

# Development of a Mathematical Model to Forecast Solar Radiation and Validating Results Using Machine Learning Technique

H.M.K.K.M.B. Herath (✉ [kasunherathlive@gmail.com](mailto:kasunherathlive@gmail.com))

The Open University of Sri Lanka <https://orcid.org/0000-0002-1873-768X>

S.V.A.S.H. Ariyathunge

The Open University of Sri Lanka

G.M.K.B. Karunasena

The Open University of Sri Lanka

---

## Research Article

**Keywords:** Solar Radiation, Machine Learning, Multiple Linear Regressions, Mathematical Modeling

**Posted Date:** June 30th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-669429/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Solar radiation or also referred to as solar power is the general expression for electromagnetic radiation emitted by the Sun. Direct solar radiation is an important factor in global solar radiation and is very influential in the efficiency evaluation of various applications for solar energy. For countries like Sri Lanka, installing a solar radiation instrument in rural areas is a challenge. Thus, both scientific and economically, measuring solar radiation without installing measuring instruments is an advantage. The aim of this study is to development of a mathematical model to predict solar radiation where solar radiation measurement instruments are not installed. The Artificial Neural Network (ANN) was used to verify the predictions of the mathematical model. Multiple Linear Regression (MLR) analysis was used for the development of a mathematical model to predict solar radiation. The model with the highest  $R^2$  value (**0.5973**) was chosen from the **127 equations** as the best model that describes the solar radiation that reaches the surface of the earth. The dataset used for this study was meteorological data from the four month *HI-SEAS* weather station and are composed of ten attributes including date, time, radiation (**H**), temperature (**T<sub>air</sub>**), pressure (**P**), humidity (**φ**), sunrise time, sunset time, wind direction (**D**), and speed (**S**). The angle of declination (**δ**) and sunshine hours (**N**) were calculated using the dataset. For the training of the neural network, **80 %** of the data from the *HI-SEAS* dataset was used. The remaining data were used for testing both mathematical and ANN models. Results obtained from the multiple linear regression method and the ANN method was compared with the measured values. The experimental results suggested that the mathematical model was predicted the solar radiation with **±100 Wm<sup>-2</sup>** tolerance for both measured and ANN values.

# Introduction

Solar radiation is the component of the radiation of the Sun that comes onto the surface of the earth. Such energy is uniquely required for other uses, such as increasing water temperature in a photovoltaic cell or transferring electrons through it [1], solar electricity, and solar ventilation. In addition, it provides energy for natural processes such as photosynthesis. Kalogirou (2013) [2] claimed that solar radiation is the fundamental source of the energy of the Planet, supplying approximately 99.97% of the heat energy available in the atmosphere, ocean, soil, and other water bodies for different chemical and physical processes. To predict global solar radiation, where solar radiation measurement instruments are not installed, the development of a mathematical model is important. For plentiful realistic uses, this model would be very beneficial for the efficient uses of a vast volume of free ecological solar power. A statistical approach that uses many explanatory variables to predict the outcome of a response variable is multiple linear regression (MLR), also known simply as multiple regression. Multiple linear regression (MLR) [3] attempts to model the linear relationship between the (independent) explanatory variables and the (dependent) response variable. This study mainly focused on the development of a mathematical model for forecast solar radiation using the multiple linear regression method. Artificial Neural Networks (ANNs) can be used for various functions such as forecasting, curve fitting, and regression. The basic unit of an

Artificial Neural Network is a neuron that utilizes a transfer function to formulate the output. In this study, Artificial Neural Networks were used to validate the mathematical model for predicting solar radiation.

## Methodology

The ultimate objective of this study is to development of a mathematical model for predict solar radiation and verify predicted data using the Artificial Neural Network. Meteorological data from the four month *H-SEAS* weather station [4] are the datasets used for this study and are composed of eleven characteristics, including date, time, radiation, temperature, pressure, humidity, sunrise time, sunset time, wind direction, and speed. The methodology of this study was comprised of four main steps such as Development of a mathematical model for predicting solar radiation, Development of an ANN model to predict solar radiation based on measured data, Perform the test run for mathematical model and verifying results with ANN, Analyze the accuracy of the results using the data analyzing method.

In order to find the relationship between selected parameters and solar radiation, linear regression analysis was performed. The response variable was set as solar radiation ( $H$ ) and the other variables such as open-air temperature ( $T_{air}$ ), atmospheric pressure ( $P$ ), relative humidity ( $\varphi$ ), wind direction ( $D$ ), wind speed ( $S$ ), sunshine hours ( $N$ ) and angle of declination ( $\delta$ ) were set as explanatory variables.

