Daily global solar radiation estimation using artificial intelligence approach

Y T Olorunfemi[1]*, O B Wojuola[1] and J A Adesina[2, 3]

[1] Dept. of Physics, Faculty of Natural and Agricultural Sciences, NWU, Mafikeng Campus, South Africa.
[2] School of Geo- and Spatial Science, Unit for Env. Sc. and Management, NWU, Potchefstroom Campus, South Africa.
[3] EPPEI Specialisation Centre for Emissions Control, NWU, Potchefstroom Campus, South Africa.
*E-mail: olorunfemiyetundet@gmail.com

Abstract. Daily global solar radiation (DSR) is sparsely measured in meteorological stations in South Africa. The prediction of DSR is crucial to solar energy conversion systems (modelling, design and operation) and decision-making of potential energy policies. The need for these solar system designs varies from the use of power and water supply for industrial purposes to agricultural and domestic services. This paper employed the use of three artificial intelligence models, which are artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS), and support vector regression (SVR) in predicting DSR from the Capes of South Africa using NASA satellite data for 30 years. Daily values of minimum and maximum temperatures, relative humidity, precipitation, wind speed, atmospheric pressure and earth's temperature were used as independent variables and the solar radiation as dependent variable when training the model. Statistical metrics such as root mean squared error (RMSE), mean absolute percentage error (MAPE) and coefficient of determination were used to evaluate the model performance. The result shows that the ANFIS algorithm has the least MAPE values ranging from 9.82 – 13.18 and lowest RMSE values ranging 0.70 – 0.78 which outperforms the SVR model (MAPE ranging 11.53 – 14.89 and RMSE values ranging 0.78 – 0.94) and ANN model (MAPE ranging 11.41 – 27.34 and RMSE values ranging 0.79 – 1.57), making it a better technique for estimating solar radiation.

1. Introduction

The potential contribution of renewable energy technology to power generation is enormous. The building of solar energy systems in a specific area is highly dependent on knowledge about that region's sun radiation [1]. High-quality measurements done by pyranometers are the best sources of data on daily global solar radiation (DSR) . Unfortunately, the high price, maintenance, and calibration criteria of these devices, as well as the necessity for a skilled professional make it challenging to measure in many regions [2].

Artificial intelligence (AI) has become highly popular in virtually all technical areas in recent decades as a result of technological advancements [3]. Numerous AI approaches, such as deep learning (DL), support vector machine (SVM), artificial neural networks (ANN), kernel nearest neighbour (k-NN), genetic algorithms (GA), adaptive neuro-fuzzy inference system (ANFIS), and others, have begun to be widely used models [4]. Previous research has found that AI algorithms have several benefits over traditional modelling. The benefits include the capacity to handle huge volumes of outliers from dynamic and unpredictable systems. This provides more precise outcomes than empirical models when it comes to predicting solar radiation.

A review of the literature revealed that no study has investigated and compared the accuracy of support vector regression (SVR), ANFIS and ANN, incorporating the earth's temperature as one of the inputs in DSR estimation in the Capes of South Africa. The goal of this research is to study and to compare the performance of these techniques in modelling DSR from three locations in the Capes of South Africa. Comparisons and performance assessments are carried out based on various commonly used statistical indicators.

2. Methodology

2.1. Areas of case study and data source

Daily measurements of atmospheric pressure, minimum and maximum temperatures, wind speed, precipitation, relative humidity, earth's temperature and solar radiation for a period of thirty years (1990-2020) for selected cities in the Western Cape, Eastern Cape, and Northern Cape of South Africa were considered in this study. These measurements were obtained from the NASA satellite in RETScreen Expert Software. Each model used in this study was trained using data for the period of January 1990 to September 2011,while data for October 2011 to December 2020 were used for testing. The cities' geographical coordinates and locations in a geospatial map are presented in Table 1 and Figure 1 respectively.

