Long-Term Solar Irradiance Forecasting Using Multilinear Predictors

Abdulamjeed Sulaiman[1], Farhad E. Mahmood[2], and Sayf A. Majeed[3]*
[1] Department of optics technique, Alnoor University College, Bartella, Iraq
[2] Department of Electrical Engineering, College of Engineering, University of Mosul, Mosul, Iraq
[3] Computer Technologies Engineering Department, Al-Hadba University College, Mosul, Iraq
Email: almajeed.ahmed@alnoor.edu.iq; Farhad.m@uomosul.edu.iq

***Abstract*—As the demand for crude oil is increasing every day, prices and pollution are both increasing in return, which has harmful effects on the environment. Thus, more attempts are being made to develop clean energy to rescue the planet and provide humanity with a cleaner energy source. As the renewable energy sector grows, new issues and challenges have emerged. Instability in electricity production from wind turbines, solar power plants, and dams creates challenges for energy transmission and storage systems. In order to achieve a more reliable and effective energy system, machine learning techniques have been used to forecast energy changes. Predicting the sun's Global Horizontal Irradiance (GHI) is one of the machine learning applications used in that sector. In this paper, many machine learning methods have been utilized, such as linear regression and long-short term memory (LSTM) methods to have long term GHI forecasting. Moreover, the significance of this paper is located in the way of prediction of the GHI irradiance prediction by using different levels of the linear regressors to find the best regressor level that provides the minimum error for the testing set based on cross-validation. Results showed that the regressor method provides a lower error compared with a single vanilla LSTM system for a shorter time computationally.**

***Index Terms*—Cross-validation, linear regression, LSTM, renewable energy, Recurrent Neural Network (RNN), Solar irradiance forecasting**

I. Introduction

To save the environment from the pollution of crude oil energy, many calls are being made to develop clean energy, save the Earth, and provide a better source of energy for humanity. Renewable energy can be a solution to save the Earth. Nowadays, the importance of renewable energy sources is increasing with the massive expansion of energy consumption in the world and with the decrease of non-renewable energy resources such as crude oil, coal, and natural gas. In addition, the use of non-renewable energy resources leads to high pollution and causes radical climatic changes [1]. As part of efforts to make our planet less polluted, much of the research in the past decade has focused on renewable energy sources, especially wind and solar energy sources.

According to the International Energy Agency (IEA), the global renewable energy capacity is expected to grow by more than 1 Terawatt, a growth of 46% by 2023 [2]. Solar and wind energy like other types of natural energy resources depend on the amount of Global Horizontal Irradiance (GHI) of solar power plants and the speed and wind direction of wind turbines and other climate phenomena, due to these effects, the output power fluctuates over time, that is why calculations are needed for obtaining suitable power storage units (batteries), determining the way of integrating the output power into the network and what control-units are needed so the highest possible generated power is handled [3].

With the uncertainty of the amount of GHI and of good energy adaptation to the power grid, it is necessary to calculate the GHI in the long run, depending on historical data that is measured from the ground or from the satellite. Many methods have been developed to predict solar radiation and predictions are also used for solar power plants to detect all possible results can be used to overcome the problems of generation and transmission, physical, empirical and statistical methods have been introduced and improved for these reasons.

The importance of this paper is reflected in the way of prediction of the GHI irradiance prediction using different levels of the regressors to find the best regressor level that provides the minimum error for the testing set. This regressor method provides a lower error than the LSTM for a shorter time computationally.

Even the partial linear regression (PLR), is not a new method, but to the best of our knowledge, we are the first to utilize this method in long-term solar irradiance prediction, which provides open doors to develop solar irradiance prediction easily and, yet, accurately method. Furthermore, we utilized the cross-validation method to test the prediction step size that provides the optimum step that fulfills the minimum error for both training and testing.

The present paper is divided into five sections. In Section II, we introduce the regression methods used for solar irradiance forecasting. In section III, we introduce performance analysis that can be used with our work. In Section IV, the implemented models and their results are explained. Finally, in Section V, the results are discussed further with regard to future work.

Manuscript received September 6, 2022; revised November 7, 2022; accepted January 6, 2023.
*Corresponding author: Sayf A. Majeed (email: sayf@hcu.edu.iq).

doi: 10.18178/ijeetc.12.2.134-141

II. Solar Irradiance Forecasting Method

This section will discuss the prediction methods that are used for solar irradiance forecasting, starting from the oldest and simplest and moving toward the most recent techniques that are used for that purpose.

A. *Physical Model*

The first attempts to solve transmission and generation problems were the physical methods, which are traditional prediction methods that are also discussed in the literature [4–6]. These methods depend on the tilt angle of the PV panel, the angle of the sun, PV properties, and other weather conditions and their effects on the absorption of sunlight in photovoltaic cells [7]. The physical methods need a large amount of data to have accurate results. Other predictor types appear to have more accurate results in the literature.

B. *Empirical Model*

One of the most important methods is called the sunshine method. This method is multiple-parameter method that depend on two empirical values: the monthly average of the GHI and the monthly time of sunshine [8]. As shown in Eq. (1)

$$\frac{GHI_{average}}{H_0} = a + b\frac{S}{S_0} \tag{1}$$

where S is the monthly average of sunshine, S_0 is the day length, a and b are location empirical values, $GHI_{average}$ is the average global horizontal irradiance predicted monthly, H_0 is the monthly real mean solar radiation [3].

C. *The Statistical Models*

To have more accurate predictions, the literature puts more effort into statistical methods. These methods depend on the previous values of GHI to predict future solar irradiance and its absorbable power [7]. Statistical methods can be categorized into time-series methods and machine learning methods.

