LSTM Deep Learning Method for Radiation Short and Long-Term Prediction

Mohamed Khala[1(✉)], Houda Abouzid[2], Sara Teidj[3], Omar Eloutassi[1], and Choukri Messaoudi[1]

[1] Optoelectronics and Applied Energy Techniques, Faculty of Science and Technology, Moulay Ismail University of Meknes, Meknes, Morocco
`mo.khala@edu.umi.ac.ma`

[2] Computer Science and New Technology Trends, Abdelmalek Essaadi University, Tetouan, Morocco

[3] Department of Physics, Faculty of Science and Technology, Moulay Ismail University of Meknes, Meknes, Morocco

Abstract. The renewable energy market, like conventional energy, is often unstable. Solar energy in particular presents disturbances linked to the spatiotemporal fluctuations that solar radiation undergoes. To remedy this problem, it is essential to predict its fluctuations that have exhausting consequences for integration into the electricity network and reduce the dependence on fossil fuels essential for socio-economic development. The main objective of this study is to propose a prognostication model based on artificial intelligence to predict with great accuracy the fluctuations of global solar radiation (GSR). Indeed, we adopted a deep dynamic recurrent memory artificial neural network (ANN) model which can manipulate the dynamics of atmospheric factors and is capable to save the history of these characteristics. For this purpose, we developed two long-short term memory (LSTM) models for short and long term forecasting of global solar radiation time series. The first model is a monovariate LSTM, the second one is a multivariate LSTM, after selecting the input variables (predictors) by correlation study, the models are trained by dividing the data set into two subsections, the first subsection is used for model training and the second subsection is reserved for testing. The indices and results established in this work demonstrate the robustness attributed to our method which can provide power grid managers at Er-Rachidia, MOROCCO (Latitude: 31°55′53″N; Longitude: 4°25′35″ W; Altitude: 1039 m) credible predictions to ensure better supervision of the energy system and to adapt solar energy to the electricity production system, especially multivariate LSTM model which proved more accuracy compared to monovariate LSTM.

Keywords: Global solar radiation · Prediction · Deep artificial neural network · Long short time memory

1 Introduction

The world faced severe crises related to energy issues. To ease these crises, it is essential to make a transition to renewable energies. Nevertheless, the integration of this type of energy in the Grid is not always obvious and often unsuitable and destabilizes the Grid.

The deployment of renewable energy sources and their storage reinforces the idea of the Smart Grid to solve the problem of interruption of renewable energy, stabilize the frequency of grid integration and prevent back-feeding to avoid the reversal of the electric flow [1]. Since December 2015, Morocco has set a challenge to accelerate its energy transition toward a smart grid by increasing the installed power from renewable energies to up 52% by 2030 [2].

Current scientific research dedicated to the integration of renewable energies is directed towards the exploitation and application of new technologies to accurately predict solar energy. Artificial intelligence with its rapidly growing in several areas [3–5] remains a very relevant choice, especially with the appearance of new information technologies and the availability of a large database. Indeed, several recent researches based on Machine Learning (ML) algorithms have been adopted with the aim of the prediction and optimization of energy systems.

In this context, Voyant et al. [6] listed several journals associated with ML processes that aim to predict global solar radiation. Several methods of predictors demonstrate that it is the ANNs that are generally used to establish a set of predictors (> 70% of cases). Some articles have reviewed various hybrid approaches involving decision tree-based ML methods and non-linear learning i.e. Deep Learning (DL) methods such as Long- and Short-Term Memory neural networks (LSTM) and Convolution Neural Networks (CNN) [7–9]. Other papers have focused on LSTM and/or Gated Recurrent Unit (GRU) memory neural networks to highlight and use time series of weather data that helps in optimizing the power of output of photovoltaic parks [10–12]. LSTM networks have shown great importance in terms of storing and transporting information over long distances and capturing long-term dependencies. Recently Jallal et al. [13, 14] proposed two DL algorithms, able to handle nonlinearities and dynamic behaviour of meteorological data and also able to generate accurate real-time predictions of RSG data, the first algorithm achieves a coefficient of correlation R = 99.38% and the second one achieves a coefficient of correlation R = 99.13%.

For this purpose, we are interested in predicting the time series of global solar radiation with less complexity. We are motivated by the research mentioned above, the optimal parameters for the proposed model will be examined and designated. Another strong point of the suggested technique is demonstrated by the considerable prediction data which reaches 217 days.

2 Methodology

2.1 LSTM Theory

The standard Recurrent Neural Networks (RNN) are challenged with the complications of the disappearance of gradient i.e. the vanishing gradient problem, which does not allow to change parameters (weight matrices) considering the previous entries [15]. However, with the appearance of LSTM type networks, answers have been attributed to these problematics [3]. An LSTM network, or recurrent short- and long-term memory network or even more explicitly, an intelligent memory cell with feedback connections. In practice, this type of network is frequently used, and it offers solutions to the difficulties of gradient disappearance.

In this section, we discuss the LSTM theory. Unlike the recurrent unit, which simply calculates a weighted sum of the input signals and applies a nonlinear function, each LSTM unit maintains one memory at a time. An LSTM unit consists of three gates, an input gate that can accept or reject the upgrade of the LSTM system, another output that checks if the state of the cell is transferred out of this system, and a for-gotten gate with internal memory providing the ability to perform a cell state reset. An LSTM unit thus can decide whether to keep or forget the existing memory via the introduced gates. Intuitively, if the LSTM unit detects an important feature of an input sequence at an early stage, it easily carries that information over a long distance, thereby capturing potential long-distance dependencies. A comparison between simple RNN and LSTM units is shown in Fig. 1.

LSTMs are the most interesting RNN models used for DL to solve time-series data prediction problems [9, 10].

The parameters of an LSTM network shown in Fig. 1 are calculated from the following formulas:

The LSTM network input gate i_t decides how much current information should be transmitted according to the following equation:

$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right) \tag{1}$$

The forget gate f_t decides which information to forget from the previous state as follows:

$$f_t = \sigma\left(x U^f + h_{t-1} W^f\right) \tag{2}$$

The output gate O_t defines the internal state information that should be transmitted using the following equation:

$$O_t = \sigma\left(x_t U^o + h_{t-1} W^o\right)$$

The cell state also called the internal memory C_t is updated in two steps. The first one consists of the information about the state of the cell

$$\widehat{C}_t = \tanh\left(x U^g + h_{t-1} W^g\right) \tag{3}$$

The U and W correspond to each of the weight matrices applied to the inputs of the associated operation. After calculating the cell state Ct, the internal memory C_t is updated from $\widehat{C}_t$ as follows:

$$C_t = \sigma\left(f^t * C_{t-1} + i_t * \widehat{C}_t\right) \tag{4}$$

where * is an element-by-element multiplication operation. The state of a cell C_t, is used to get the state of the next time step using the following formula:

$$h_t = \tanh(C_t) * O_t \tag{5}$$

where x_t, h_{t-1} and C_{t-1} are respectively the input to the hidden layer, the previous hidden state of the previous time step, and the previous state of the cell of the previous time step.
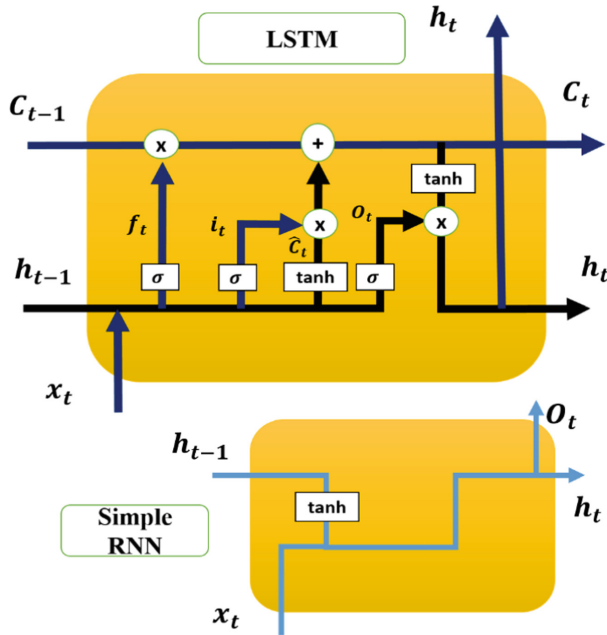
Fig. 1. Comparison between simple RNN and LSTM cells

Activation Function

There are several types of activation functions, the choice of the activation function depends on the application, whether or not it is necessary to have binary outputs. The mentioned activation functions are the Sigmoid (σ), and Tangent-hyperbolic (tanh) functions [14]:

Sigmoid function:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{6}$$

Hyperbolic tangent function;

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{7}$$

3 Database Collection and Preprocessing

The PROPRE.MA project initiated by Cadi Ayyad University, MOROCCO and sponsored by the Research Institute for Solar Energy and Renewable Energies (IRESEN), the data used during the design of this RSG prediction system are drawn from the meteorological station installed within this project in the Faculty of Sciences and Techniques of Er-Rachidia. The data collected between 01/01/2016 and 01/10/2020 are analyzed and processed to obtain information that may be useful for learning our predicting system.

3.1 Data Collection

The weather data measured by the sensors and instruments of the weather station are sent in real-time by wireless connection in an orderly manner to the weather station console. The Weather link data management software installed on our PC retrieves the recorded weather data at a frequency of 30 min, and allows us to visualize them in the form of bulletins or graphs over several periods of time as shown in Fig. 2.
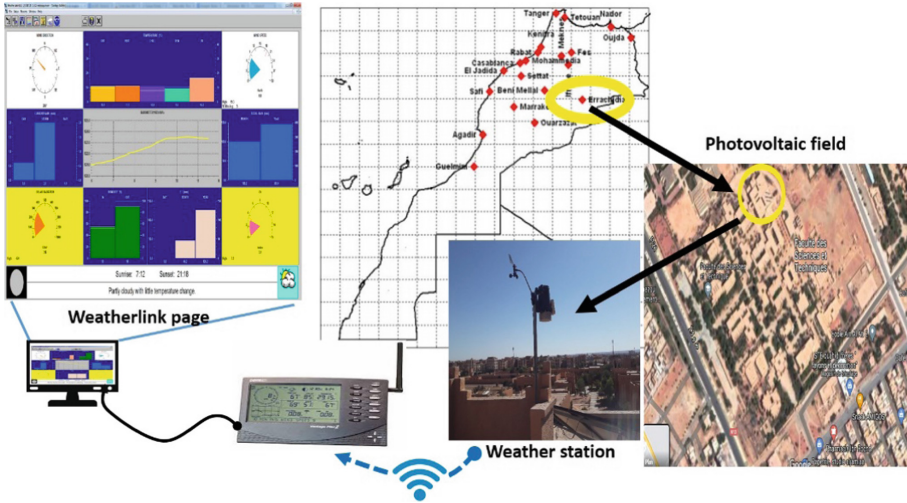

Fig. 2. Data collection process

3.2 Data Cleaning

Before analyzing the collected data and adjusting the prediction models, it is essential to process and purify its content. As a matter, missing data can mislead the learning process or be contributing to misinterpretation of the data. To avoid any such issues, we replace the missing values with the average of measured values of the previous and the following days at the same time step, provided that the total number of missing values on the same day does not exceed 12-time steps (6 h).

3.3 Selection of Predictors by a Correlation Study

Statistical studies are proposed to master the database which allows the analysis of the relationship between the elements that contribute to this database.

Figures Fig. 3 and Fig. 4 show Pearson's (R) and Phik's (φk) correlation indicators respectively [16]. It can be considered that the three variables: THWindex, Solar Energy, and UVindex can build the best predictors for the prediction of global solar radiation since the last three fields are strongly correlated and interactive with the field from GSR.
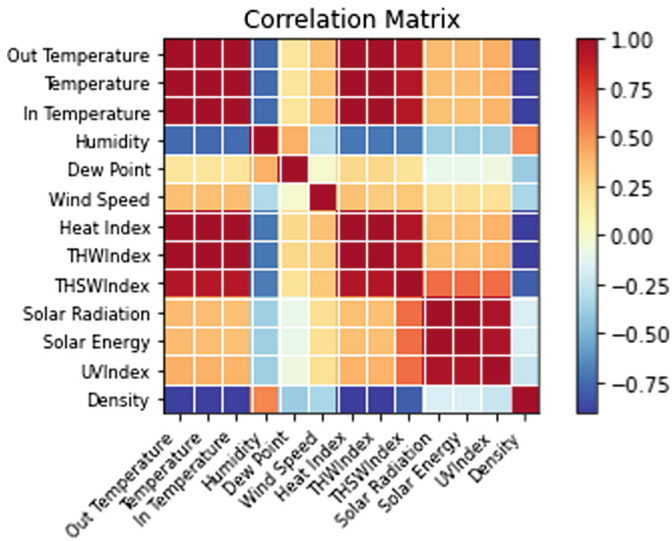
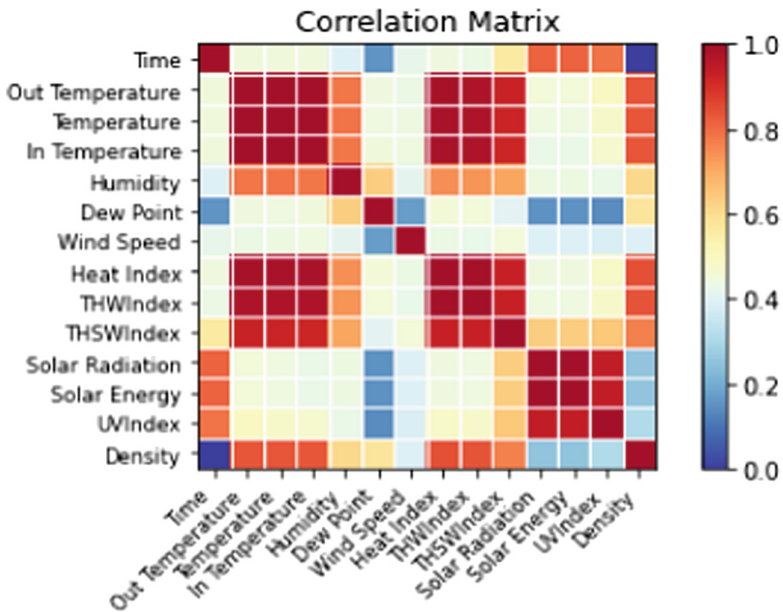**Fig. 3.** Pearson's correlation of the global solar radiation parameters.