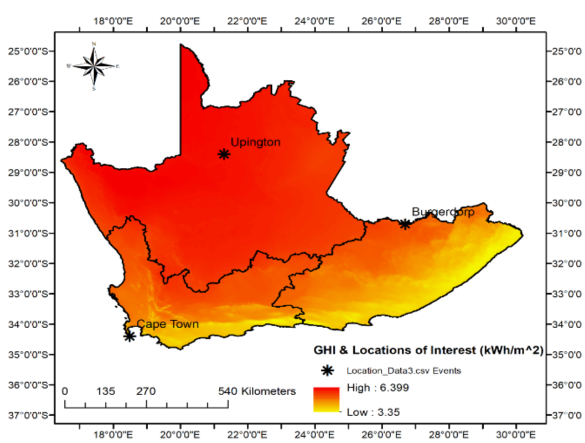

Figure 1: The geospatial mapping of the Cape cities (South Africa) used in this study.

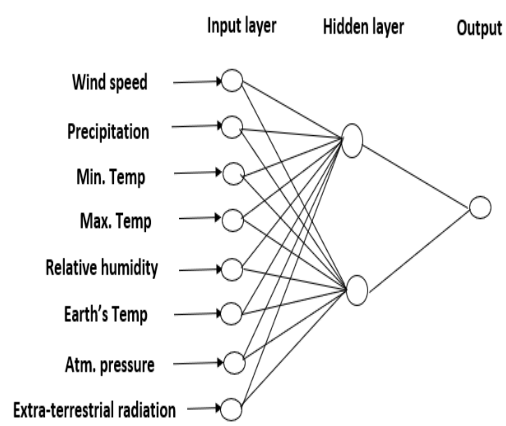

Figure 2: Schematic diagram of ANN model for the present work.

Table 1. Climate data locations.

Province	Location	Latitude (°S)	Longitude (°E)	Altitude (m)
Western Cape	Cape point	34.4	18.5	42
Eastern Cape	Burgersdorp	30.7	26.7	1518
Northern Cape	Upington	28.4	21.3	838

2.2. Support vector regression

The support vector regression (SVR) used for modelling the DSR for each chosen city was built according to a supervised learning approach. Meteorological variables formed the inputs to the model, and the solar radiations constituted the models' output.

In the SVR model, the Gaussian function was used as the kernel function for model training.The Gaussian kernel with an automatic kernel scale was used, and this selection was due to the efficiency of the Gaussian function:

$$K(x, y) = (x^T y + c)^d, \tag{1}$$

where x and y are feature vectors derived from training or testing data, and c is a constant that optimizes the effect of higher-order against lower-order polynomial components.

2.3. Artificial Neural Network

A back-propagation ANN was utilized to generate the intended output properly with an optimal network topology. Figure 2 above shows the schematic diagram of the ANN model.

2.4. Adaptive Neuro-Inference System

The ANFIS model combines ANN with the Fuzzy Inference System (FIS) to provide the best membership function distribution from feedback mapping [5]. The fuzzy layer, product layer, normalized layer, de-fuzzy layer, and total output layer in layers one to five respectively make up a typical ANFIS network [5, 6], as can be seen in Figure 3 and in equations (2) – (7).