1) *The Time-Series Methods*

By using old data and some formulas, we can have an idea that describes how solar irradiance will be in the future. These methods are not very accurate due to abnormal weather changes, and we cannot obtain all the conditions in our mathematical formulas [9].

The Auto-Regressive Integrated Moving Average (ARIMA) is one of the most significant of these methods. Derived from the ARIMA method, which depends on two parts (auto-regression based on the moving average), these methods are more flexible and accurate than other time series methods [10–12].

2) *Computational Intelligence*

Depending on statistics and to have more accurate results, automated methods have been developed, which are called computational intelligence, or “Machine Learning (ML) methods.” As we can see from Fig. 1, computational intelligence methods are the second branch of statistical methods. Moreover, they can also be used to improve both the generating and transferring of power from solar energy plants [9].

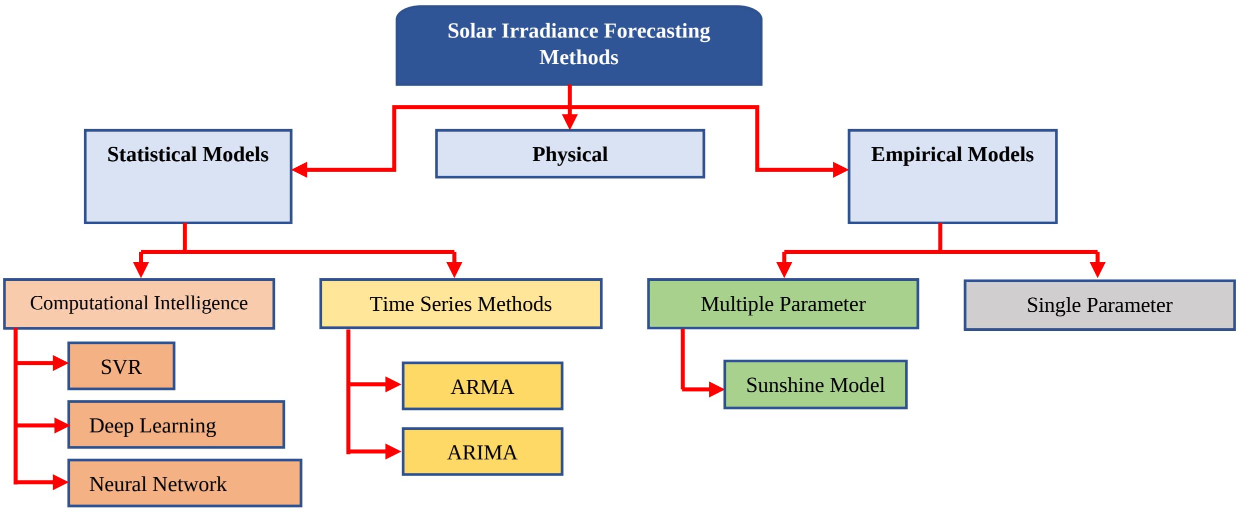

Fig. 1. Solar irradiance forecasting methods.

Numerous computational intelligence techniques for both short- and long-term predictions have emerged. The most fundamental is linear regression, which is usually utilized for short-term prediction [13, 14]. In addition, multiple linear regression is also used to predict solar irradiance where it has been used to increase the number of effect parameters [12, 15]. Furthermore, extreme learning machine (ELM) is an advancement over multiple linear regression [16, 17]. For more accurate data, a support vector machine (SVM) was developed, which predicts short-term solar irradiance via data mapping [18]. Although it yields improved results, it still has limitations, particularly in long-term prediction; hence, the researchers investigated more complex nonlinear methods.

For improved outcomes (long- and short-term prediction), this approach grew more complex and difficult to apply by using all parameters from prior irradiance states.

On the other hand, feed-forward neural networks (FNN) are the simplest approach that resembles the human brain and nervous system. Typically, this kind has three data-processing layers: the input layer, the output layer, and hidden layers [7, 19, 20]. As an evolution of the FNN, the Recurrent Neural Network (RNN) emerged. To produce fewer errors, the RNN approach uses sequential data with prior states and errors saved internally in iterators [7, 21, 18]. In addition, the Deep Recurrent Neural Network (DRNN), which has several hidden layers, generates more precise results than its counterparts. However, it requires more computing effort than others [7, 22, 23].

There are many types of DRNN, which all have poor long-term memorization traceability when using short-memory connections, and this makes the gradient either explode or disappear. This resulted in the creation of the LSTM approach, which addresses these issues [7, 9, 24], [25]. LSTM is one of the deep learning methods that was designed to avoid long-term dependencies by keeping the information in a separate control unit. It consists of two components: a hidden layer memory cell and a working cell, allowing it to circumvent the gradient exploding (vanishing) issue [26]. The LSTM unit contains three gates instead of two. In addition to the input and output gates, there is also a "forget" gate, which eliminates noise and other unwanted signals from the cell for improved output [16, 26]. Eq. (2) describes the LSTM unit:

$$\begin{aligned} f_t &= \sigma_g(W_f \times x_t + U_f \times h_{t-1} + b_f) \\ i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \\ o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \\ \check{C}_t &= \sigma_g(W_c x_t + U_c h_{t-1} + b_c) \\ C_t &= f_t c_{t-1} + i_t \check{C}_t) \\ h_t &= o_t \sigma_t(C_t) \end{aligned} \tag{2}$$

where f_t is the forget gate; i_t is the input gate; O_t is the output of h_t, which is the output of the hidden gate; h_{t-1} is the hidden state output of the previous LSTM block; C_t represents the cell state; $\check{C}_t$ represents the candidates for cell state at timestamp t; σ is the sigmoid function; x_t is the input of the current timestamp. W_f, W_i, W_o, W_c, U_f, U_i, U_o, and U_c are the weights of the respective gates, and b_f, b_i, b_o, and b_c are the biases for the respective gates. Weights and bases remain constant throughout all steps and do not change from one time step to the next [27-30], LSTM neural network architecture is shown in Fig. 2.