Fig. 4. Phik (φk) correlation matrix of the global solar radiation parameters.

3.4   Model Architecture

The architecture of the model is, in general, a set of an input layer, which receives the collected data, an output layer, which produces information, calculated in the form of

predictions and a hidden layer used to connect the input layer, and the output layer using processing unit neurons [17]. However, recent established studies seek to improve the performance of the models by increasing the number of hidden layers and the number of nodes [10], the optimization of the model is done randomly at the beginning and then validated later by error estimators and the correlation coefficient [17, 18].

In this study, all the basic LSTM models are structured in the same form of the Adam learning algorithm [19] to train the LSTM models. Each model receives as input a sequence of successive input vectors of size $N - 1$ with N the number of base observations to predict the elements y(t) [20]. In the case of multivariate LSTM, these vectors include subsets of the vectors for the $P = 3$ parameters (the predictors selected in section A). However, oscillating windows of univariate LSTM contain the RSG parameter only. The samples of the resulting input data will be in the form of the 2D vectors of dimension $[n, P] = [200, 1]$ for monovariate LSTM, and in the form of the 3D vectors $[N, n, P] = [N, 200, 3]$ in case of multivariate LSTM. Each sample of the output data is a 1-dimensional vector.

3.5   Optimization Parameters Choice

Before proceeding to train each neural network model it is essential to normalize the database for scaling high contrast data for harmonized and easier computation afterwards. Indeed, we converted the characteristic values to a range of (0, 1) by the MinMaxScaler() transformer from the formula:

$$X_n = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{8}$$

where X_n is the normalized value, X is the actual value, X_{max} is the maximum value among all the values of the variable linked to X, and X_{min} is the minimum value among all the values of the variable associated with X.

Many tests have been performed to collect the appropriate hidden layers numbers and the number of neurons in each layer based on the values of Mean Squared Error (MSE) and Mean Absolute Error (MAE) calculated by the formulas mentioned in the research papers [7, 13]. MSE is practically essential in Machine Learning, it is often the error metric that is frequently used (loss function).

3.6   Classification Type

Deep neural networks such as LSTM can simulate single-variable problems and problems with multiple input variables. This research work aims to develop univariate LSTM models and multivariate LSTM models that can accurately predict short-term GSR at Er-Rachidia. Then, compare the results obtained by each approach. In the study conducted in this reference [21], the authors developed and compared two mentioned LSTM classes to predict global horizontal irradiation. However, they combined several variables to build multivariate models. Moreover, among these combinations, the model that

includes the temperature with the global horizontal irradiation as an input has the highest performance.

Monovariate LSTM Model

In this class, we consider that the GSR varies only in time. Consequently, the input tensor contains one feature as a variable. Nevertheless, we win in computation time and optimizing sensors. But, we lose the diversity of the variables that can advantageously characterize the problem of prediction to solve.

Multivariate LSTM Model

The time series forecasting method using multiple input variables for LSTM networks offers a great advantage in handling complex problems that are difficult to handle by conventional methods. In this class, we established a sense of the relationship of all variables with the GSR using the correlation matrix and then used the selected variables to develop a multivariate LSTM model. Therefore, the input tensor shapes for LSTM models this time is in the form (N, 200, 4).

### 3.7 Number of Hidden Layers Optimization

Two steps were defined to select the hidden number layers:

- **Step1**: following several tests, we fixed, while comparing the error obtained in each test, the number of neurons to 50 in the model with a single hidden layer.
- **Step2**: we test the model performance by increasing the number of hidden layers by one for each test. The results are summarized in Table 1 and Fig. 5.

Table 1. Hidden layers number variation versus calculated MSE and MAE.

Learning algorithm «Adam»			
Number of epochs	Number of hidden layers	MSE	MAE
30	1	0.0234	0.060
	2	0.0107	0.057
	3	**0.0022**	**0.0269**
	4	0.0023	0.0272
	5	0.0024	0.0271
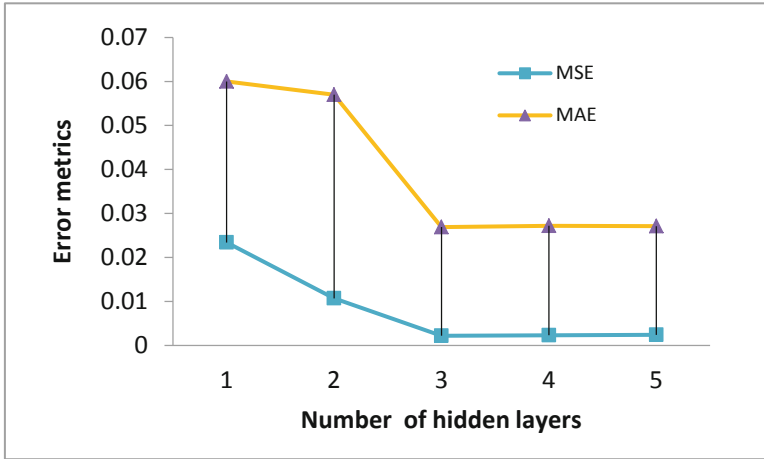
Fig. 5. Hidden layers number variation diagram versus the error metrics.

The model showed a convergence near a minimum error of MSE = 0.0022 and MAE = 0.0269 from the third layer, beyond this layer the model learning became almost stationary. Based on the above figure, the number of 3 hidden layers turned out to be the best one for the prediction model adopted in this study. In what follows, we will fix the number of layers in 3 and seek the number of optimal neurons for each layer.

3.8 Number of Neurons Optimization

By the same process as earlier, we carried out error tries, which narrowed the range of tests on the number of neurons in the three layers in (20, 60) neurons. Indeed, after several tests it has been observed that beyond 60 neurons the error committed by the model has become almost stable, similarly, below 20 neurons the performance of the model no longer improves. Consequently, the tests are carried out according to the combinations of Table 2 and the calculation results are grouped in Table 3.

The results show that on one hand, the increase in the number of neurons in the hidden layers beyond (50-50-50) neurons do not improve the learning rate of the model according to the values of the MSE, but also the MAE drops slightly, on the other hand, the reduction in the number of neurons in the three layers below (50-30-30) neurons increases the learning error of the model. Though, we scanned our tests between these extremities until the most adequate configuration (50-40-30) neurons which corresponds to an MSE = 0.0017 and MAE = 0.0180.

Table 2. Combination interval for neuron number in the hidden layer.

Hidden layer 1	Hidden layer 2	Hidden layer 3
(20-60)	(20-60)	(20-60)

Table 3. MSE and MAE values for different hidden neurons number.

