*Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi*
*Sayı 7, S. 201-206, 7, 2023*

***Araştırma Makalesi***

IJANSER

https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

*International Journal of Advanced Natural Sciences and Engineering Researches*
*Volume 7, pp. 201-206, 7, 2023*

***Research Article***

# Robust Estimation for Solar Radiation in Renewable Energy Systems

Muhammad Saad*, Mirza Jahanzaib[1], Atta Ur Rehman[2]

[1] *Engineering Management Department, University of Engineering and Technology, Taxila, 47050 Pakistan*
[2] *Industrial Engineering Department University of Engineering and Technology, Taxila, 47050 Pakistan*
[3] *Electrical engineering Department, University of Engineering and Technology, Mardan, 23200 pakistan*

*(*msaadkhan841@gmail.com *)*

*(Received: 15 August 2023, Accepted: 29 August 2023)*

(1st International Conference on Recent and Innovative Results in Engineering and Technology ICRIRET 2023, August 16-18, 2023)

**ATIF/REFERENCE:** Saad, M., Jahanzaib & Rehman, A. U. (2023). Robust Estimation for Solar Radiation in Renewable Energy Systems. *International Journal of Advanced Natural Sciences and Engineering Researches*, 7(7), 201-206.

***Abstract*** **–** Renewable energy is gaining huge respect due to the increase in demand for energy globally, low cost production, and sustainability of the environment. The one to mention out of all renewable energies is solar energy. Due to its availability in abundance. A wide range of people are driven to invest in this industry. Because of its high co-linearity, it's important that proper estimation of solar radiation is necessary for long-term performance, management of energy, and economic aspects. In this paper global horizontal irradiance is predicted from input parameters namely, air temperature, relative humidity, wind speed, wind spend of ghust, Wind direction in degrees north counted clockwise (standard deviation), Wind direction in degrees north counted clockwise, day length, and air pressure. Based on these factors the paper investigates the performance of different machine-learning models for solar radiation estimation. $R^2$(coefficient of determination) of the models is compared, and the best one, i.e., gradient boost regressor model is selected which predicts solar radiation in Pakistan from metrological data. Considering wind speed of gust and wind direction in a single model with high accuracy, and aids knowledge to predictions using machine learning models. The paper is part of an ongoing effort in the research community to predict solar radiation with ease and accuracy. As real-time data is difficult and costly due to high calibration costs.

*Key Words – Artificial, İntelligence, Machine Learning, Prediction, Solar Radiation, Renewable Energy.*

## I. INTRODUCTION

In recent years, renewable energy has gained much attention from governments, industrial sectors, and the research community due to the rapid growth of energy demands around the world. Solar energy plays a significant role in providing sustainable energy. The sun's energy is the main source of solar energy at the surface of the earth that is useful for numerous applications, particularly, increasing water's temperature, excitation of the electrons in a photovoltaic cell, providing energy to a natural process, and many more. More specifically, the solar radiation obtained by the earth 's surface significantly varies on a daily or monthly basis because of the existence of both suns as well as clouds. Furthermore, it also differs on an hourly basis because of the relative position of the sun to the east or west. Therefore, obtaining accurate information on the intensity of solar radiation at a particular location is significant to develop projects related to solar energy and economic assessment of solar energy systems. For a good analysis, the solar radiation data of the real-time that is recorded for a brief period is always preferred. However, numerous factors significantly affect the data, namely, location, time, maintenance, as well as costly calibration. Generally, professional research labs or universities perform such measurements for a brief period. Consequently, in developing countries and particularly in Pakistan, recording accurate solar radiation data is challenging [1].

The solar radiation predictor models convert the climatological as well as geographical data in the absence of properly real-time measured solar radiation data. The research community has mainly focused on the prediction of solar radiations for the approaching hours and days. As it is obvious that an hour is usually the smallest interval of time used in official measurements of meteorological stations. Moreover, hourly recorded solar radiation data show more details and are helpful for numerous applications. Therefore, hourly solar radiation data through statistically tested prediction models have a significant role in the application as well as the design of solar systems. But the availability of devices to measure solar radiation in most regions is difficult due to two factors, namely, technology requirements and expensive cost. Therefore, estimating the amount of solar radiation data through algorithms or mathematical models, particularly daily solar radiation is extremely significant to designing and developing systems for solar energy [1].