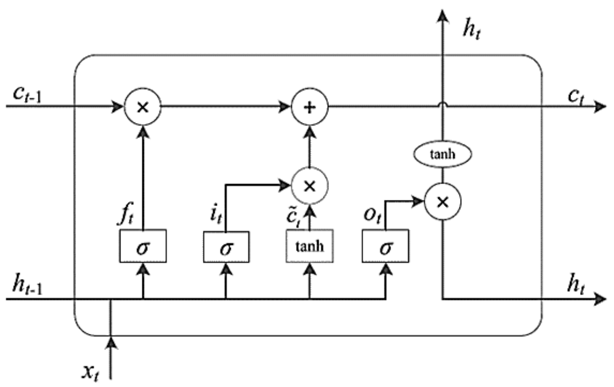

Fig. 2. LSTM structure network.

Instead of all this complex method, and by returning to linear regression which has the main limitation of poor outcomes for long-term predictions of the natural phenomenon, some sources used multiple linear regression to overcome the limitations of linear regression to predict long-term states [31], In this paper, it is used multiple linear regression to predict solar radiation for long-term forecasting which is uncommon in previous studies, predicting solar irradiance for short-terms (hours) is very easy by using any machine learning approach, whereas predicting solar for long-term (days, months or year "as done here"), required more efforts [32]. Moreover, the decision to use multiple linear regression for solar irradiance is because the solar irradiance is changing slowly during the day, which makes long-term prediction more important than short-term and mid-term prediction in this field.

III. Performance Analysis

Due to the nonideal nature of the GHI, errors in solar irradiation predictions occur; to eliminate these errors, the system must be analyzed and the faults that have been committed must be examined. One of the following assessment types may be used.

Graphically: predicted data can be compared with measured data so we can visually recognize errors and differences. Many kinds of literature use time series plots, scatter plots, or Receiver Operating Characteristic (ROC) curves [15].

Statistically: the graphical "traditional" methods are not enough to test the results. More accurate and more detailed results are needed. The statistical method can give us what we need. Many statistical formulas are used for our objectives. The most commonly used method is, the root mean square error (RMSE), which is also used for analyzing the systems. RMSE can be calculated using Eq. (3):

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(y_i - \hat{y}_i\right)^2} \tag{3}$$

Database description: the database used in this work contains many weather details besides the date, rainfall, snowfall, wind direction, and other information. This database is obtained from the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), which is being developed to be a "milestone" for future integrated Earth system analysis (IESA) that is under construction. MERRA-2 has many quality enhancements and decreases in erroneous trends and system-related faults, the database contains data of 1490 days with solar irradiance took for every 24 hours, at the time of maximum daily irradiance, these 1490 records are divided into the training set and testing set [33]. This sort of technology provides GHI, rainfall, and snowfall probabilities together with temperature, relative humidity, and pressure at a distance of 2 m from the land, while wind speed is measured at a distance of 10 m, the data of this technology is widely used in literature and for many fields [34–36]. The time period covered by the dataset begins on December 1, 2016, and ends on December 31, 2020, for a site located at a latitude of 36,356 (positive means north) and a longitude of 43,125 (positive means east) (Mosul, Iraq).

In this study, solar irradiance is predicted using only the days of the year, which is more important for us due to the slowness of irradiance changing during the day and other parameters like wind speed and rainfall are neglected for now

IV. Models Implementation

The MERRA-2 was analyzed using Python programming in order to generate one-year forecasts. Four of the five implemented models rely on linear prediction, while the fifth model is the LSTM model. For each model, the data is divided into two parts: two years for training data and one year for testing data. The outliers are set to the average, so they have no effect on the results.

A. Linear Predictors

Initially, a single predictor is used for the entire year. In the second model, the year is divided into four parts for which four predictors are applied. The third model divides the year into 12 parts. The fourth model divides the year into twenty-four parts. The following formula clearly demonstrates linear prediction.

$$Y_i = (X_i\beta) + e_i \tag{4}$$

where Y_i is dependent variable (to be predicted), X_i is the independent variable (to be used in Y prediction), β is the Regression coefficient, e_i is the error.

V. Results and Discussion

The forecasting results will span five-time frames: one-year, seasonal forecasting (four months), monthly forecasting, half-month forecasting, and weekly forecasting. Accordingly, each of these scenarios will involve a number of regressors.

$$Y_i(t) = \begin{cases} X_1\beta + e_1 \\ X_2\beta + e_2 \\ \vdots \\ X_n\beta + e_n \end{cases} \tag{5}$$

Every equation is used to predict a specific period of time, in another word there are many simple linear models that work together to form the final prediction curve. Later, we will compare the results with the LSTM results.

A. Whole Year Forecasting (One Regressor)

Short-term results from the single-regressor model are acceptable, but as can be seen in Fig. 3, over longer time periods, the model introduces extremely bad results.

The RMSE is depicted in Fig. 4 and demonstrates that the result differs significantly from the actual data (it gives the mean irradiance for the entire year). It is obvious that the system produces results with high errors (most days have an RMSE of 1 up to 4). Fig. 4 depicts the distribution of points along the range of RMSE, where around 90 points may have an RMSE equal to 0.3. As RMSE increases, we observe that the number of points decreases.