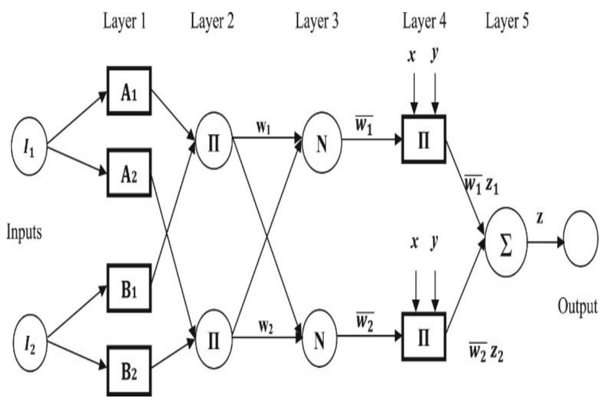

Figure 3: ANFIS structure

$$O_m^1 = \mu_{A_m}(x), m = 1,2 \tag{2}$$

$$O_m^1 = \mu_{B_m}(y), m = 1,2 \tag{3}$$

$$O_m^2 = w_m = \mu_{A_m}(x) \, . \, \mu_{B_m}(y), m = 1,2 \tag{4}$$

$$O_m^3 = \overline{w}_i = \frac{w_m}{w_1 + w_2}, m = 1,2 \tag{5}$$

$$O_k^4 = \overline{w}_i(p_m x + q_m y + r_m), m = 1,2 \tag{6}$$

$$O_m^5 = \sum_m \overline{w}_i \, z_m =, m = 1,2 \tag{7}$$

Figure 3 shows the fuzzy rules of the first order Takagi-Sugeno fuzzy model, which have the following structure:

Rule 1: if I_1 is A_1 AND I_2 is B_1 then $f_1 \ = p_1 I_1 + q_1 I_2 + r_1$

Rule 2: if I_1 is A_2 AND I_2 is B_2 then $f_2 \ = p_2 I_1 + q_2 I_2 + r_2$

where x and y are the crisp inputs to node, $A_m$and B_m are the fuzzy sets in the antecedent, f_i is the output inside the fuzzy area defined by the fuzzy rule, m is the adaptive node, $\overline{w}_i$ is the normalized third layer firing strength, O_m^i is the output of adaptive node m in layer $i = 1, 2, \ldots, 5$ and p_m, q_m, and r_m are the design criteria decided during the training phase.

2.5. Evaluation of the model's performance

Mean absolute percentage error and root mean square error were used for evaluating the performance of the models. The mean absolute percentage error (MAPE) reflects the mean absolute percentage variation between observed and expected values, whereas the root mean square error (RMSE) compares the variance between predicted and the actual data to assess the model's accuracy. The RMSE always has a positive value. The coefficient of determination (R^2) is a metric that indicates the strength of the linear connection between the predicted and measured values. It is important to note that lower MAPE and RMSE values indicate greater precision in the global solar radiation prediction, and in an ideal situation, they are both zero. The R^2 is a number that ranges from 0 to 1. The presence of a perfect linear relationship is indicated by R^2 values around 1. The performance of the model was evaluated using the MAPE, RMSE, and R^2 as expressed in the following equations:

$$MAPE = {}^{[\sum(s_i - t_i)]}/_{n} \tag{6}$$

$$RMSE = \left[\sum {}^{(s_i - t_i)^2}/_{n}\right]^{\frac{1}{2}} \tag{7}$$

$$R^2 = 1 - \frac{\sum(s_i - t_i)^2}{\sum(t_i)^2 - \frac{\sum(t_i)^2}{n}} \tag{8}$$

where $s_i$and t_i are observed and predicted values, respectively, n is the observation number. The MLP model script for this study was written using MATLAB (R2021a) software.

3. Results

Figure 4-6 depict the scattering diagrams of measured DSR estimated using three AI models for the location of Burgersdorp, Upington, and Cape Point located in the South African provinces of Eastern Cape (EC), Northern Cape (NC), and Western Cape (WC) respectively. The prediction accuracy of DSR differed significantly among the various model types. Based on the statistical values of the three AI models, the ANFIS model has the least MAPE values ranging from 9.82 – 13.18, and least RMSE values ranging from 0.70 -0.78 , if compared to the other models. However, the SVR model has the highest R^2 value of 0.89 amidst the three AI models, with a better model forecast ranging from 85.11% to 88.47 % when compared to the ANN model.

SVR model (MAPE ranging 11.53-15.52, RMSE ranging 0.78 - 0.94 and R^2 ranging 0.77 - 0.89) performed better than the ANN model (MAPE ranging 11.41-27.34, RMSE ranging 0.79 – 1.57 and R2 ranging 0.60 - 0.83). In general, the ANFIS model outperformed the other two AI models in the testing stage, with reduced data dispersion and better fitting of predicted data to actual values, especially in the Northern Cape (Upington).