Number of neurons in a hidden layer			MSE	MAE
Layer 1	Layer 2	Layer 3		
60	60	60	0,0022	0,0261
55	55	55	0,0022	0,0267
50	50	50	0,0022	0,0269
45	45	45	0,0023	0,0281
40	40	40	0,0023	0,0287
55	30	20	0,0028	0,0330
50	**40**	**30**	**0,0017**	**0,0180**
50	30	30	0,0025	0,0298
50	45	30	0,0029	0,0338
50	40	20	0,0028	0,0331

Therefore, an LSTM recursive neural network with deep machine learning by three hidden layers and 50, 40, and 30 neurons respectively in each layer was found to be the best implementation for predicting the semi-hourly global solar radiation at the Er-Rachidia site with a minimum of MSE. Suffice to say that the evolution of MSE and MAE during model training adheres to model convergence towards minimum error as shown in Fig. 6.
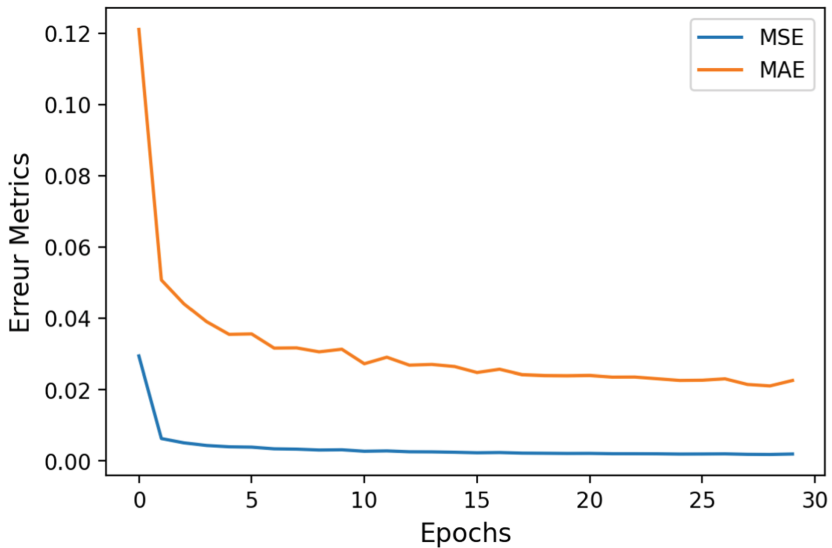
Fig. 6. MSE and MAE evolution versus epoch's number.

4 Realization of Predictions and Discussions

Two variants of the LSTM network (LSTM-multivariate and LSTM-monovariate) were experimented with within this part to predict RSG at the Er-Rachidia level as accurately as possible with learning rate (Learning rate = 0.001) and introduce one (Dropout = 0.1) after the first hidden layer to avoid the risk of overfitting [22]. The performance of the two varieties is evaluated according to the MSE and the MAE. The batch size is given as 100 and 48 for monovariate LSTM and multivariate LSTM respectively.

The raw data run through the model in the form of oscillating windows [10] as input and following the flowchart in Fig. 7 to estimate the output RSG. The first step in machine learning is data acquisition. Then, the acquired data was shared by various interested parties and summarized into useful information. The steps included in this process are data selection, purification and preprocessing, also data scaling. The data is divided into two disjoint subsets, a training set and another reserved for testing and evaluating the model according to different appropriate statistical indicators. The correlation coefficient R is chosen to assess the performance of the models in this study. R is calculated by the following formula:

$$R = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(x_i - \overline{y})^2}} \tag{9}$$

where, x_i is the observed value of semi-hourly RSG, y_i is the estimated value of semi-hourly RSG, $\overline{y}$ is the average of observations, and N is the total number of observations.
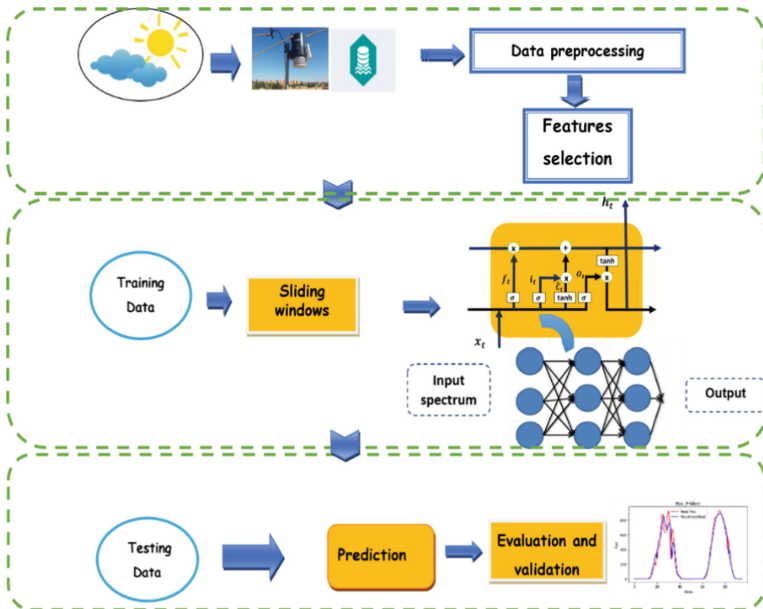

Fig. 7. DATA processing diagram and learning process.

4.1 First Experience

Over the period from 01/10/2016 until 13/03/2017 two simulations were carried out by the two methods monovariate LSTM and multivariate LSTM by the data shown in figure Fig. 8. Over this period, the two ANNs are tested by the totality of the data (training data and test data): the training data to control and visualize the training phase and the test data are processed in installments: 20%, 30%, and 40% of the database to study the stability of the models.

Based on the coefficient of determination R and the error metrics, both methods showed high accuracy, with prediction quality stability up to 40% as shown in Table 4. Emphasizing that the multivariate LSTM model converges faster than monovariate LSTM. A comparison between the estimated values in different forms and the measured values for the two methods is presented in Fig. 9 and Fig. 10.
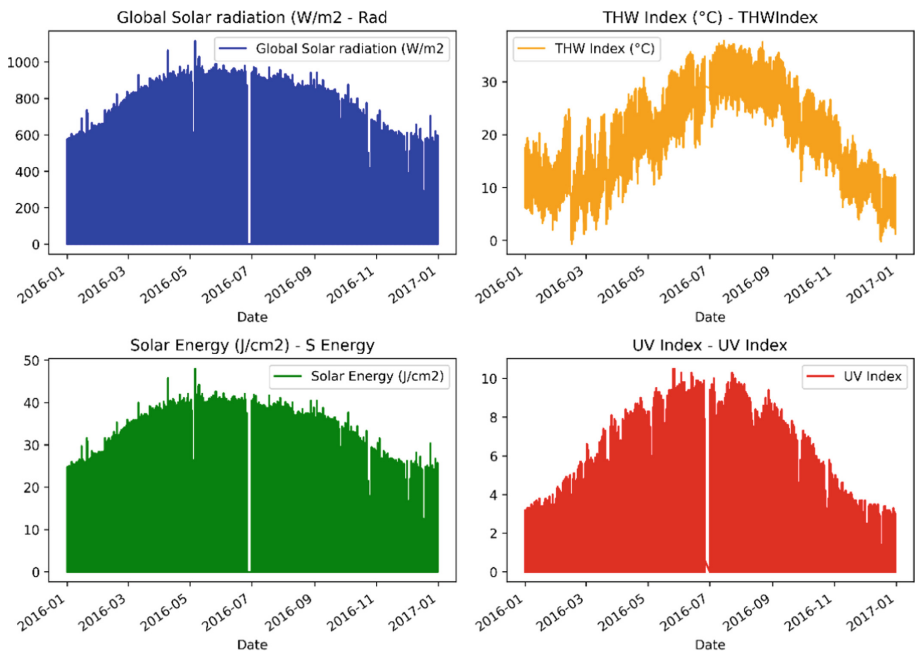

