with no prior assumption. There are two types of deep learning (DL) of models: (i) recurrent neural networks (RNNs) [14] and (ii) convolutional neural networks (CNNs) [15]. In addition, RNNs are a class of ANNs which figure out how much information is required for each forecasting step, and they can access data from any point in the acoustic series [16]. Bidirectional long short-term-memory (BI-LSTM) recurrent neural networks (RNNs) are widely utilized in natural language processing (NLP) as an approach that produces the lowest recorded value, particularly with time series forecasting error metrics [16], [17]. RNNs have achieved promising performance in processing sequences of arbitrary length and modeling tasks [18]. This technology made a great leap in the field of forecasting, such as solar energy forecasting, and made it a more reliable method with notable accuracy.

In this paper, a prediction model will be built based on BI-LSTM to estimate the future solar irradiance values, which are the global horizontal irradiance (GHI), direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI), for one week ahead as hourly values (169—h). The remainder of this paper is organised as follows: Section II contains the model architecture

and algorithm. Section III debates the results and discussion. Finally, the conclusions present in Section IV.

II. PREDICTIVE MODEL ARCHITECTURE AND ALGORITHM

In this section, the main concepts and the architecture of the RNN, LSTM and BI-LSTM will be discussed along with their algorithms.

A. Recurrent Neural Networks (RNN)

The RNN emulates a dynamical system of discrete time that has sequence vectors as input variables at a (t) time step [19],[8]. The RNN input patterns are defined as $(x(1), x(2), \dots, x(t))$, and every sample/datapoint $x(t)$ is considered to be real-valued variable [8]. The $(y(1), y(2), \dots, y(t))$ represents the goal sequence vectors or the predicted target datapoint of the RNN, and its hidden layers are $(h(1), h(2), \dots, h(t))$ (see Fig. 1a).

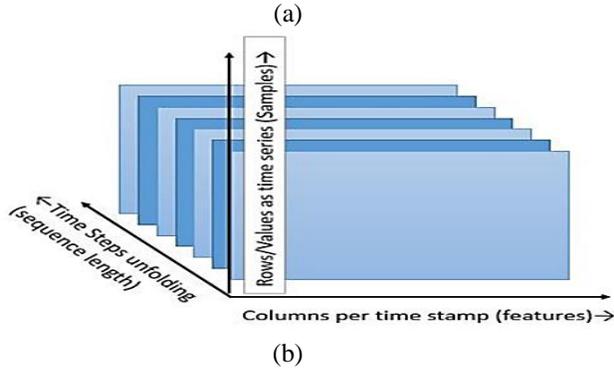
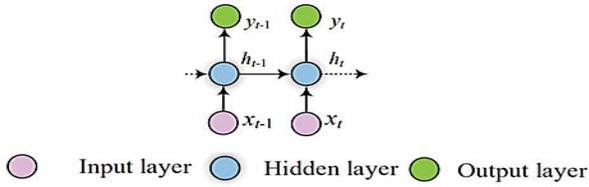


Fig. 1. Structure of RNN (a) [19] and input representation for the RNN (b) [8].

The RNN architecture is described by relations between all of its layers, this enables mapping from the present input to the required vectors, as well as mapping the required vectors from the entire previous inputs. [20]. The RNN has 3 parameters of the input as follows: (i) The minibatch size, which is the sample length (datapoints); (ii) The columns' number in the vector per time steps, which are the features in the vector input; and (iii) The timesteps number, which presents a distinguishing feature of RNN, and it can unfold the input vector through the time (see Fig. 1b) [8]. The input pattern is injected into the neurons in each layer, where it is multiplied by the activation function to generate the neuron's median output, which is then used as the input to the next neuron in the second phase layer [8]. S_i is an overall input to the next layer's neuron, which is the connection weight (W) multiplied by the previous neuron's output and aggregated to the bias ($b \in \mathbb{R}$), as presented in equation 1 [8].

$$S_i = W_i \cdot x_i + b \quad (1)$$

In addition, g is the function of activation applied to the S_i to generate neuron's output (a_i), as shown in equations 2 and 3 [8].