In Pakistan, the types of energy, namely, solar and wind are considered significant sources of renewable energy. More specifically, renewable energy is sustainable energy, which plays a significant role in protecting us from climatic changes [2,3]. Precise estimation of solar radiation is mandatory for solar energy systems. Estimating the wind speed having minimum errors is convenient for the economics and security of wind power. Moreover, wind energy is reflected as extremely non-linear as well as changeable, however, it increases safety, minimizes air pollution, is sustainable energy, and is environmentally friendly [4,5]. The wind speed estimation models are further categorized into three categories, particularly, short-term, medium-term, and long-term estimating models. The short-term forecasting model performs well compared to other models [6]. Two energy sources, particularly, wind and solar do not emit greenhouse gas due to the reason that their source of energy or fuel is free from carbon. Emissions of the amount of Carbon dioxide ($CO_2$) can significantly be minimized in electricity sectors by utilizing a small electric grid consisting of both solar and wind energy. Compared to wind energy, solar energy has rare non-linearity and is usually affected by three factors, particularly, wind, temperature, and solar panels [7, 8]. Numerous types of prediction models including machine learning and deep learning-based models can be utilized for both solar and wind energy forecasting. However, deep learning-based models are considered significant for data processing to show the complicated relationships between input and output. It also provides good capabilities for generalization as well [9,10]. Moreover, it is well suited for the uncertain behavior of solar and wind energy data. This work aims to determine both the environmental and societal issues because it estimates the near future conditions of the environment and natural disasters [11].

The research community examined many artificial neural networks (ANN) based methods including evolutionary-ANN, multi-layer perceptron (MLP), and regression neural network (RNN) [12]. Has developed an automatic selection of the parameters for the forecasting of global solar irradiation. The selection was done through evolutionary-ANN by utilizing numerous operators, namely, selection, mutation, and crossover. Moreover, other significant factors including the potential generation effect, population size, mutation percentage, and cutting points nature were explored. Comparative analysis was performed to validate the efficiency of evolutionary-ANN using both forward stepwise and heuristic search regression models. The study also explored three temperature-based and k-nearest neighbor (KNN) models to compare the performance for forecasting [13]. Explored four different models, namely, ANN, MLP, generalized RNN, and radial basis neural networks for forecasting the daily global solar radiation utilizing the meteorological parameters including temperature, relative humidity, pressure, sunshine, and water vapor pressure. Furthermore, the Bristow-

Campbell method has been enhanced by including other factors, namely, cloud cover, relative humidity, etc. Among all the above four models, the ANN performed better in forecasting the daily solar radiation [14]. Artificial intelligence-based hybrid technique comprised of an integration of an ANN and simulated annealing (ANN/SA) to determine the daily solar radiation. ANN/SA approach utilized temperature based on the simulated annealing for enhancing the performance of an ANN. Furthermore, a comparative analysis was performed with support vector machine (SVM), ANN, and other machine learning models, however, the ANN/SA outperformed all the existing algorithms [15]. Investigated three different models, namely, an extreme learning machine, a hybrid genetic algorithm with neural networks, and a random forest for the forecasting of global horizontal irradiance. Parameters, namely, minimum temperature, maximum temperature, duration of sunshine, and humidity were utilized for training the above-mentioned models. Comparative assessments of performance against two other models, namely, Angstrom-Prescott and Ogelman were conducted. Moreover, it was reported that the duration of sunshine has the most influential effect in all three models. The extreme learning machine has outperformed compared to the other two models [16]. Presented a comparative analysis of numerous machine learning algorithms, namely, hybrid mind evolutionary, ANN, random forest, wavelet neural network, and neural network as well as the empirical models, namely, Hargreaves-Samani, Bristow-Campbell, Fan, and Jahani for the forecasting of global solar radiation by utilizing only the temperature as an input parameter. However, the hybrid mind evolutionary and ANN performed better compared to the other models [17]. A model based on sunshine hours, a model based on temperature, a model based on clouds, a model based on meteorological factors, and a model based on meteorological and geographic parameters are designed. The most accurate methods for measuring sun radiation in Pakistan, according to the results, are those based on meteorological and geographic parameters [18].