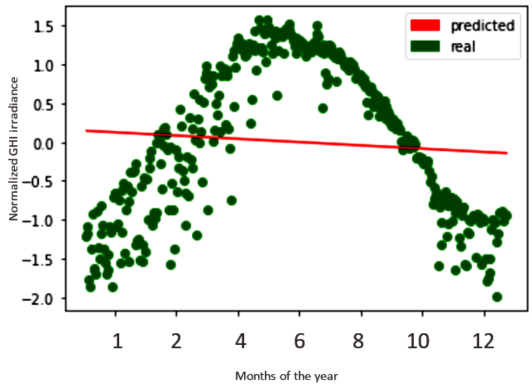

Fig. 3. One-part regression (whole year) output.

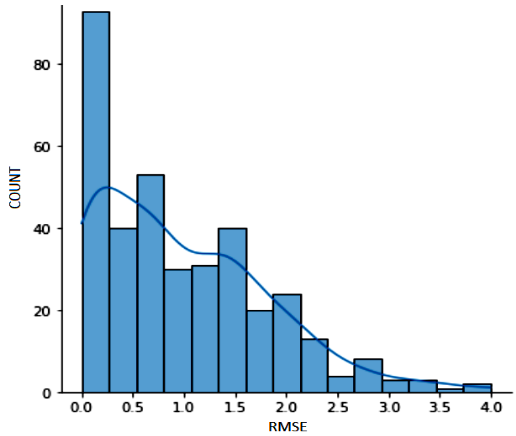

Fig. 4. RMSE for one-part regressor

B. Season Forecasting (Four Regressors)

Due to the inaccuracy of the (one regressor) model, the year has been divided into four parts, each part represents a distinct season. Using the same equation for each regressor, the outcome of our design is shown in Fig. 5.

The output curve shows that the model performs better than a single regressor model. The average of each season can be taken, where the summer has higher irradiance due to the low probability of cloud formulation. The probability distribution of the RMSE of the model is depicted in Fig. 6. It indicates that more points are located closer to the left, resulting in a lower RMSE, which is preferable to the RMSE for a one-part regressor. Evidently, error levels have decreased since the majority of errors are smaller than 2.

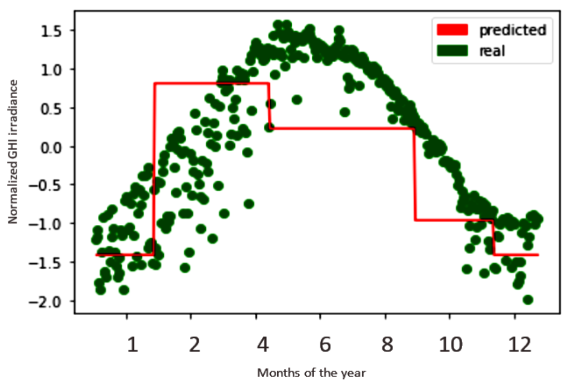

Fig. 5. Four regressors output.

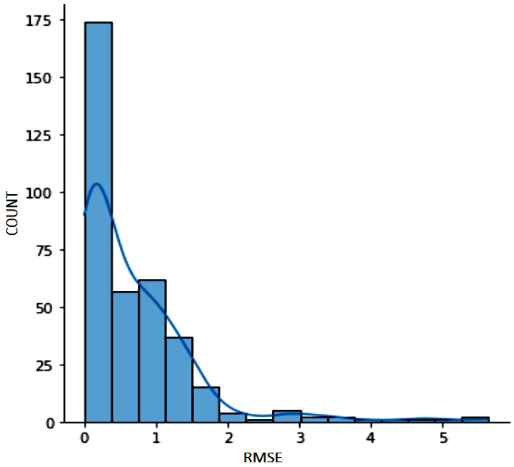

Fig. 6. RMSE for four regressors.

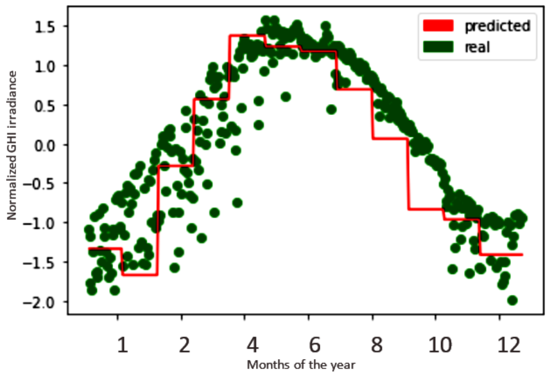

Fig. 7. Twelve regressors output.

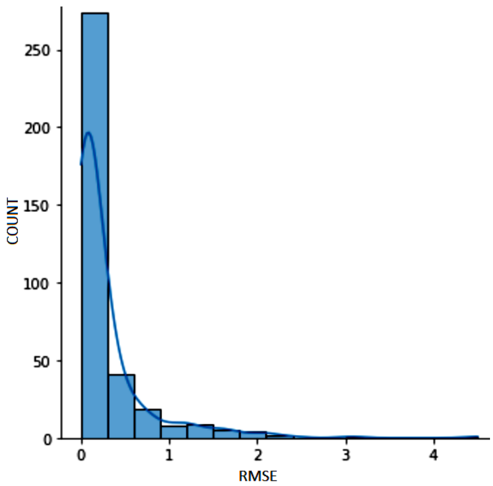

Fig. 8. RMSE for twelve regressors.