$$a_i = g(S_i) \quad (2)$$

$$a_i = g(W_i \cdot x_i + b) \quad (3)$$

At time (t), the neurons with recurrent edges fed by the input from the current sample x_t with another input from the hidden layer $h(t-1)$, as shown in equation 4. In equation 4, H is the hidden layer, W_{hx} represents the conventional weight matrix between the input and the hidden layer, and W_{hy} is the recurrent weights matrix [8]. The output $y(t)$ at each time step (t) can be found by applying equation 5 to hidden layer $h(t)$. In addition, the b_y and b_h represent bias parameters.

$$h_{(t)} = H(W_{hx} x_t + W_{hh} h_{t-1} + b_h) \quad (4)$$

$$y_{(t)} = g(W_{hy} h_t + b_y) \quad (5)$$

B. Long Short-Term Memory (LSTM)

LSTM is a RNN variant that was proposed in 1997 [21], [22]. LSTM can overcome the issues of vanishing and exploding gradient challenges during the training in RNNs [21]. The best way to find and exploit the long-range context is LSTM architecture-built memory cells that store data [16], [22]. Fig. 2 shows the configuration of an LSTM memory cell including the main components of the LSTM, such as the addition of element levels; where the multiplication sign refers to the element-level multiplication and the vector merge is denoted by con. LSTM consists of four functions with the same size: (i) The forget gate (f_t), which regulates the amount of data that remains in the cell. (ii) The input gate (i_t), which monitors the amount of data that flows into the cell. (iii) The input node (g_t), which can be a cell activation vector. (iv) The output gate (o_t), which performs the output activation of the LSTM. In addition, the LSTM functions can be presented as equations such as equations 6-10. In these equations, σ is the logistic sigmoid function, and h_t is the hidden vector.

$$f_{(t)} = \sigma(W_{xf} x_t + W_{hf} h_{t-1} + W_{cf} c_{t-1} + b_f) \quad (6)$$

$$i_{(t)} = \sigma(W_{xi} x_t + W_{hi} h_{t-1} + W_{ci} c_{t-1} + b_i) \quad (7)$$

$$g_{(t)} = f_t g_{t-1} + i_t \tanh(W_{xg} x_t + W_{hg} h_{t-1} + b_g) \quad (8)$$

$$o_{(t)} = \sigma(W_{xo} x_t + W_{ho} h_{t-1} + W_{co} c_t + b_o) \quad (9)$$

$$h_{(t)} = o_t \tanh(g_t) \quad (10)$$

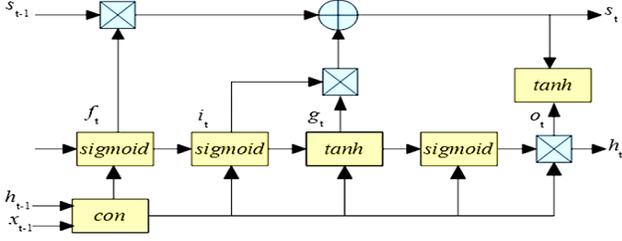


Fig. 2. The structure of the long short-term memory (LSTM) unit.

C. Bidirectional long short-term memory (BI-LSTM)

The main advantage of Bidirectional RNNs (BRNNs) is that they process data in two directions with two separate hidden layers and can feed forward to the same output layer [17], [23]. The main components of bidirectional RNNs are the forward hidden sequence \vec{h} , the backward hidden sequence \overleftarrow{h} and the output sequence (y) that iterate the backward layer from $t = T$ to 1, as illustrated in Fig. 3. In addition, the output layer of a bidirectional RNN can be presented in equations 11—15.

$$\vec{h}_{(t)} = H(W_{x\vec{h}} x_t + W_{\vec{h}\vec{h}} \vec{h}_{t-1} + b_{\vec{h}}) \quad (11)$$

$$\overleftarrow{h}_{(t)} = H(W_{x\overleftarrow{h}} x_t + W_{\overleftarrow{h}\overleftarrow{h}} \overleftarrow{h}_{t-1} + b_{\overleftarrow{h}}) \quad (12)$$

$$y_{(t)} = (W_{y\vec{h}} \vec{h}_t + W_{y\overleftarrow{h}} \overleftarrow{h}_t + b_y) \quad (13)$$