Solar radiation isn't predicted by the wind speed of gusts and wind direction. variables in Pakistan. This research considers both aspects to predict solar radiation. The objective of the study is to design a robust estimation model for solar radiation based on geographical and meteorological variables which accurately estimate solar radiation and aid the ongoing effort in predicting solar radiation. The remaining paper is organized as section II describes the methodology followed by section IV results and discussion. Section V covers the conclusion.

## II. METHODOLOGY

### A. *DATA SOURCE*

Four different metrological sites were selected for data fetch from ENERGYDATA.INFO i.e., Islamabad, Karachi, Peshawar, and Lahore having a frequency of 10min in CSV file. The services were provided by a German base company namely CSP Services GmbH. The guideline for metrological instruments and method of observation for the world metrological Organization (WMO) NO.8 were followed for quality control. Sunset and sunrise time datasets were obtained from online sources to differentiate between tha stations. Variables are all taken in natural logrithium.

### B. METHOD OF ANALAYSIS

In At the 1st leisure, all the required libraries were imported i.e., pandas, numpy, seaborn, matplotlib, DateTime, and Sklearn in phyton. Organizing of da ta is carried out through pandas, for data manipulati on and analysis. The data is read, hours are convert ed to minutes, the date from a string form is conver ted for differentiation and the loop function is enfor ced for all 4 cities. The data from all the cities are c ombined in a file named combined CSV using the C oncat function. Ghi(Global Horizontal Irradiance)is the dependent variable, where null the row is dropp ed. Further on the null power supply and station ma intenance rows are dropped. Where wind from dire ction standard deviation, wind speed, wind from dir ection, and wind speed of gust are collectively zero , drop the row. Time and actual date columns are al so dropped out.

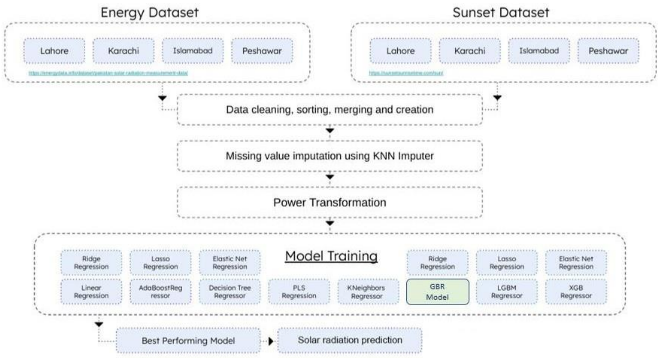


Fig. 1 Flow Diagram.

The shuffling of data is done to avoid overfitting the model. Test size is taken as 25% of the combined data and trained as 75%. Correlation is carried out to check whether we can drop a column or not. K-Nearest Neighbors Algorithm is used for filling in the missing values in the rows, with having a **K** value of 3. Power transportation is carried out to normalize, standardize and remove skewness to obtain data in a clean and desired form. Different models are applied as shown in Fig.1of which the best one, Gradient Boost Regressor was selected, a decision tree classifier with respective hyperparameters of values estimator's=3000, learning rate=0.05, max depth=4, max features='sqrt', min samples leaf=15, min samples split=10, loss='huber', random state =42.Further model is validated on the 25% testing data, day-wise solar radiation predictions for different cities of Pakistan, and instant time predictions are carried out.

## III. RESULT AND DISCUSSION

The graphs plotted using the Matpot library show d ay-wise predictions of cities i.e; Peshawar, Islamab ad, Karachi, and Lahore.