Fig. 8. Selected parameters for building ANN models for 2016

The four parameters shown in Fig. 8 are very important. Indeed, the actors of alternative energies especially solar energy need to track the amount of solar radiation for solar energy production purposes. However, the THWindex field uses humidity, temperature and wind measurements to calculate an apparent temperature, and the UV index assigns a number between 0 and 16 to the current UV intensity, which will help to analyze changing UV radiation levels.
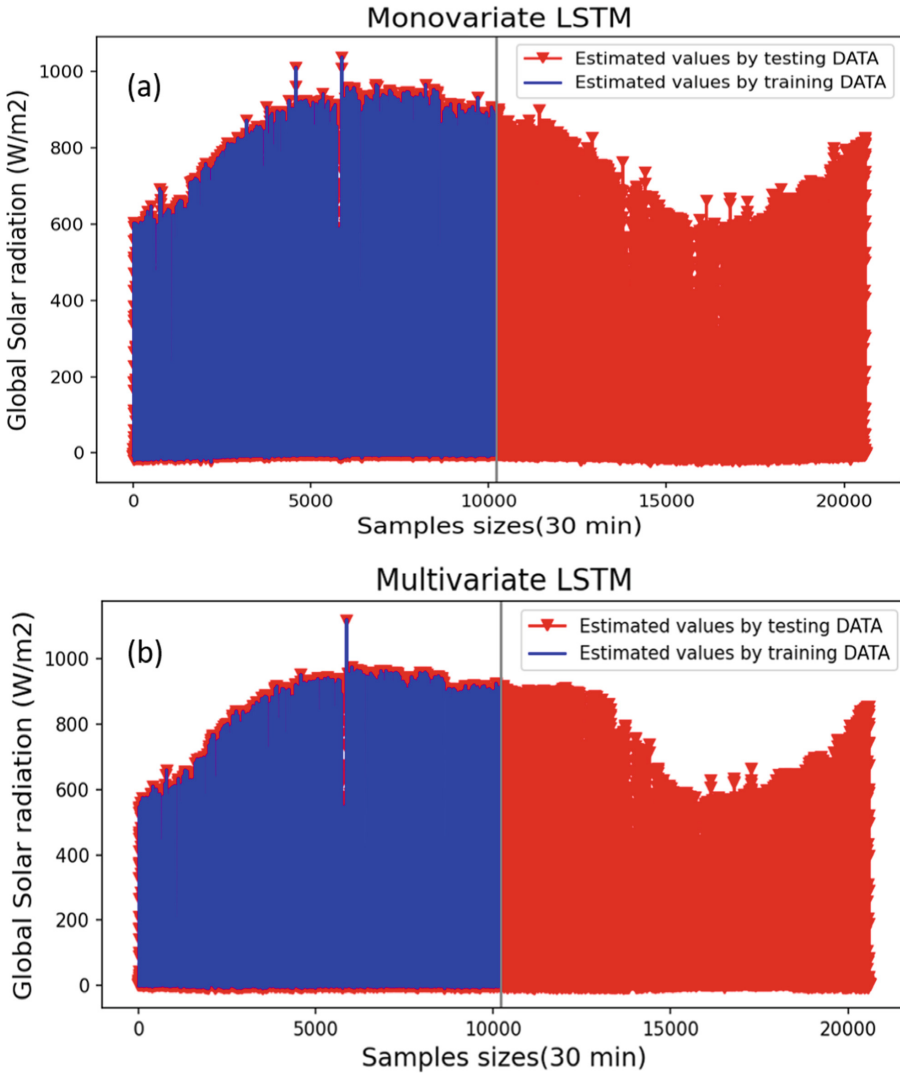
Fig. 9. Estimated GSR in the training and testing phases—(a) by univariate LSTM, (b) by multivariate LSTM.

Considering the convergence of the loss function as a function of epochs and the values of the significant correlation coefficients obtained at the end of the tests (Table 4), can guarantee that the learning rate and the other optimization parameters are selected appropriately. The maximum correlation coefficient obtained for this first simulation is R = 98.97% (Fig. 12).
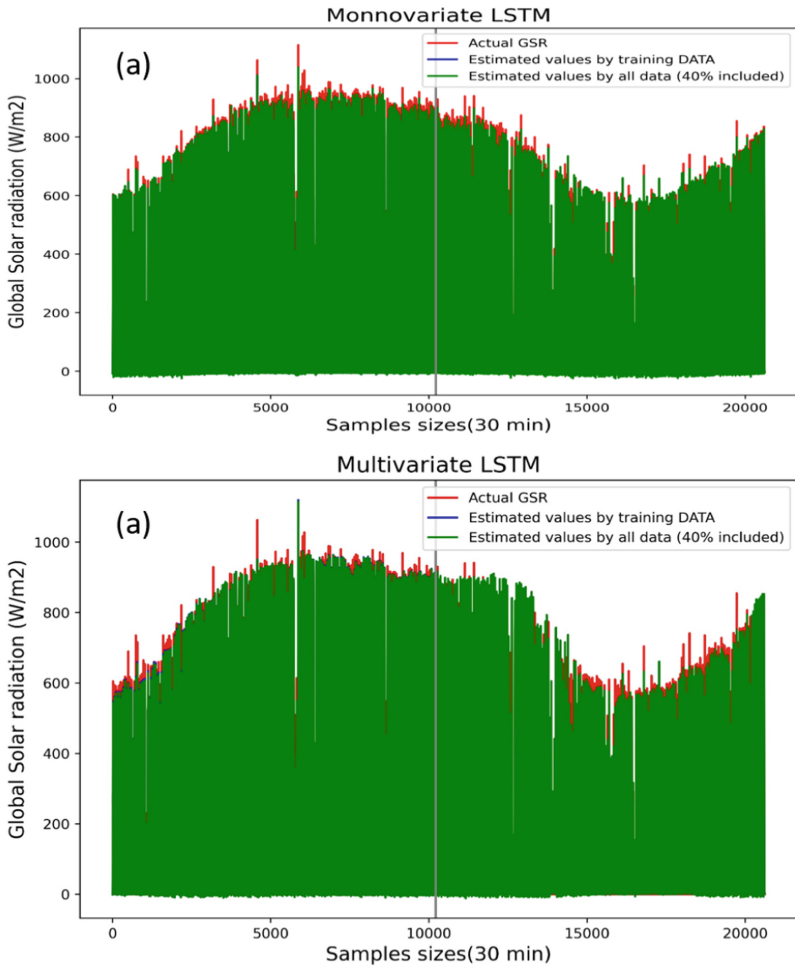
Fig. 10. Estimated and measured GSR in the training and testing phase—(a) by univariate LSTM, (b) by multivariate LSTM.