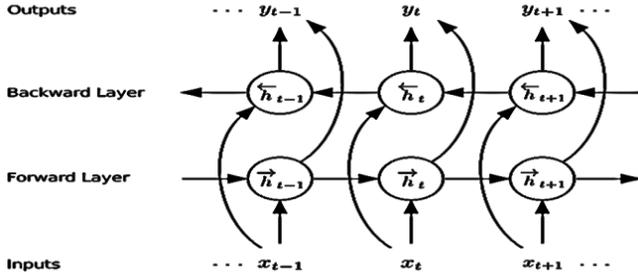


Fig. 3. The bidirectional RNN unit [17].

BI-LSTM will be generated by combining BRNNs with LSTM, which can process long-range contexts in both input directions [17]. Deep RNNs can be built by integrating several RNN hidden layers on top of one another, including the output sequence of one layer that shaping the input iteration process of the next. The same hidden layer can be employed for N numbers of layers in the stack, and the hidden vector series h^n can be iteratively calculated from $n=1$ to N ($t=1$) as presented in equation 14, where $h^0 = x$. y_t refers to the final network output, as illustrated in equation 15.

$$h^n_{(t)} = H(W_{h^{n-1}h^n} h^{n-1}_t + W_{h^n h^n} h^n_{t-1} + b_{h^n}) \quad (14)$$

$$y_{(t)} = (W_{h^N y} h^N_t + b_{(y)}) \quad (15)$$

BRNNs can replace the hidden sequences (h^n) with the backward sequences ($h^{n\leftarrow}$) and the forward sequences ($h^{n\rightarrow}$). In BI-LSTM, every single hidden layer obtains input from both

backward and forward layers. BI-LSTM uses LSTM for the hidden layers.

D. Data description and preparation

Tabuk city in Saudi Arabia was selected as a case study. It is located between $28^\circ 23' 59.28''$ N in latitude and $36^\circ 34' 17.44''$ E in longitude. A historical data was collected to predict the future performance of the three types of solar irradiance (W/m^2), including global horizontal irradiance (GHI), direct normal irradiance (DNI) and diffuse horizontal irradiance (DHI), one week ahead hourly values (169—h). The historical data were obtained from the station situated in the meteoblue weather service [24]. The historical wind speed (m/s) and temperature ($^\circ\text{C}$) data of Tabuk city were employed as external data to provide more details to the BI-LSTM model and to support the model due to the high correlation between them. All the gathered historical data used in this study cover the period from 01-01-1985 to 16-09-2020, and hourly values are shown in Figs (4—10). To achieve better future prediction values, the quality and quantity of the datasets are important. The data were reviewed, checked, and analyzed to maintain any missed or duplicated values. All the datasets were maintained as stationary, and the p-value hypotheses were checked by applying the Dickey-Fuller test to confirm that, the p-values are less than 0.05 (p-values < 0.05).

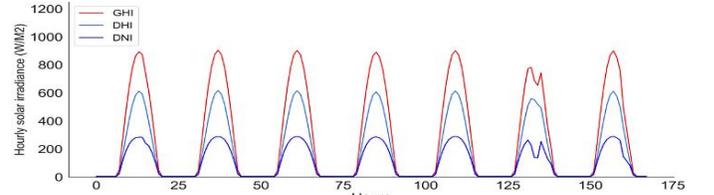


Fig. 4. The DHI, GHI and DNI performance as hourly values of the first week during January 1985.

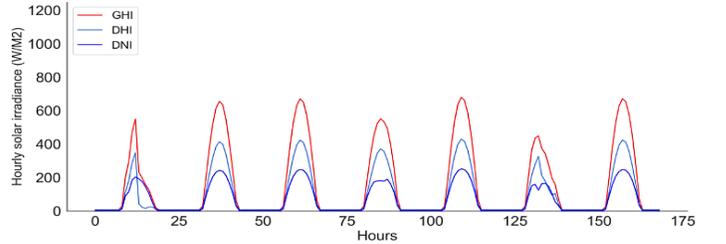


Fig. 5. The DHI, GHI and DNI performance as hourly values of the second week during September 2020.