Declination angle ( $\delta$ ) was calculated as follows, **see formula 1 in the supplementary files.**

Where, n is the Julian day of the year

In the process of applying multiple linear regressions, one variable is chosen out of seven parameters, and seven equations ( ${}^7C_1$ ) for the selected explanatory variable were obtained. Similarly, out of the seven parameters, two variables were chosen and twenty one ( ${}^7C_2$ ) equation numbers were obtained. This process was repeated for 3, 4, 5, 6, and 7 numbers of variables. In the above process total of 127 ( ${}^7C_1 + {}^7C_2 + {}^7C_3 + {}^7C_4 + {}^7C_5 + {}^7C_6 + {}^7C_7$ ) number of equations were obtained. Out of the many models developed, seven models ( $H_1 - H_7$ ) were found to have acceptable levels of accuracy.

**See formula 2 in the supplementary files.**

The determination coefficient, denoted by  $R^2$ , is a number in mathematics that shows how well a mathematical model suits data. The model with the highest  $R^2$  value was chosen from the 127 equations as the best model that describes the solar radiation that reaches the surface of the earth.

Table 1- The  $R^2$  value of the obtained mathematical models

Model	$H_1$	$H_2$	$H_3$	$H_4$	$H_5$	$H_6$	$H_7$
<b>R-Square</b>	0.5973	0.5948	0.5659	0.5588	0.5546	0.3116	0.3113

Table 1 depicts the R-square ( $R^2$ ) value for seven models and results suggested that the higher value of  $R^2$  is the best performance of the model. Therefore model  $H_1$  was selected as the best model at high accuracy.

See formula 3 in the supplementary files section.

The proposed architecture of the neural network was developed using the method of trial and error. We analyzed the precision of the various architectures of the neural network and used it to select architectures that, depending on the sample dataset, would provide better accuracy. The ANN was used to verify the mathematical model results. The mathematical model used the same data from the testing dataset and compared the forecast of ANN to the prediction of the mathematical model.

## Results And Analysis

Table 2 shows the validation of model  $H_1$  with the results of ANN predicted data. Mathematical model predictions were the closest value to the ANN prediction with a tolerance of  $\pm 100 \text{ Wm}^{-2}$  according to test results data.

Table 2- Validation of the model  $H_1$  with seven variables ( $T_{air}$ ,  $P$ ,  $\varphi$ ,  $D$ ,  $S$ ,  $n$ , and  $\delta$ )

Test	<i>Measured Value (<math>\text{Wm}^{-2}</math>)</i>	<i>Predicted Value</i>	
		<i>Mathematical Model (<math>\text{Wm}^{-2}</math>)</i>	<i>ANN Model (<math>\text{Wm}^{-2}</math>)</i>
1	1039.22	965.09	1032.57
2	1027.27	1001.17	1039.73
3	915.28	958.75	1048.32
4	1040.64	956.74	906.95
5	1045.82	1017.41	961.21

The Levenberg Marquardt (LM) algorithm was used to training the dataset and Figure 1 depicts the regression graph of the training. The best training performance was observed with  $1.4537 \times 10^{-20}$  at epoch 35. Figure 2 illustrates the variation of solar radiation for the regression analysis model and ANN model. The experimental results show that the mathematical model has a  $\pm 100 \text{ Wm}^{-2}$  tolerance with measured values. Moreover, we have observed that the results are close to the approximate values of the ANN predictions.

## Conclusion

In this study, we have developed a mathematical model for predicting solar radiation and verify results with ANN prediction that develop and train with the same dataset. Multiple linear regression analysis was performed to obtain a total of 127 numbers of equations. The mathematical model with the highest  $R^2$  (0.5973) was selected as the best model. ANN model was trained with the same dataset that was used to test the mathematical model. Results obtained from mathematical model was compared with the ANN prediction and the experimental result shows that the mathematical model predicts solar radiation with  $\pm 100 \text{ Wm}^{-2}$  tolerance for both measured and ANN predicted values.