4.2 Error Metrics Synthesis of the Developed Models

To prove the stability of our models, we have calculated the MSE, MAE, and R for the two approaches developed in this work during the training and the testing phase.

Table 4. Model performance in terms of MSE, MAE, and R indicators during the training and testing phases.

	Test size	Train			Test		
		MSE	MAE	R	MSE	MAE	R
LSTM_MONO	20%	0.0014	0.0152	0.9908	0.0010	0.0125	0.9800
	30%	0.0018	0.0221	0.9895	0.0010	0.0160	0.9890
	40%	0.0017	0.0219	0.9897	0.0019	0.0222	0.9886
LSTM_MULTI	20%	0.0010	0.0153	0.9929	0.0010	0.0139	0.9803
	30%	0.0011	0.0152	0.9865	0.0011	0.0153	0.9897
	40%	0.0011	0.0162	0.9932	0.0017	0.0207	0.9827

Based on the mean square error and the mean absolute error values given in Table 4, the scores obtained for MSE and MAE indicate that the two models examined showed great stability, especially the LSTM_MULTI model which committed an error of MSE = 0.0017W/m^2 at 40% test size compared to LSTM_MONO model which reached an error of MSE = 0.0019W/m^2. The results obtained in the training and testing phase mean that our model is correctly calibrated, also it's able to overcome learning problems such as vanishing gradient.

Validation Loss Function

For the validation task, we choose an MSE loss function to validate the predictive model. Therefore, a graph is also created in Fig. 11 showing two-line plots, the top showing the MSE for the training dataset over epochs, and the bottom showing the MSE for the testing dataset over epochs. The graph shows that the training process has indeed converged. But we can see that MSE shows an irregular evolution in the test phase, although the dynamics of MSE do not seem to be too affected compared to the training phase, given that the nature of GSR sometimes shows large outliers.
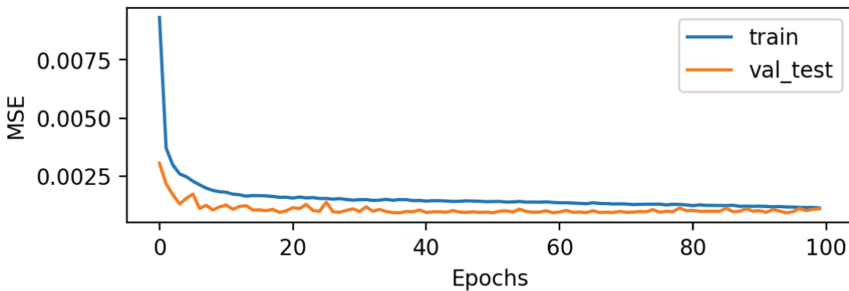
Fig. 11. Evolution of the loss function versus the number of epochs for the training and testing phase.

Fig. 12. Scattering diagram for testing data obtained in the first experience.

4.3 Second Experience

The second experiment is performed by monovariate LSTM based on the RSG time series between 01/10/2016 and 01/10/2020 (Fig. 13). However, this period is tested by 20% of all data chosen between 2019-08-05 and 2020-10-01 of the GSR time series. Model convergence is established at 30 epochs. Figure Fig. 14 shows that predicted SRG versus actual GSR wish is very accurate. The predictions of the RSG time series show high accuracy and a strong correlation between the measured values and the estimated values with a correlation coefficient of R = 98.83%, (Fig. 15).

The performance demonstrated by our model in the two experiments examined shows that it could be applied to optimize the operation of solar energy-based electricity services by predicting solar radiation. Several studies have recently been carried out internationally to simulate solar radiation [14], the technique proposed in this study marks a high accuracy in terms of correlation coefficient compared to the majority of studies shown in Table 5.

The main common point between the studies that have marked a great correlation between the estimated and the real values of GSR based on the coefficient of determination R is the integration of dynamic and recurrent ANN models such as LSTM models that can handle the complexity of prediction problems. The prime keys to success in this work are: The data analysis examined based on the correlation study which allows selecting the most strongly correlated features that contribute to learning models with less error and the adoption of a dynamic neural network LSTM to deal with the dynamics of GSR which is influenced by the weather conditions changes.
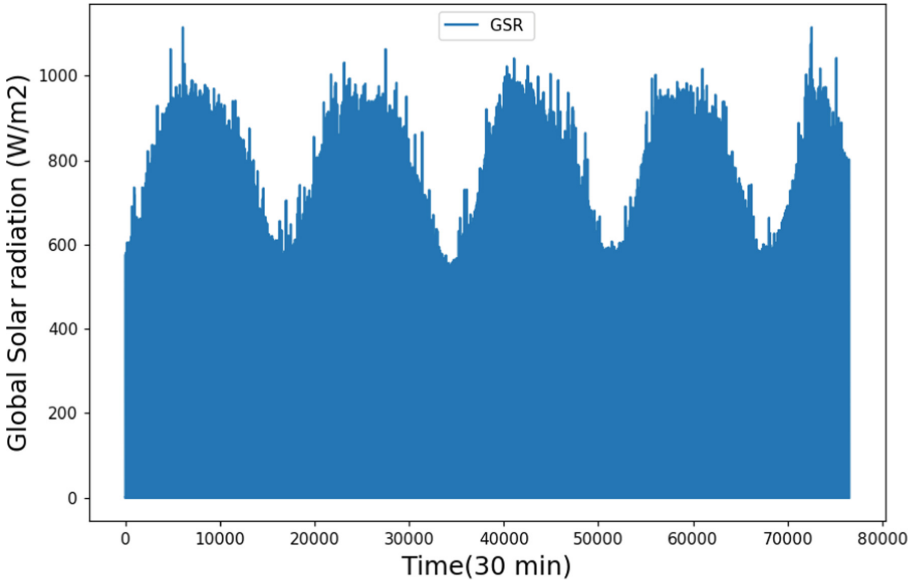
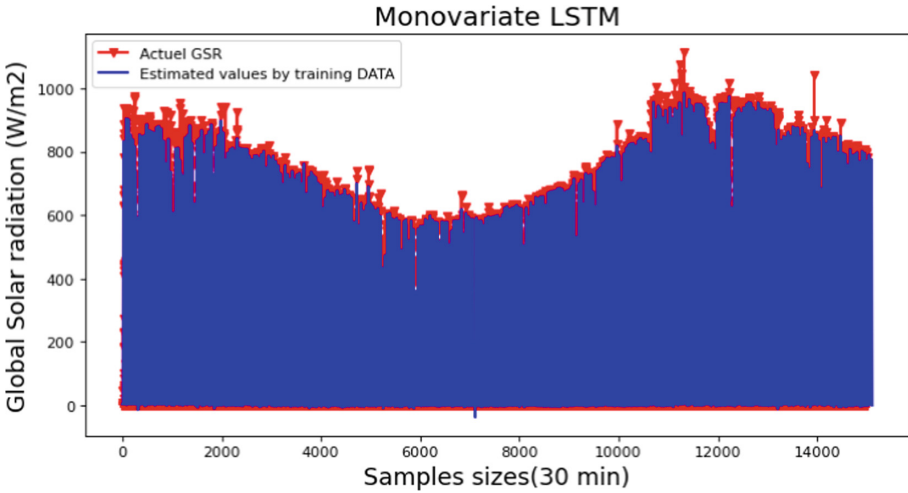**Fig. 13.** Global solar radiation time series measurement between 2016 and 2020.