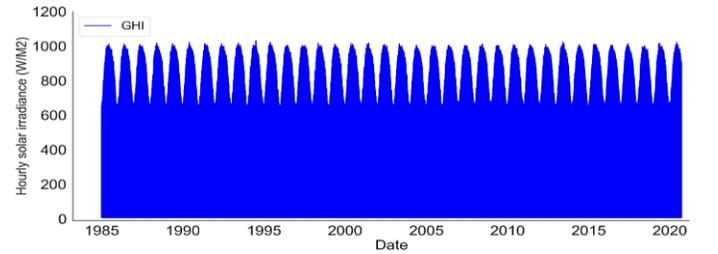


Fig. 6. The GHI performance as hourly values over 35 years.

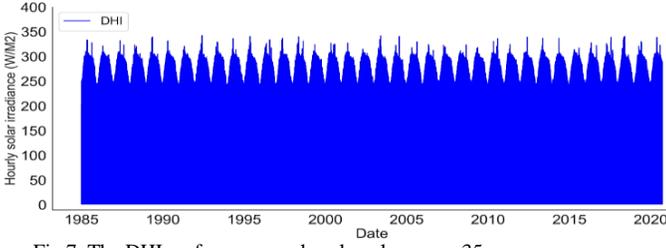


Fig. 7. The DHI performance as hourly values over 35 years.

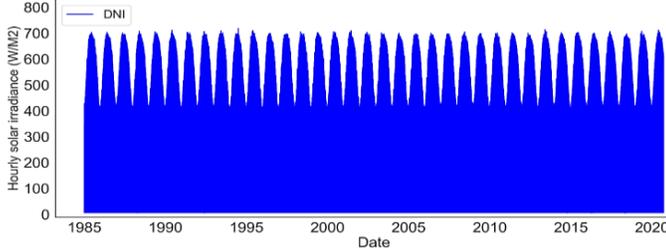


Fig. 8. The DNI performance as hourly values over 35 years.

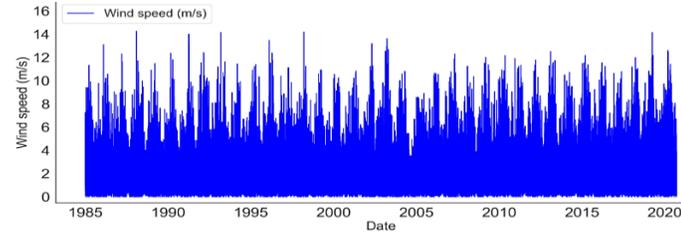


Fig. 9. The wind speed performance as hourly values over 35 years.

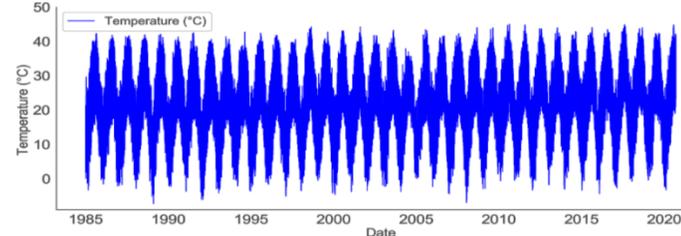


Fig. 10. The temperature performance as hourly values over 35 years.

E. Model structure

Python programming language was employed to build the predicted model due to the large memory of Python, which is able to handle large amounts of historical data. The historical data were prepared and cleaned to be processed. The BI-LSTM model was built as a code based on the mathematical algorithm (section II), and the model parameters were defined as 3 hidden layers (Fig. 11). The first and second layers were set to 100 neurons, and the third layer was set to 50 neurons. If the number of hidden neurons increases, the predictive errors decrease sharply. However, the root mean square error remains almost constant with only small variations when the number of neurons is more than 150 and less than 300, created by the randomization of input weights for extreme learning machine networks [12]. If the number of inputs is greater than 2, the predictive error will be dramatically higher. In addition, the number of inputs and the number of neurons of hidden layer are the two main parameters for BI-LSTM in factorization learning machine networks. The number of epochs is adjusted as 30, batch size is 64 and the validation split is 0.011. The target training data were adjusted to the maximum size, and the test size was set at the minimum

size to reduce the simulation time and to obtain a better fit of the model due to the specific size of the historical data. Moreover, for each future predicted value of the solar irradiance, the other historical data, such as wind speed, temperature and the other two types of solar irradiance data, were used as external support data, which can provide the model with more details due to the high coloration between the data.