## Declarations

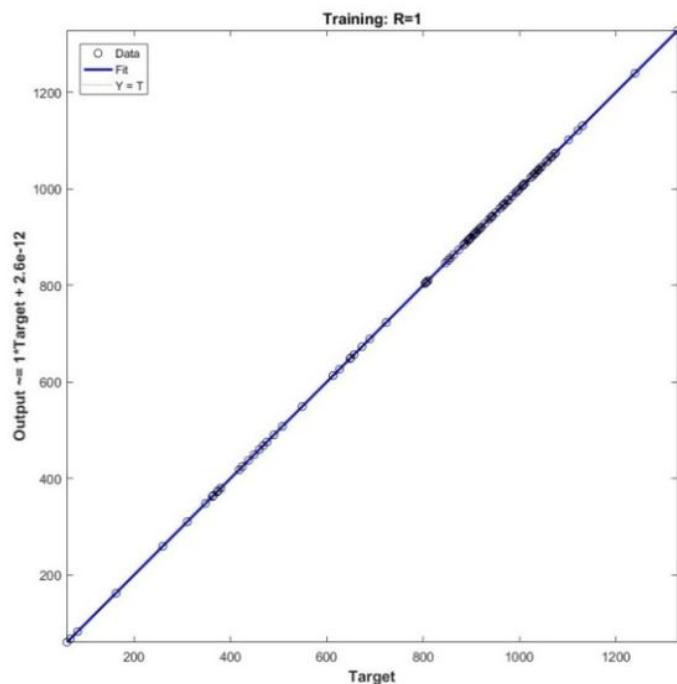
### Competing Interests

The authors declare no competing interests.

## References

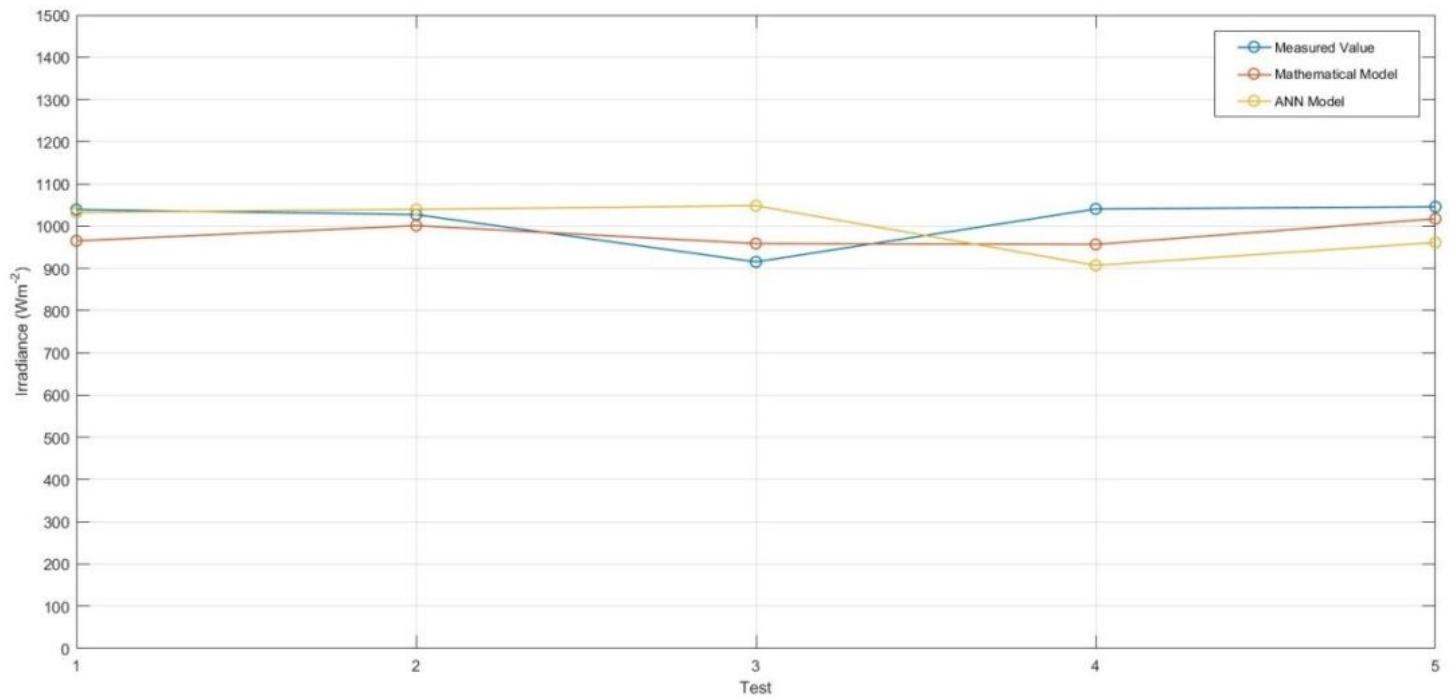
1. Mubiru, J., (2011) *Advances in Artificial Neural Systems*, **2011**, 1-6.
2. Premalatha, N. and Arasu, A.V., (2016) *Journal of Applied Research and Technology*, **14**, 206-214.
3. Wijayapala, W.D.A.S. and Kushal, D.H.K., (2015) *Engineer: Journal of the Institution of Engineers, Sri Lanka*, **XLVIII**, 19-25.
4. NASA, (2015) *HI-SEAS Mission IV Research*, **2015**.

## Figures



## Figure 1

Regression graph of the training model.



## Figure 2

Variation of the solar radiation with measured values.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [formula.docx](#)