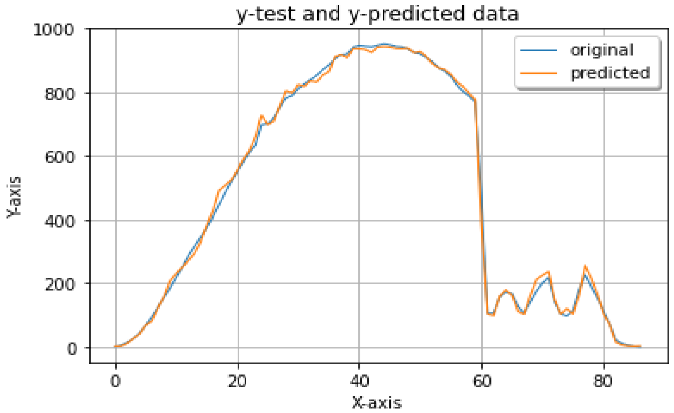


Fig. 2 Solar Radiation of Peshawar, Pakistan the date,17/07/2015

The graph shows the comparison of day-wise solar radiation between actual and predictive values on the given date of city Peshawar, KPK Figure 2. Solar Radiation of Peshawar, Pakistan. The orange line shows the actual solar radiation while the blue line shows the predicted solar radiation. The x-axis includes the data points numbers and y-axis includes the solar radiation.

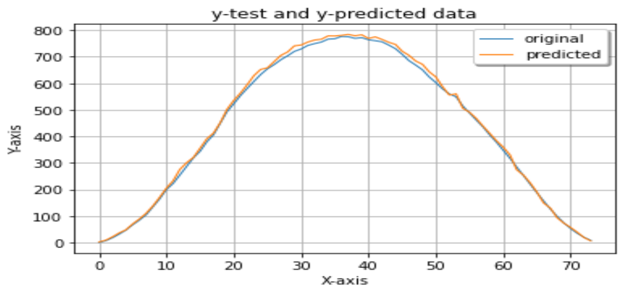


Fig. 3 Solar radiation of Islamabad, Pakistan at date: 17/07/2015.

The graph shows the comparison of day-wise solar radiation between actual and predictive values on th e given date of city Islamabad, Pakistan

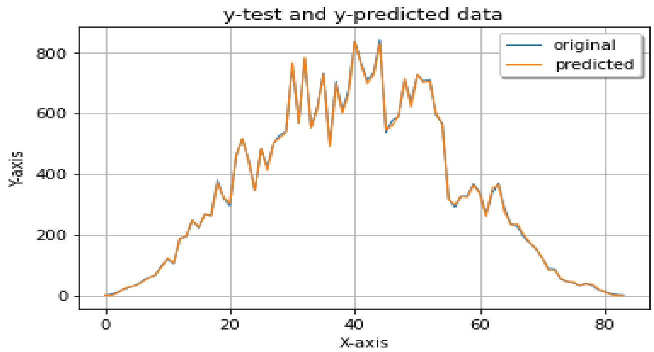


Fig. 4 Solar Radiation of Karachi, Pakistan the date,17/07/20 15.

The graph shows the comparison of day-wise solar radiation between actual and predictive values on th e given date of the city Karachi, Sindh province of Pakistan.

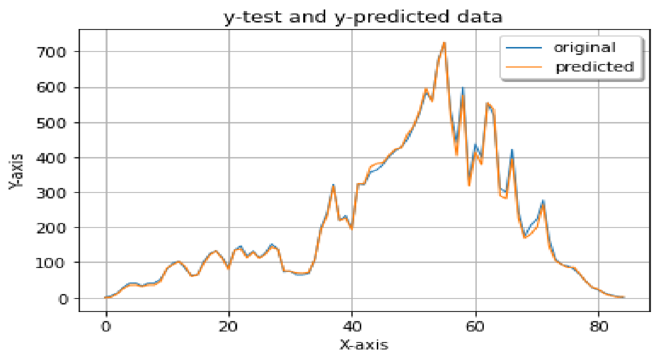


Fig. 5 Solar Radiation of Lahore, Pakistan the date,17/07/201 5.

The graph shows the comparison of day-wise solar radiation between actual and predictive values on th e given date of city Lahore, Punjab province of Pak istan.