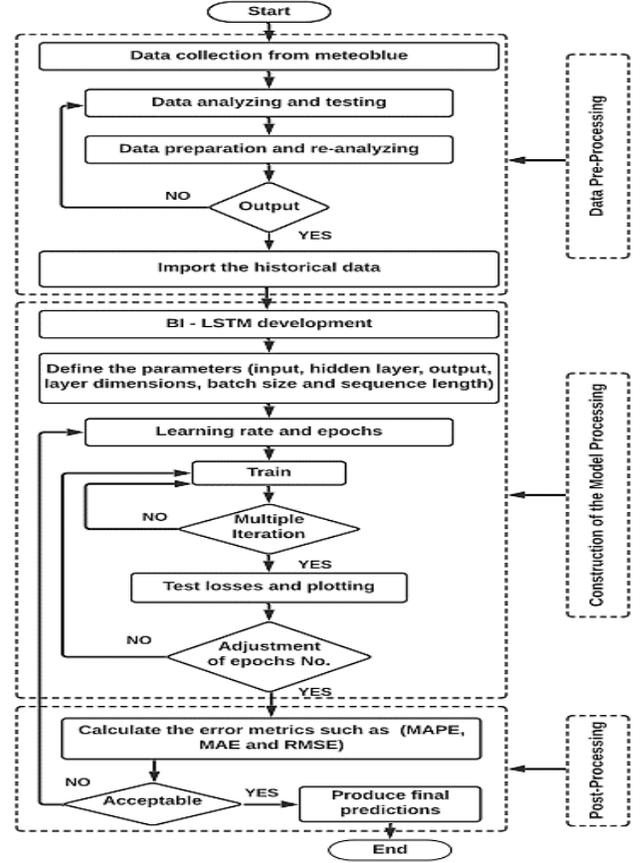


Fig. 11. Flowchart of the proposed model.

F. Error metrics and Evaluation

The BI-LSTM model's accuracy can be evaluated using various statistical error metrics such as the root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Moreover, the R-squared (R^2) is a statistical metric reflecting the percentage of the variance of a dependent variable accounted for an independent variable. Although the intensity of the relationship between an independent and dependent variable is explained by the correlation, the R^2 describes how the variance of one variable explains the variance of the second variable. The various statistical error metrics that are used to evaluate the output of the BI-LSTM model are described by equations 16-19. In these equations, j is the generic hour of the predicted data, n is the number of forecasted datapoints, y_j is the real values and \hat{y}_j is the predicted observation value. SS_{res} is the sum of squared residuals; and SS_{Tot} is the absolute number of squares, which is proportional to the variance of the data.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^n (y_j - \hat{y}_j)^2} \quad (16)$$

$$\text{MAE} = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (17)$$

$$\text{MAPE} = \frac{1}{n} \sum_{j=1}^n \left| \frac{y_j - \hat{y}_j}{y_j} \right| * 100 \quad (18)$$