C. Monthly Forecasting (Twelve Regressors)

Although the RMSE for four regressors is less than 2 units, a more precise result is required. As illustrated in Fig. 7, this model uses a predictor for each month to achieve better results.

Based on Fig. 7, the predicted irradiance is closer to the actual data, and we can observe the RMSE probability distribution to examine our results in further depth.

According to the graph in Fig. 8, the number of days with errors of three or four units approaches zero, while the number of days with errors of less than half a unit increases (more than 250 days have that RMSE).

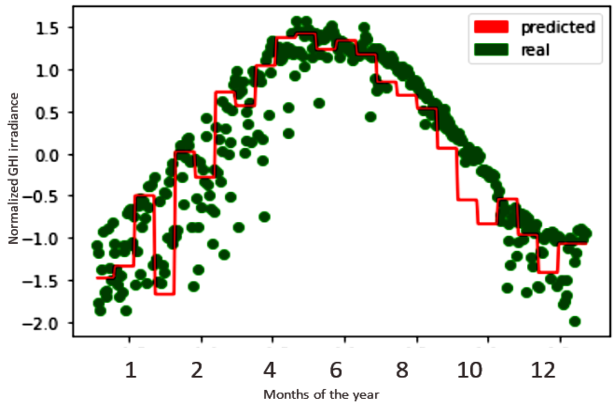

Fig. 9. Twenty-four-part regressors output

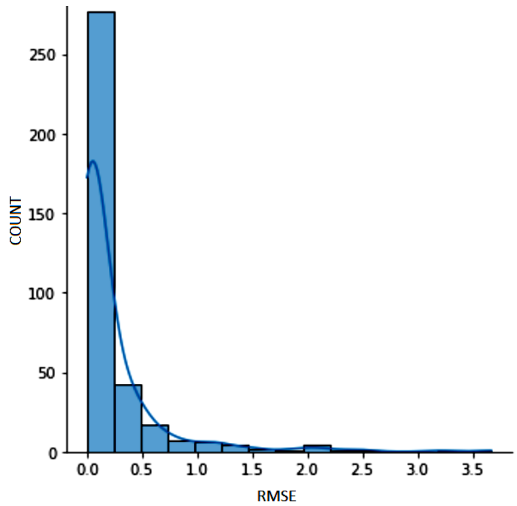

Fig. 10. RMSE for twenty-four regressors.

D. Half-Month Forecasting (Twenty-Four Regressors)

To improve the precision of the results, more regressors are added, one for every 15 days, as seen in Fig. 9. While Fig. 10 displays the RMSE probability distribution.

E. Weekly E-Budget (Fifty-Two Regressors)

For more precise predictions, fifty-two linear regressors are used separately. Fig. 11 depicts the forecasts. And the RMSE probability distribution is shown in Fig. 12.

Because there was no discernible difference in the RMSE curves between the (24-regressors and 52-regressors) scenarios, no more regressors were introduced.

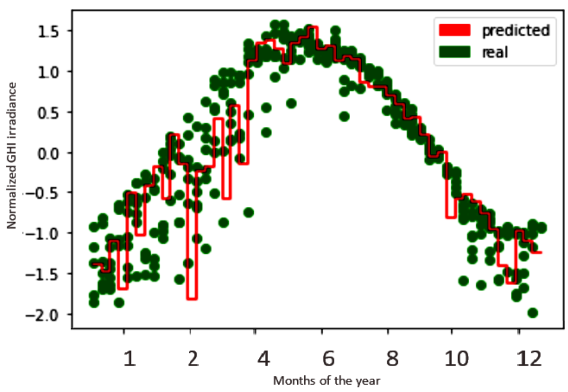

Fig. 11. Fifty-two-part regressors output.

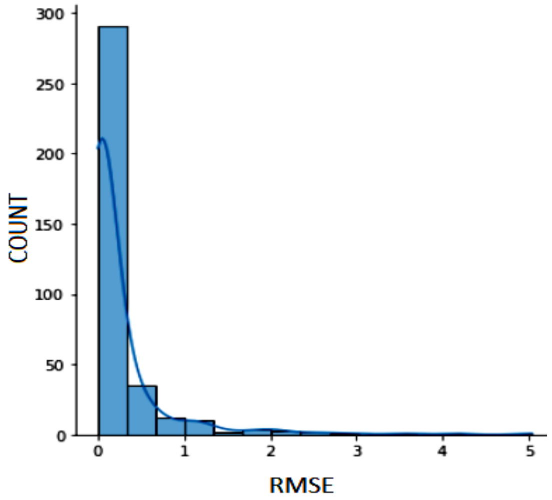

Fig. 12. RMSE for fifty-two-part regressors.

F. Cross-Validation

As training and testing data are separated, increasing the number of regressors enhances training performance. The same is done to the testing phase, increasing the number of regressors in the testing set due to "overfitting" for the training, which may not be suitable for the testing. The RMSE for training and testing is listed in Table I. We can observe that as the number of regressors increases, the training error decreases. On the other hand, the error decreases in the testing set until the twenty-fourth regressor; beyond that, the RMSE of the testing set increases.

In Table I, we can see the Min-RMSE in the fourth column, which shows the minimum value of RMSE shared between the training and testing, which is the goal of cross-validation. The min RMSE value that we obtained in the Table is at twenty-four linear predictors, which makes sense, refer to Fig. 9 comparing to Fig. 3, Fig. 5, Fig. 7, and Fig. 9.