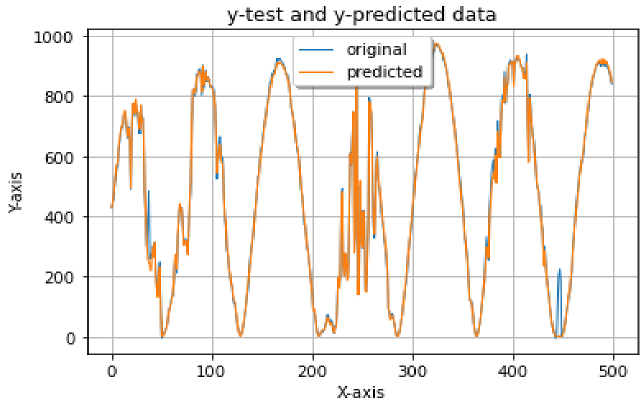


Fig. Comparison of actual and predictive solar radiation

The graph shows a comparative analysis of actual and predictive data from point 100000 to 100500 i.e; in range 500, while calculating the mean absolute percentage error(MAPE), MAPE= 2.95%

Table 1 Results Comparison

| Model Name | Coefficient of determination ($R^2$) |
|---|---|
| Linear Regression | $R^2_{(Train)}$ = 0.8617<br>$R^2_{(Test)}$ = 0.8616 |
| Rigid Regression | $R^2_{(Train)}$ = 0.8617<br>$R^2_{(Test)}$ = 0.8616 |
| Elastic Net Regression | $R^2_{(Train)}$ =0.8288<br>$R^2_{(Test)}$ = 0.8285 |
| Lasso Regression | $R^2_{(Train)}$ = 0.8615<br>$R^2_{(Test)}$ =0.8613 |
| GBR Regression | $R^2_{(Train)}$ = 0.9647<br>$R^2_{(Test)}$ = 0.9648 |
| LGBM Regression | $R^2_{(Train)}$ = 0.9601<br>$R^2_{(Test)}$ =0.9607 |
| XGB Regression | $R^2_{(Train)}$ = 0.9621<br>$R^2_{(Test)}$ =0.9623 |
| K-Neighbor Regression | $R^2_{(Train)}$ = 0.9844<br>$R^2_{(Test)}$ =0.9547 |
| PLS Regression | $R^2_{(Train)}$ = 0.8414<br>$R^2_{(Test)}$ =0.8410 |
| Decision Tree Regression | $R^2_{(Train)}$ = 0.7801<br>$R^2_{(Test)}$ = 0.7771 |
| ADA Boost Regression | $R^2_{(Train)}$ = 0.8521<br>$R^2_{(Test)}$ = 0.8513 |

Table 1 shows the coefficients of determination of the applied models. Having the highest value for GBR Model. which indicates that the Gradient Boost Regressor model can predict solar radiation accurately, which can be helpful for economic assessment and power management.

## IV. Conclusion

Due to environmental sustainability, a shift has been seen towards renewable energy from non-renewable. Solar energy is the main source of fulfilling energy needs and providing sustainable energy with low-cost production. The energy of the sun available on the earth's surface is known as solar energy. The variation in solar radiation occurs due to the position of the sun and the cloud's existence. It is of utmost importance for economic purposes and the intensity of solar radiation at a particular location. For a brief period, solar radiation in real-time is recorded for some good analysis. However, recording data in developing countries is somewhat challenging due to the costly calibration, universities or professional research centers perform such measurements. In this study, different prediction machine learning models were applied. For data organization, manipulation, and analysis using the Pandas library. Correlation analysis was carried out to check either if a column can be dropped. Sklearn was used for model training out of which the best one was selected i.e., Stochastic gradient boost regressor (GBR Regressor) used, gradient boost regressor performing tree analysis having the following output parameters; $R^{2=}$0.9651 (Train), $R^{2=}$0.9615 (Test), MSE=3085.2848, and RMSE=55.5453. Hence it shows that the research will aid in the economic and managerial complications related to solar radiation and solar plants. By giving robust results with an accuracy of 97%, which helps in precisely estimating solar radiation.

## References

[1]. A. Mosavi, M. Salimi, SF. Ardabili, T. Rabczuk, S Shamshirband, AR. Varkonyi-Koczy "State of the art of machine learning models in energy systems, a systematic review" Energies, vol. 12, pp. 1301-1313, 2020.