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{Tot}}} \quad (19)$$

III. RESULTS AND DISCUSSION

Based on the historical solar irradiance data and the external historical wind speed and temperature data, the future values of the three types of solar irradiances were predicted. The results of the BI-LSTM model illustrated a significant performance and notable fitting of the model (see Fig. 12). The number of epochs and the error values have a direct relationship, which means that if the number of epochs increases, the errors will decrease. In addition, using a large historical dataset with a large amount of training data resulted in slow simulation times with notable error values. As presented in Fig. 12, the BI-LSTM model does not show major overfitting, even with this type of training data. Overfitting would be minimized by the size of the dataset. In some cases, such as future prediction, rather than splitting the dataset into training (80%) and testing (20%), we used all of our data for training and reduced the testing data to the minimum size; however, the model can still handle this technique and generated acceptable results, as shown in Fig. 13. The BI-LSTM model perfectly considered the zero values of the solar irradiances during the sunset period (6 PM—5 AM) and predicted them as zero values, which means that this model has promising performance. Furthermore, 169 hours from 17-09-2020 to 23-09-2020 were predicted, as illustrated in Fig. 13, with acceptable error metrics. The MAPE is 4.3%, the MAE is 11 W/m² and the RMSE is 31 W/m², as shown in Table 1. In addition, the R² value is 0.98%, which shows acceptable correlation and variance between the variables. The accuracy of the predicted models primarily depends on the number of variables. The formulas do not balance dimensional characters, this creates uncertainty. If multiple variables most used, it is required to specify the units and quantities in the equations.

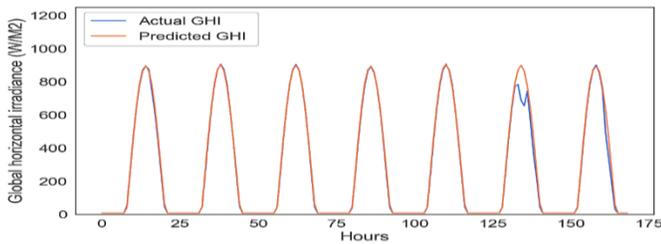


Fig. 12. The actual and predicted GHI performance show the model fitting.

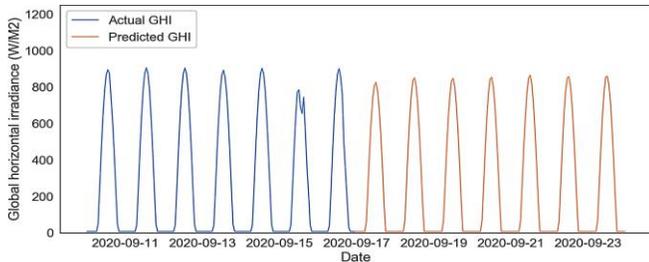


Fig. 13. The future predicted values of the GHI 169 hours ahead.

The BI-LSTM model has made incredible progress in discriminative tasks for the DNI, as depicted in Figs. 14 and 15. This has been fueled by the development of model architectures and the use of big data with the support of external data, which have played significant roles in developing the model. The RMSE and MAE are improved to 11.6 W/m² and 6 W/m². In addition, R² also improved to 0.99%, and the MAPE is 19%, as illustrated in Table 1. However, one of the most challenging issues is enhancing the generalization ability of this model and achieving high-accuracy outputs. Generalizability indicates the difference in the performance of the model when evaluating previously seen data, such as training data versus data the model has never seen before, such as testing data. In addition, the poor generalizability of the model can result in overfitting of the training data. The BI-LSTM model predicts the DHI with acceptable error metrics. Its MAPE is 29%, and its RMSE is 19.8 W/m², as shown in Table 1 (see Figs. 16 and 17). The MAE is 5.9 W/m² and R² is 0.97%, which means that the values have high correlation. It is clear that the BI-LSTM model can cope with the three types of solar irradiance data. Predictive models are important in the predictive control of energy systems to optimize energy-saving performance. Therefore, it is essential to upgrade the model by learning the observed data on a regular basis in order to enhance the model's potential results. The BI-LSTM model would be suitable for predicting solar irradiances for other location applications.

TABLE I. THE FORECASTING ACCURACY OF THE BI-LSTM MODEL.