TABLE I: THE CROSS-VALIDATION OF TESTING AND TRAINING

Method	RMSE in training	RMSE in testing	Min-RMSE training/ testing
One linear predictor	0.984	0.9799	
Four linear predictors	0.456	0.4611	
Twelve linear predictors	0.261	0.2812	
Twenty-four linear predictors	0.221	0.2277	0.221/ 0.2277
Fifty-two linear predictors	0.212	0.2475	
One hundred and four linear predictors	0.211	0.4382	

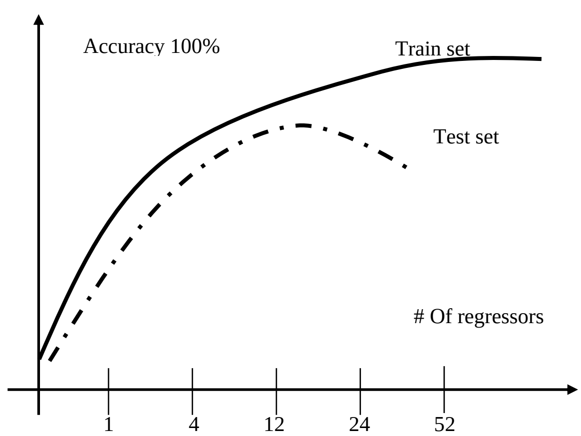

Fig. 13. The cross-validation of the training and testing.

The accuracy as a function of the number of regressors for the training set and testing set is depicted in Fig. 13. A number of regressors boost the accuracy of the training set. Nonetheless, the testing set improves accuracy. After the twenty-four regressor, the problem of overfitting occurs in the linear predictor. Consequently, the twenty-fourth regressor is the optimal regressor level for predicting the weather based on training data.

G. LSTM Predictor

Vanilla LSTM is the simplest LSTM model, with a single hidden layer and a single output. The input units are 24 units, a unit for every 15 days. The inputs are processed at the first layer, then transferred to the hidden layer, and finally transferred to the output layer, which gives the predictions for the selected period of the year. In this model, the data is divided, so every time 32 samples are processed at the same time. In other words, the batch size is chosen to be 32. Taking a number less than 32 may give better results, but it will take too much time. The epoch is selected to be 100. That means the model will iterate up to 100 times for every batch to have better results. The Augmented Dynamic Adaptive Model (ADAM) optimization model in python is used as an optimizer because of its high performance and fast convergence compared to the other optimizers. In Fig.14 are the LSTM prediction results for the given model.

It is clear that the LSTM model can give a pure curve (not over-fitted) as forecasting for the entire year. RMSE is given in Fig. 15 below.

Most of the time, the error in the forecasts is less than 0.6, which is a desirable number, making LSTM one of the most accurate methods.

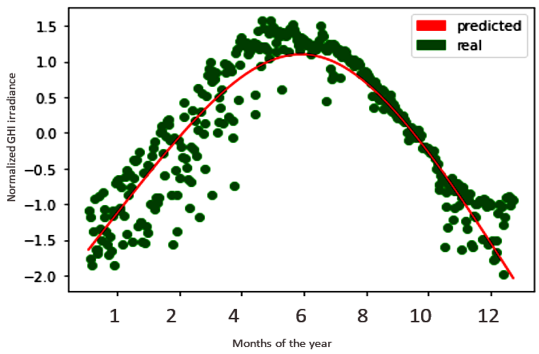

Fig. 14. LSTM output

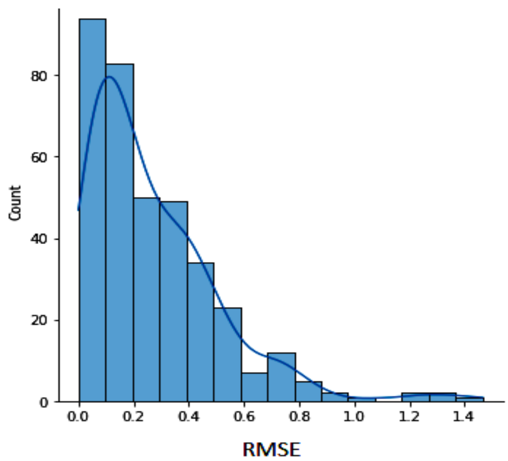

Fig. 15. RMSE for LSTM output

H. Discussion

The RMSE and processing time for each model are summarized in Table II.

TABLE II: RMSE AND PROCESSING TIME FOR DIFFERENT MODELS

Method	RMSE in average	Processing time
One linear predictor	0.9799	0.4264 second
Four linear predictors	0.4611	0.4308 second
Twelve linear predictors	0.2812	0.4329 second
Twenty-four linear predictors	0.2277	0.6397 second
Fifty-two linear predictors	0.2475	2.739 seconds
Vanilla LSTM	0.360	6.8090 second

From Table II, linear regression is used in different ways to obtain results close to the actual values and to vanilla LSTM, where the single linear regressor's amount of error is high throughout the year (giving an average value for the year). As the number of regressors increases, the time required to complete the forecast increases, and the value of the RMSE decreases. The RMSE decrement becomes shallower when twelve regressors are set in the model; better results are obtained with twelve and twenty-four regressors. However, increasing the number of regressors does not always lead to better results. In the case of the fifty-two regressors, it provides less accurate results for a longer time. And this result reinforces the cross-validation result in Table I, where the twenty-four linear predictors provide the best performance in terms of accuracy.

The LSTM model, on the other hand, has pure curvilinear results and takes a very long time to reach results, while RMSE is worse than twelve and twenty-four regressors. The LSTM took 6.8 seconds and had an RMSE of 0.36, whereas the 24-regressor model took 0.64 seconds and had an RMSE of 0.227. These results give the twenty-four linear model an advantage over the LSTM model. The comparison of the results figures and RMSE values reveals that the predicted values are more closely aligned with the real irradiance values, supporting the system's basic idea.