[2]. SP. Mishra, PK. Dash "Short term wind speed prediction using multiple kernel pseudo inverse neural network." Int J Autom Comput, vol. 15, pp. 66–83, 2017.

[3]. M. Wiston, KM. Mphale "Weather forecasting: from the early weather wizards to modern–day weather

forecasting.” J Climatol Weather Forecast, vol. 6, no. 2, pp. 1–9, 2018.

[4]. G. Memarzadeh, F. Keynia “A new short term wind speed forecasting method based on fine-tuned LSTM neural network and optimal input sets.” Energy Convers Manag, vol. 213, pp. 112824-112840, 2020.

[5]. RP. Shetty, A. Satyabhama, PS. Pai “Wind speed prediction and modelling: a comparison of feed forward neural networks.” Int J Mech Product Eng Res Develop, vol. 10, pp. 116–127, 2020.

[6]. SR. Moreno, RG. da Silva, VC. Mariani, LS. Coelho “Multi- step wind speed forecasting based on hybrid multi-stage decompo- sition model and long short-term memory neural network.” Energy Convers Manage, vol. 213, pp. 12869-12882, 2020.

[7]. AH. Elsheikh, SW. Sharshir, AE. Mohamad, AE. Kabeel, W. Guilan, Z. Haiou “Modeling of solar energy system using artificial neural network: a comprehensive review”. Sol Energy, vol. 180, pp. 622–639, 2019.

[8]. JM. Barrera, Reina A, Maté A, Trujillo JC “Solar energy prediction model based on artificial neural networks and open data”, Sustainability vol. 12, pp. 6915-6924, 2020.

[9]. H. Wang, Y. Liua, B. Zhoub, C. Lib, G. Caoa, N. Voropaic, E. Barakht- enkoc “Taxonomy research of artificial intelligence for deterministic solar power forecasting.” Energy Convers Manage, vol. 214, pp. 1–17, 2020.

[10]. J. Chorowski, JM. Zurada “Extracting rules from neural networks as decision diagrams.” IEEE Trans Neural Networks, vol. 22(12), pp. 2435–2446, 2011.

[11]. SK. Biswas, N. Sinha, B. Purkayastha, L. Marbaniang “Weather prediction by recurrent neural network dynamics.” Int J Intell Eng Inf, vol. 2, no.2/3, pp. 166–180, 2014.

[12]. V. Wu, G. Huang, J. Fan, F. Zhang, X. Wang, W. Zeng. “Potential of kernel-based non- linear extension of Arps decline model and gradient boosting with categorical features support for predicting daily global solar radiation in humid regions.” Energy Convers Manag, vol. 183, pp. 280–95, 2019.

[13]. M. Marzouq, Z. Bounoua, H. El Fadili, A. Mechaqrane, K. Zenkouar, Z. Lakhliai. “New daily global solar irradiation estimations model based on automatic selection of input parameters using evolutionary artificial neural networks.” J Clean Prod, vol. 209, pp. 1105–18, 2019.

[14]. L. Wang, O. Kisi, M. Zounemat-Kermani, GA. Salazar, Z. Zhu, W. Gong. “Solar radiation prediction using different techniques: model evaluation and comparison.” Renew Sustain Energy Rev, vol. 61, pp. 384–97, 2016.

[15]. Mousavi SM, Mostafavi ES, Jiao P. “Next generation prediction model for daily solar radiation on horizontal surface using a hybrid neural network and simulated an- nealing method Energy”, Convers Manag, vol. 153, pp. 671–82, 2017.

[16]. Y.Feng,N. Cui, Y. Chen, D. Gong, X. Hu. “Development of data-driven models for prediction of daily global horizontal irradiance in northwest China”, J Clean Prod, vol. 223, pp. 136–146, 2019.

[17]. Y. Feng, D. Gong, Q.Zhang, S.Jiang, L. Zhao, N.Cui. “Evaluation of temperature-based machine learning and empirical models for predicting daily global solar radiation”, Energy Convers Manag, vol. 198, pp.117-122, 2019.

[18]. H. Khalid, M. Zakaria.). “Estimation of solar radiation in southern areas of Pakistan using radiation models”. Journal of Renewable and Sustainable Energy, vol. 8(4), 2016