No.	Solar irradiance	P-value (%)	RMSE (W/m ²)	MAE (W/m ²)	MAPE (%)	R ² (%)
1	GHI	5.8 e-22	31	11	4.3	0.98
2	DNI	6.4 e-24	11.4	6	19	0.99
3	DHI	2.5 e-23	19.8	5.9	29	0.97

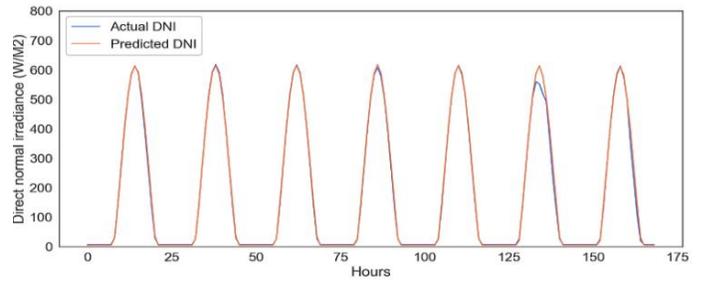


Fig. 14. The actual and predicted DNI performance show the model fitting.

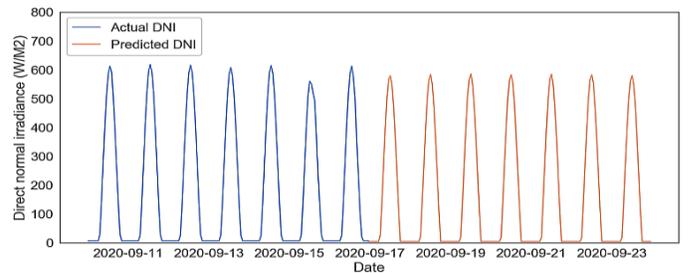


Fig. 15. The future predicted values of the DNI 169 hours ahead.

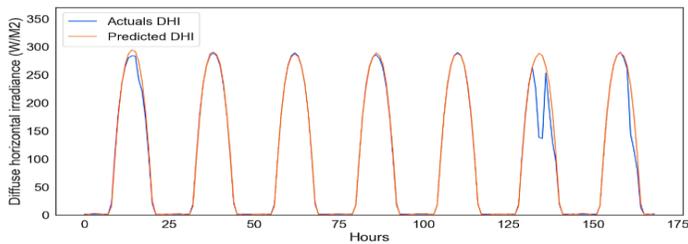


Fig. 16. The actual and predicted DHI performance show the model fitting.

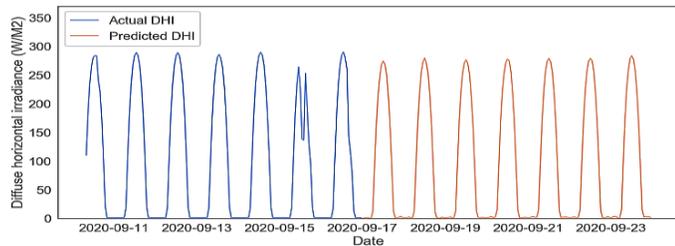


Fig. 17. The future predicted values of the DHI 169 hours ahead.

IV. CONCLUSIONS

The BI-LSTM prediction model was developed by using strength of neural connections and layers based on both backward and forward layers, which are the best described prediction patterns of solar irradiances in this model. The BI-LSTM model was verified to predict 169 hours ahead based on the historical weather data of Tabuk city in Saudi Arabia. The data included three types of solar irradiances, which are the GHI, DNI and DHI. The historical wind speed and temperature data of Tabuk city were used as external data to support the model due to the high correlation between them. The hourly values from 01-01-1985 to 16-09-2020 were used to study the ability of BI-LSTM to cope with these types of historical data. The BI-LSTM model predicted the GHI, DNI, and DHI 169 hours ahead with excellent predictive performance supported by RMSEs of 40, 13.3 and 20.4 W/m², respectively. The BI-LSTM model was shown to achieve maximum and minimum MSEs of 13–5.9 W/m² and MAPEs of 4.8%, 21% and 27%. In addition, this model is characterized by high computational performance. However, the model can be improved in several ways including applying hyperparameter tuning, reinforcing the hidden layers and further testing the numbers of epochs or learning iterations to maximize the learning performance; however, extreme learning machine performance environments can necessitate a significant amount of time as well as a high-performance learning system or overfitting could occur. Finally, extending the proposed BI-LSTM model to predict wind speeds will be a potential future research subject in this domain.

REFERENCES

[1] M. H. Alsharif, M. K. Younes, and J. Kim, "Time series ARIMA model for prediction of daily and monthly average global solar radiation: The case study of Seoul, South Korea," *Symmetry*, vol. 11, no. 2, p. 240, 2019.