**Fig. 14.** Estimated GSR and actual GSR for the second experience

For a prediction dedicated to the short applications [30] of the GSR time series by our model, a zoom on 2 days (Fig. 16) and 10 days (Fig. 17) of prediction. The figures show the difference between real GSR and GSR predicted by our approach. For this purpose, the developed prediction model allows the power grid managers to optimize the operation of energy systems, and to generate simulated data of the GSR when missing values of this parameter are detected.
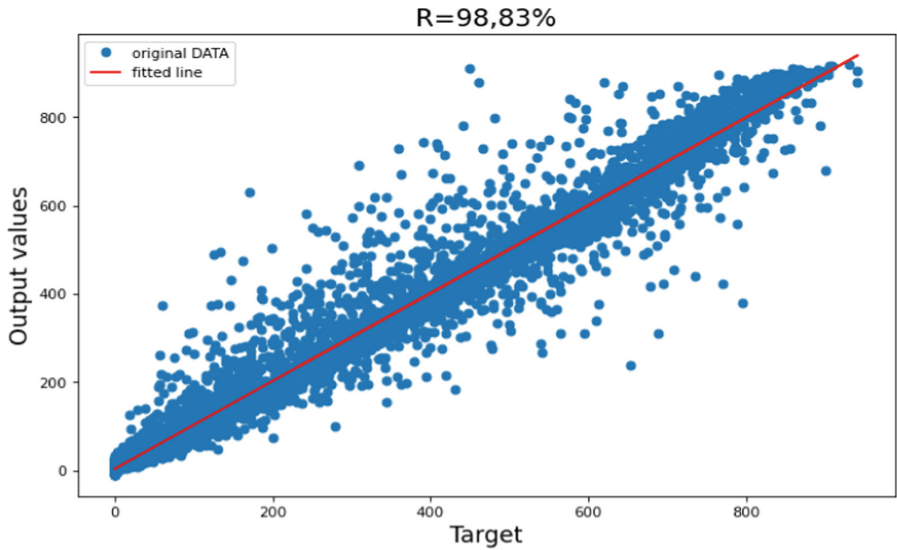
Fig. 15. Scattering diagram for testing data obtained in the second experience.

Table 5. Some recent studies on the prediction of GSR time series and their correlation coefficient.

Reference	Applied technique	R (%)
[13] (2020)	ANN	99.13
[23] (2018)	SVR-PLS	93.3
[24] (2018)	Empirical model	96.7
[24] (2018)	Empirical model	82.5
[25] (2019)	DENFIS	94.2
[25] (2019)	DENFIS	97.9
[26] (2020)	PVGIS	97.1
[26] (2020)	HILIOSAT	97.2
[26] (2020)	HILIOSAT	97.6
[27] (2020)	ANN	98.49
[28] (2020)	ANN	97.93
[29] (2021)	ANN (SR and Wind)	96.93–99.97
Present work (2022)	**Proposed model (ANN/LSTM)**	**98.97**
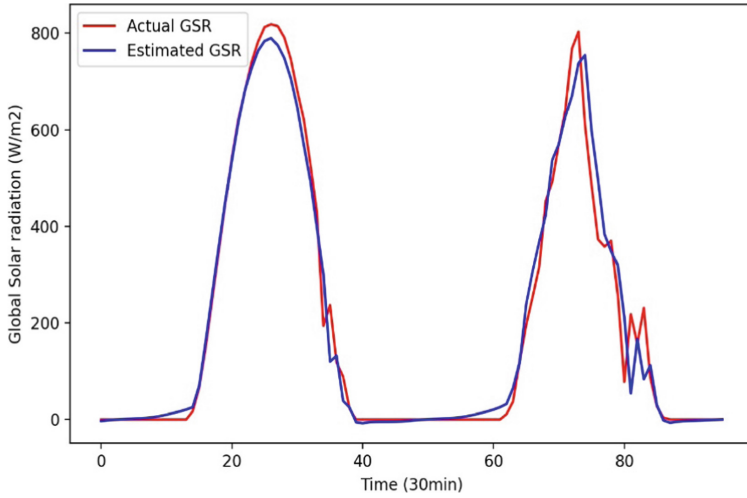
Fig. 16. Predicted GSR and actual GSR for two days.
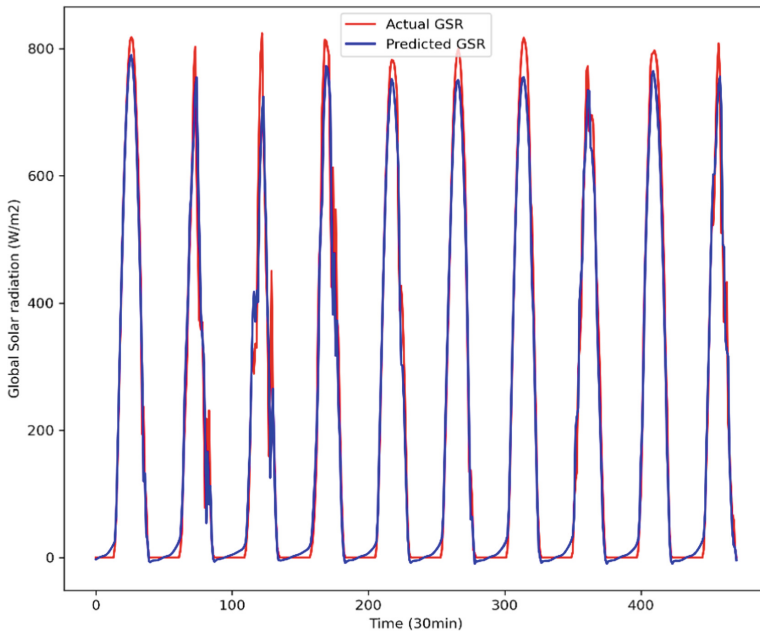
Fig. 17. Predicted GSR and actual GSR for ten days.

5   Conclusion

The study we conducted in this article was meant to improve the solar irradiation pre-
diction tools using Deep Learning methods. The aim is to innovate in terms of new
technologies and then invest in smart systems to respond to a crucial need which is the

integration of solar energy into the grid, in a fast dynamic and appropriate way. After-wards studying statistically, the database and selecting strongly correlated fields to the RSG parameter, we adopted a model based on LSTM neural networks to identify the half-hourly RSG changes at Er-Rachidia, Morocco. We also controlled the dynamism over time of this parameter during model training with a well-optimized calibration. The pertinent results obtained show the robustness of our method. We reached a strong Pearson correlation at R = 98.97% given by the multivariate LSTM model.