VI. CONCLUSION

Dividing the long term into smaller periods and using the simplest ways can give better results than forecasting the whole long term by using one regressor. Because of the non-ideality of natural phenomena, results will have a non-curvilinear shape. As the linear regressors increase, more accurate data will appear. While the LSTM has a non-over-fitted curve, it has complicated parameters that must be chosen by trying different values till the model gives better results. That takes more time besides processing time. For this reason and as a future work may need more efforts to have output shape closer to the curve and to have better predictions.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Abdulamjeed Sulaiman and Farhad E. Mahmood conceptualized, developed, validated, implemented, and wrote the initial draft of the article. Sayf A. Majeed supervised, reviewed, edited, and finalized the work. All authors have approved the final version.

ACKNOWLEDGMENT

The authors wish to thank the University of Mosul, Alnoor University College, and Al-Hadba University College for their support.

REFERENCES

[1] K. Yan, H. Shen, L. Wang, H. Zhou, M. Xu, and Y. Mo, "Short-term solar irradiance forecasting based on a hybrid deep learning methodology," *Information*, vol. 11, no. 1, Jan. 2020, doi: 10.3390/info11010032.

[2] M. Gao, J. Li, F. Hong, and D. Long, "Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM." *Energy*, vol. 187, #115838, Nov. 2019.

[3] E. Quansah, L. K. Amekudzi, K. Preko, J. Aryee, O. R. Boakye, D. Boli, and M. R. Salifu, "Empirical models for estimating global solar radiation over the Ashanti Region of Ghana," *Journal of Solar Energy*, vol. 2014, pp. 1–6, Jan. 2014.

[4] M. U. Siddiqui, O. K. Siddiqui, A. B. S. Alquaity, H. Ali, A. F. M. Arif, and S. M. Zubair, "A comprehensive review on multi-physics modeling of photovoltaic modules," *Energy Conversion and Management*, vol. 258, #115414, Apr. 2022.

[5] A. Dolara, S. Leva, and G. Manzolini, "Comparison of different physical models for PV power output prediction," *Solar Energy*, vol. 119, pp. 83–99, Sep. 2015, doi: 10.1016/j.solener.2015.06.017.

[6] IEA (2018), *Renewables 2018*, IEA, Paris https://www.iea.org/reports/renewables-2018

[7] A. Alzahrani, P. Shamsi, C. Dagli, and M. Ferdowsi, "Solar irradiance forecasting using deep neural networks," *Procedia Computer Science*, vol. 114, pp. 304–313, 2017.

[8] 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.

[9] J. A. Ruiz-Arias, H. Alsamamra, J. Tovar-Pescador, and D. Pozo-Vázquez, "Proposal of a regressive model for the hourly diffuse solar radiation under all sky conditions," *Energy Conversion and Management*, vol. 51, no. 5, pp. 881–893, May 2010.

[10] S. Siami-Namini, N. Tavakoli, and A. Siami Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *Proc. 17th IEEE Int. Conf. on Machine Learning and Applications*, Dec. 2018. doi: 10.1109/ICMLA.2018.00227.

[11] C. Voyant, G. Notton, S. Kalogirou, M. L. Nivet, C. Paoli, F. Motte, and A. Fouilloy, "Machine learning methods for solar radiation forecasting: A review," *Renewable Energy*, vol. 105, pp. 569–582, 2017. doi: 10.1016/j.renene.2016.12.095.

[12] R. Gelaro, W. McCarty, M. J. Suárez, *et al.*, "The modern-era retrospective analysis for research and applications, Version 2 (MERRA-2)," *Journal of Climate*, vol. 30, no. 14, pp. 5419–5454, Jul. 2017.

[13] A. Y. Saber and A. K. M. R. Alam, "Short term load forecasting using multiple linear regression for big data," in *Proc. IEEE Symposium Series on Computational Intelligence (SSCI)*, Nov. 2017, doi: 10.1109/ssci.2017.8285261.

[14] K. Hirose, "Interpretable modeling for short- and medium-term electricity demand forecasting," *Frontiers in Energy Research*, vol. 9, Dec. 2021, doi: 10.3389/fenrg.2021.724780.

[15] T. Lehtola and A. Zahedi, "Technical challenges in the application of renewable energy: A review," *International Journal of Smart Grid and Clean Energy*, pp. 689–699, 2020, doi: 10.12720/sgce.9.3.689-699.

[16] H. Suyono, H. Santoso, R. Nur Hasanah, U. Wibawa, and I. Musirin, "Prediction of solar radiation intensity using extreme learning machine," *Indonesian Journal of Electrical Engineering*

and Computer Science, vol. 12, no. 2, Nov. 2018, doi: 10.11591/ijeecs.v12.i2.pp691-698.

[17] M. Abdillah, W. Pramudito, T. Nugroho, and D. Fitria, "Solar irradiance forecasting using kernel extreme learning machine: Case study at Lamongan and Muara Karang Regions, Indonesia," *Journal of Engineering Science and Technology*, vol. 17, no. 2, pp. 1561-1576, 2022.

[18] M. Z. Hassan, M. E. K. Ali, A. B. M. S. Ali, and J. Kumar, "Forecasting day-ahead solar radiation using machine learning approach," in *Proc. 4th Asia-Pacific World Congress on Computer Science and Engineering* (APWC on CSE), Dec. 2017. doi: 10.1109/APWConCSE.2017.00050.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[20] S. Dutta, S. Jha, S. Sankaranarayanan, and A. Tiwari, "Learning and verification of feedback control systems using feedforward neural networks," *IFAC-Papers on Line*, vol. 51, no. 16, pp. 151–156, 2018, doi: 10.1016/j.ifacol.2018.08.026.