The results demonstrate the ability of SVR, ANFIS, and ANN models to adapt to existing conditions in the Capes of South Africa. The lowest value of MAPE and highest value of the coefficient of determination at the study's sites are comparable those found by other researchers [7, 8].

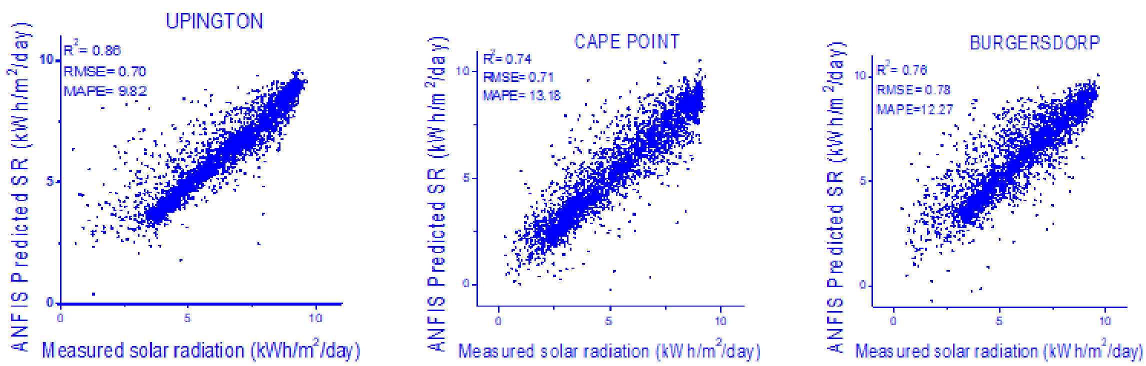

Figure 4: Scattering diagrams of measured DSR estimated using ANFIS model for Upington, Cape point and Burgersdorp.

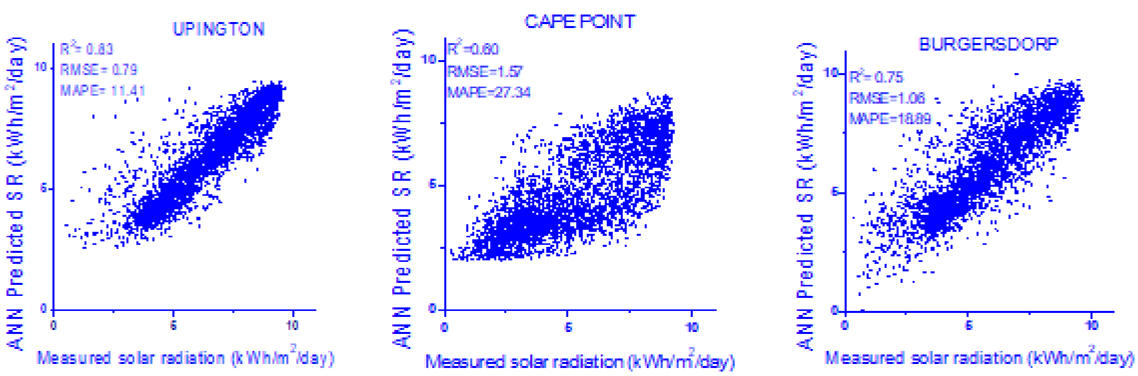

Figure 5: Scattering diagrams of measured DSR estimated using ANN model for Upington, Cape point and Burgersdorp.