References

1. Tan, K.M., Babo, T.S., Ramachandaramurthy, V.K., Kasinathan, P., Solanki, S.G., Raveen-dran, S.K.: Empowering smart grid: a comprehensive review of energy storage technology and application with renewable energy integration. J. Energy Storage **39**, 102591 (2021). https://doi.org/10.1016/j.est.2021.102591
2. Le Ministère de l'Energie, des Mines et de l'Environnement. https://www.mem.gov.ma/Pages/secteur.aspx?e=2. Accessed 5 Sept 2021
3. Cheng, P.M., Tejura, T.K., Tran, K.N., Whang, G.: Detection of high-grade small bowel obstruction on conventional radiography with convolutional neural networks. Abdom. Radiol. **43**(5), 1120–1127 (2017). https://doi.org/10.1007/s00261-017-1294-1
4. Schmidhuber, J.: Deep Learning in neural networks: an overview. Neural Netw. **61**, 85–117 (2015). https://doi.org/10.1016/j.neunet.2014.09.003
5. Goldberg, Y.: A primer on neural network models for natural language processing. J. Artif. Intell. Res. **57**, 345–420 (2016). https://doi.org/10.1613/JAIR.4992
6. Voyant, C., et al.: Machine learning methods for solar radiation forecasting: a review. Renewable Energy **105**, 569–582 (2017). https://doi.org/10.1016/j.renene.2016.12.095
7. Huang, X., et al.: Hybrid deep neural model for hourly solar irradiance forecasting. Renewable Energy **171**, 1041–1060 (2021). https://doi.org/10.1016/j.renene.2021.02.161
8. Citakoglu, H.: Comparison of artificial intelligence techniques via empirical equations for prediction of solar radiation. Comput. Electron. Agric. **118**, 28–37 (2015). https://doi.org/10.1016/j.compag.2015.08.020
9. Al-Hajj, R., Assi, A., Fouad, M., Mabrouk, E.: A hybrid lstm-based genetic programming approach for short-term prediction of global solar radiation using weather data. Processes **9**(7), 1187 (2021). https://doi.org/10.3390/pr9071187
10. Aslam, M., Lee, J.M., Kim, H.S., Lee, S.J., Hong, S.: Deep learning models for long-term solar radiation forecasting considering microgrid installation: a comparative study. Energies **13**(1), 147 (2019). https://doi.org/10.3390/en13010147
11. Qing, X., Niu, Y.: Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. Energy **148**, 461–468 (2018). https://doi.org/10.1016/j.energy.2018.01.177
12. Alzahrani, A., Shamsi, P., Dagli, C., Ferdowsi, M.: Solar irradiance forecasting using deep neural networks. Procedia Comput. Sci. **114**, 304–313 (2017). https://doi.org/10.1016/j.procs.2017.09.045
13. Jallal, M.A., El Yassini, A., Chabaa, S., Zeroual, A., Ibnyaich, S.: AI data driven approach-based endogenous inputs for global solar radiation forecasting. Ing. des Syst. d'Inf. **25**(1), 27–34 (2020). https://doi.org/10.18280/isi.250104
14. Jallal, M.A., El Yassini, A., Chabaa, S., Zeroual, A., Ibnyaich, S.: A deep learning algorithm for solar radiation time series forecasting: a case study of el kelaa des sraghna city. Rev. d'Intelligence Artif. **34**(5), 563–569 (2020). https://doi.org/10.18280/ria.340505
15. Bengio, Y., Simard, P., Frasconi, P.: Learning long-term dependencies with gradient descent is difficult. IEEE Trans. Neural Netw. **5**(2), 157–166 (1994)

16. Phik (φk). https://towardsdatascience.com/phik-k-get-familiar-with-the-latest-correlation-coefficient-9ba0032b37e7. Accessed 28 Jan 2022
17. Premalatha, N., Valan Arasu, A.: Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. J. Appl. Res. Technol. **14**(3), 206–214 (2016). https://doi.org/10.1016/j.jart.2016.05.001
18. Apaydin, H., Feizi, H., Sattari, M.T., Colak, M.S., Shamshirband, S., Chau, K.-W.: Comparative analysis of recurrent neural network architectures for reservoir inflow forecasting. Water **12**(5), 1500 (2020). https://doi.org/10.3390/W12051500
19. Adam. https://towardsdatascience.com/adam-latest-trends-in-deep-learning-optimization-6be9a291375c. Accessed 3 Oct 2021
20. Cui, Z., Ke, R., Wang, Y.: Deep Bidirectional and Unidirectional LSTM Recurrent Neural Network for Network-Wide Traffic Speed Prediction. IEEE (2018)
21. Mandal, A.K., Sen, R., Goswami, S., Chakraborty, B.: Comparative study of univariate and multivariate long short-term memory for very short-term forecasting of global horizontal irradiance. Symmetry **13**(8), 1544 (2021). https://doi.org/10.3390/sym13081544
22. Mele, B., Altarelli, G.: Lepton spectra as a measure of b quark polarization at LEP. Phys. Lett. B **299**(3–4), 345–350 (1993). https://doi.org/10.1016/0370-2693(93)90272-J
23. García-Hinde, O., et al.: Evaluation of dimensionality reduction methods applied to numerical weather models for solar radiation forecasting. Eng. Appl. Artif. Intell. **69**, 157–167 (2018). https://doi.org/10.1016/J.ENGAPPAI.2017.12.003
24. Bailek, N., et al.: A new empirical model for forecasting the diffuse solar radiation over Sahara in the Algerian Big South. Renew. Energy **117**, 530–537 (2018). https://doi.org/10.1016/j.renene.2017.10.081
25. Kisi, O., Heddam, S., Yaseen, Z.M.: The implementation of univariable scheme-based air temperature for solar radiation prediction: new development of dynamic evolving neural-fuzzy inference system model. Appl. Energy **241**, 184–195 (2019). https://doi.org/10.1016/j.apenergy.2019.03.089
26. Ener Rusen, S., Konuralp, A.: Quality control of diffuse solar radiation component with satellite-based estimation methods. Renew. Energy **145**, 1772–1779 (2020). https://doi.org/10.1016/J.RENENE.2019.07.085
27. Guti, P.A., et al.: Evolutionary artificial neural networks for accurate solar radiation prediction. Energy **210**, 118374 (2020). https://doi.org/10.1016/j.energy.2020.118374
28. Üstün, İ, Üneş, F., Mert, İ, Karakuş, C.: A comparative study of estimating solar radiation using machine learning approaches: DL, SMGRT, and ANFIS. Energy Sour. Part A **44**(4), 10322–10345 (2020). https://doi.org/10.1080/15567036.2020.1781301
29. Shboul, B., et al.: Original article A new ANN model for hourly solar radiation and wind speed prediction: a case study over the north & south of the Arabian Peninsula. Sustain. Energy Technol. Assessments **46**, 101248 (2021). https://doi.org/10.1016/j.seta.2021.101248
30. Mishra, M., Byomakesh, P., Nayak, J., Naik, B., Kumar, S.: Deep learning and wavelet transform integrated approach for short-term solar PV power prediction. Measurement **166**, 108250 (2020). https://doi.org/10.1016/j.measurement.2020.108250