[21] A. Gasparin, S. Lukovic, and C. Alippi, "Deep learning for time series forecasting: The electric load case," *CAAI Trans. on Intelligence Technology*, vol. 7, no. 1, pp. 1–25, Sep. 2021. doi: 10.1049/cit2.12060.

[22] J. Cao and X. Lin, "Study of hourly and daily solar irradiation forecast using diagonal recurrent wavelet neural networks," *Energy Conversion and Management*, vol. 49, no. 6, pp. 1396–1406, Jun. 2008.

[23] T. Ozdemir, F. Taher, B. O. Ayinde, J. M. Zurada, and O. Tuzun Ozmen, "Comparison of feedforward perceptron network with LSTM for solar cell radiation prediction," *Applied Sciences*, vol. 12, no. 9, #4463, Apr. 2022, doi: 10.3390/app12094463.

[24] G. Reikard, "Predicting solar radiation at high resolutions: A comparison of time series forecasts," *Solar Energy*, vol. 83, no. 3, pp. 342–349, Mar. 2009.

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

[26] A. Milad, I. Adwan, S. A. Majeed, N. I. M. Yusoff, N. Al-Ansari, and Z. M. Yaseen, "Emerging technologies of deep learning models development for pavement temperature prediction," *IEEE Access*, vol. 9, pp. 23840–23849, 2021.

[27] P. Anki and A. Bustamam, "Measuring the accuracy of LSTM and BiLSTM models in the application of artificial intelligence by applying chatbot programme," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 1, Jul. 2021, doi: 10.11591/ijeecs.v23.i1.pp197-205.

[28] 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.

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

[30] P. P. Phyo, C. Jeenanunta, and K. Hashimoto, "Electricity load forecasting in Thailand using deep learning models," *International Journal of Electrical and Electronic Engineering & Telecommunications*, vol. 8, no. 4, pp. 221–225, 2019.

[31] W. González-Manteiga, M. Febrero-Bande, and M. Piñeiro-Lamas, "Semiparametric prediction models for variables related with energy production," *Journal of Mathematics in Industry*, vol. 8, no. 1, Aug. 2018, doi: 10.1186/s13362-018-0049-0.

[32] M. Mohamed, F. E. Mahmood, M. A. Abd, A. Chandra, and B. Singh, "Dynamic forecasting of solar energy microgrid systems using feature engineering," *IEEE Trans. on Industry Applications*, vol. 58, no. 6, pp. 7857-7869, Nov.-Dec. 2022.

[33] Modern-Era Retrospective analysis for Research and Applications, Version 2 'MERRA-2'. https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/

[34] S. B. Shams, V. P. Walden, J. W. Hannigan, W. J. Randel, I. V. Petropavlovskikh, A. H. Butler, and A. de la Cámara, "Analyzing ozone variations and uncertainties at high latitudes during sudden stratospheric warming events using MERRA-2," *Atmospheric Chemistry and Physics*, vol. 22, no. 8, pp. 5435–5458, Apr. 2022.

[35] L. Hoffmann, P. Baumeister, Z. Cai, J. Clemens, S. Griessbach, G. Günther, Y. Heng, M. Liu, *et al.*, "Massive-parallel trajectory calculations version 2.2 (MPTRAC-2.2): Lagrangian transport simulations on graphics processing units (GPUs)," *Geoscientific Model Development*, vol. 15, no. 7, pp. 2731–2762, Apr. 2022.

[36] L. Shikwambana, "Global distribution of clouds over six years: A review using multiple sensors and reanalysis data," *Atmosphere*, vol. 13, no. 9, Sep. 2022, doi: 10.3390/atmos13091514.

Abdulamjeed Sulaiman (almajeed.ahmed" at"alnoor.edu.iq) recived the B.S. degree in electronic engineering from University of Mosul in 2012 and M.S. degree in communication engineering from Cankaya University-Ankara-Turkey. Currently he is lecturer assistant in Alnoor University College-Mosul-Iraq. His research interest in machine learning and python.

Farhad E. Mahmood (farhad.m" at"uomosul.edu.iq) (S'13-M'19) received the B.S. and M.S. degrees from the University of Mosul in 2005, and 2008, respectively, and received the Ph.D. degree from The University of Kansas, the USA in 2019, all in electrical engineering. Currently, he is a lecturer in the Electrical Engineering Department of the College of Engineering, University of Mosul. His research interests include 5G wireless communication systems, green communications, RF antenna design, radio network planning, positioning location. Dr. Mahmood was a scholar of CARA organization for two years starting from 2010. Then, in 2012, he was a Higher Committee of Education and Development in Iraq scholar to obtain his Ph.D., when he received the Telemetry Award in 2019.

Sayf A. Majeed was born in Nineveh, Iraq, in 1981. He received his B.S. degree in Electrical, Electronics and Communication Engineering from the University of Mosul, Iraq, in 2004, an MSc in Computer and Communication Engineering from the National University of Malaysia (UKM) in 2008, and a Ph.D. degree in Electrical, Electronics, and Systems Engineering from the National University of Malaysia (UKM), in 2017. Currently, he is a lecturer at the Department of Technical Computer Engineering, Al-Hadba University College, Mosul, Iraq. He has written articles for several international conferences and journals. His research interests are in speech recognition, speech and audio processing, communications and multimedia computing, machine learning, deep learning applications, and pattern recognition.