[2] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PLoS one*, vol. 13, no. 3, p. e0194889, 2018.

[3] R. H. Inman, H. T. C. Pedro, and C. F. M. Coimbra, "Solar forecasting methods for renewable energy integration," *Progress in energy and combustion science*, vol. 39, no. 6, pp. 535-576, 2013.

[4] A. Tuohy et al., "Solar forecasting: methods, challenges, and performance," *IEEE Power and Energy Magazine*, vol. 13, no. 6, pp. 50-59, 2015.

[5] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, "Solar photovoltaic generation forecasting methods: A review," *Energy Conversion and Management*, vol. 156, pp. 459-497, 2018.

[6] C. A. Severiano, P. C. L. Silva, H. J. Sadaei, and F. G. Guimarães, "Very short-term solar forecasting using fuzzy time series," 2017: IEEE, pp. 1-6.

[7] A. Zagouras, H. T. C. Pedro, and C. F. M. Coimbra, "On the role of lagged exogenous variables and spatio-temporal correlations in improving the accuracy of solar forecasting methods," *Renewable Energy*, vol. 78, pp. 203-218, 2015.

[8] S. Mishra and P. Palanisamy, "Multi-time-horizon solar forecasting using recurrent neural network," 2018: IEEE, pp. 18-24.

[9] D. Yang, J. Kleissl, C. A. Gueymard, H. T. C. Pedro, and C. F. M. Coimbra, "History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining," *Solar Energy*, vol. 168, pp. 60-101, 2018.

[10] B.-k. Jeon and E.-J. Kim, "Next-Day Prediction of Hourly Solar Irradiance Using Local Weather Forecasts and LSTM Trained with Non-Local Data," *Energies*, vol. 13, no. 20, p. 5258, 2020.

[11] J. Wojtkiewicz, M. Hosseini, R. Gottumukkala, and T. L. Chambers, "Hour-Ahead Solar Irradiance Forecasting Using Multivariate Gated Recurrent Units," *Energies*, vol. 12, no. 21, p. 4055, 2019.

[12] Y. Zhao, L. Ye, Z. Li, X. Song, Y. Lang, and J. Su, "A novel bidirectional mechanism based on time series model for wind power forecasting," *Applied energy*, vol. 177, pp. 793-803, 2016.

[13] A. K. Yadav and S. S. Chandel, "Solar radiation prediction using Artificial Neural Network techniques: A review," *Renewable and sustainable energy reviews*, vol. 33, pp. 772-781, 2014.

[14] K.-i. Funahashi and Y. Nakamura, "Approximation of dynamical systems by continuous time recurrent neural networks," *Neural networks*, vol. 6, no. 6, pp. 801-806, 1993.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84-90, 2017.

[16] A. Graves, N. Jaitly, and A.-r. Mohamed, "Hybrid speech recognition with deep bidirectional LSTM," 2013: IEEE, pp. 273-278.

[17] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," 2013: IEEE, pp. 6645-6649.

[18] G. Liu and J. Guo, "Bidirectional LSTM with attention mechanism and convolutional layer for text classification," *Neurocomputing*, vol. 337, pp. 325-338, 2019.

[19] A. Alzahrani, P. Shamsi, M. Ferdowsi, and C. Dagli, "Solar irradiance forecasting using deep recurrent neural networks," 2017: IEEE, pp. 988-994.

[20] F. Wang, Y. Yu, Z. Zhang, J. Li, Z. Zhen, and K. Li, "Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting," *Applied Sciences*, vol. 8, no. 8, p. 1286, 2018.

[21] Y. Liu et al., "Wind power short-term prediction based on LSTM and discrete wavelet transform," *Applied Sciences*, vol. 9, no. 6, p. 1108, 2019.

[22] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

[23] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE transactions on Signal Processing*, vol. 45, no. 11, pp. 2673-2681, 1997.

[24] Meteoblue G, Basel, Switzerland. Accessed: Aug. 20, 2020. [Online]. Available: <https://www.meteoblue.com/en/historyplus>