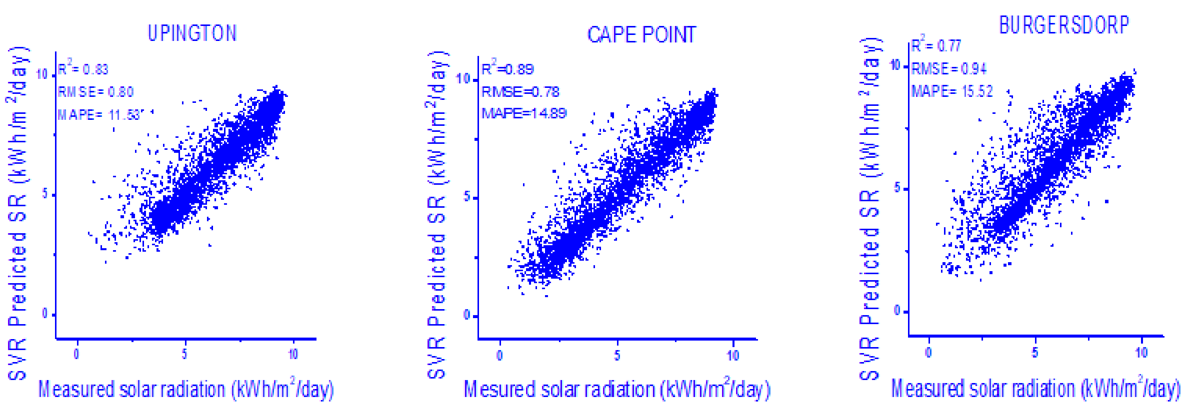

Figure 6: Scattering diagrams of measured DSR estimated using SVR model for Upington, Cape point and Burgersdorp.

Table 2. Model performance evaluation

	SVR			ANFIS			ANN		
	WC	EC	NC	WC	EC	NC	WC	EC	NC
R^2	0.89	0.77	0.83	0.74	0.76	0.86	0.60	0.75	0.83
RMSE	0.78	0.94	0.80	0.71	0.78	0.70	1.57	1.06	0.79
MAPE	14.89	15.52	11.53	13.18	12.27	9.82	27.34	18.89	11.41

4. Conclusions

The possibility of employing three artificial intelligence models (artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS), and support vector regression (SVR) to predict the horizontal daily global solar radiation using meteorological information was investigated in this study. Three cities in South Africa's Cape that serve as the country's solar centres were chosen as case studies. Aside from using the models for forecasting the daily global solar radiation, the three artificial intelligence models were further verified and compared. Statistical metrics, (MAPE, RMSE, and R^2, were used to evaluate the models' performance. Findings show that ANFIS give the best estimation of daily global solar radiation.

References

[1] Bakirci K, Kirtiloglu Y. Effect of climate change to solar energy potential: a case study in the Eastern Anatolia Region of Turkey. Environmental Science and Pollution Research. 2021:1-14.

[2] Yıldırım HB, Çelik Ö, Teke A, Barutçu B. Estimating daily Global solar radiation with graphical user interface in Eastern Mediterranean region of Turkey. Renewable and Sustainable Energy Reviews. 2018;82:1528-37.

[3] Kuo P-H, Huang C-J. A green energy application in energy management systems by an artificial intelligence-based solar radiation forecasting model. Energies. 2018;11(4):819.

[4] Long H, Zhang Z, Su Y. Analysis of daily solar power prediction with data-driven approaches. Applied Energy. 2014;126:29-37.

[5] Jang J-S. ANFIS: adaptive-network-based fuzzy inference system. IEEE transactions on systems, man, and cybernetics. 1993;23(3):665-85.

[6] Adedeji P, Madushele N, Akinlabi S, editors. Adaptive Neuro-fuzzy Inference System (ANFIS) for a multi-campus institution energy consumption forecast in South Africa. Proceedings of the International Conference on Industrial Engineering and Operations Management Washington DC, USA, September 27-29; 2018.

[7] Alizamir M, Kim S, Kisi O, Zounemat-Kermani M. A comparative study of several machine learning based non-linear regression methods in estimating solar radiation: Case studies of the USA and Turkey regions. Energy. 2020;197:117239.

[8] Quej VH, Almorox J, Arnaldo JA, Saito L. ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment. Journal of Atmospheric and Solar-Terrestrial Physics. 2017;155:62-70.