Environmental and Climate Technologies
2024, vol. 28, no. 1, pp. 422–436
https://doi.org/10.2478/rtuect-2024-0033
https://content.sciendo.com

An Empirical Approach to Solar Photovoltaic Cell Temperature Prediction

Kudzanayi CHITEKA[1], Christopher ENWEREMADU[2*]

[1,2]*Department of Mechanical Engineering, University of South Africa, Science Campus, Florida 1710, South Africa*

Received 15.03.2024; accepted 12.09.2024

Abstract **– Solar cell temperature is critical in the determination of solar energy generated by a solar photovoltaic power plant. High temperatures are associated with a reduction in the energy generated and hence prediction of photovoltaic cell temperature is essential in temperature mitigation and solar energy forecasting, especially in commercial power plants. The present study focused on the development of a hybrid machine learning based predictive model for solar photovoltaic cell temperature prediction in solar photovoltaic arrays. A physical experimental set up was developed to measure solar cell temperature under different weather and other related parameters. Satellite data were also collated for these parameters and were used to compliment experimental data used in this study. Satellite data used in the study were statistically transformed to mimic experimentally measured data. Feature selection and dimensionality reduction were performed to reduce the input variables and maintain relevant data in the modelling process. A solar cell temperature predictive model based on selected weather parameters was developed using a machine learning approach (Random Forests), and parameters used were selected from the statistical analysis. The prediction accuracy of the developed model was analysed using the coefficient of determination (R^2) and the Mean Absolute Percentage Error (MAPE). The results indicated a higher model performance compared to generic models used in cell temperature prediction. The prediction MAPE for the developed model was 0.08 % while an R^2 value of 0.99 was obtained which was indicative of a good model. The developed model was also comparable to other contemporary models developed to predict solar photovoltaic cell temperature. Simulations were also done to determine the annual energy generated with the incorporation of the solar cell temperature prediction model. The results revealed an average of 25.52 % daily energy difference between a simulation which considered solar cell temperature and that which ignored solar cell temperature.**

Keywords **– Cell temperature; empirical approach; experimental analysis; predictive model; Solar Photovoltaics.**

Nomenclature		
A_c	Solar module collector area	m^2
G	Global Horizontal Irradiance	W/m^2
G_t	Irradiance on a tilted solar PV collector	W/m^2
h_w	Wind convection coefficient	W/m^2K
$h_{w,NOCT}$	Wind convection coefficient at NOCT	W/m^2K
K	Clearness Index	–

* Corresponding author.
E-mail address: enwercc@unisa.ac.za

k	Ross coefficient	Km^2/W
H	Number of observations	–
P	Precipitation	mm
P_{max}	PV module power	W
P_S	Surface pressure	mmH_g
RF	Random Forests	–
R_H	Relative humidity	%
T_a / T_{2M}	Ambient temperature	°C
T_{act}	Actual measured cell temperature	°C
T_C	Solar cell temperature	°C
T_{pred}	Predicted temperature	°C
T_{ref}	Reference temperature	°C
v_w	Wind speed close to the PV module surface	m/s
W_{D10}	Wind Direction at 10m	°
W_{S10}	Wind Speed at 10m	m/s
W_{S2}	Wind Speed at 2m	m/s
A	Solar absorptance of the PV module	–
β_{ref}	Temperature coefficient of maximum power	$°C^{-1}$
δ_c	Packing factor	–
η_e	Solar PV module electrical efficiency	%
η_{ref}	Reference PV module efficiency	%
T	Solar transmittance of PV module	%

1. INTRODUCTION

Solar energy is an important source of clean and alternative energy required to complement efforts being made to mitigate global warming and environmental degradation caused by fossil fuel-based energy systems. Much attention has been placed on global warming and its mitigation through the development and utilization of alternative energy sources, solar energy in particular, is an important environmentally friendly energy source [1]–[3]. The importance of solar energy is verified by studies which show that the solar irradiance received in one hour is enough to meet annual global energy requirements [4], [5]. Solar energy utilization is done using two main forms i.e., solar thermal and solar photovoltaics (PV). Solar PV performance is affected by three main variables which are irradiance, temperature and soiling [6], [7]. In addition variables such as humidity, wind, altitude, barometric pressure, and installation configuration also contribute to the performance of a solar PV array [8], [9].

Temperature is an important factor in solar cell performance and has a linear reduction performance which can result in losses amounting to as much as 50 % due to heating [10], [11]–[13]. Studies have indicated that elevated temperatures reduce cell performance by affecting both power output and efficiency of a solar PV power plant [14], [15].

In Brazil, studies have shown that non-temperature controlled PV modules can actually degrade at a 3 % faster rate and result in 30 % lower efficiency [16]. Temperature has an effect of modifying both the power output and the resulting efficiency of a solar cell and this is all dependent on the solar cell material, the absorption characteristics, the thermal dissipation characteristics and the environmental conditions under which the solar cell is operating within. Furthermore, high temperatures also result in degradation of the solar cell and thus its life is shortened [17]. An increase in cell temperature results in a slight increase in the short circuit current I_{sc} while the open circuit voltage V_{oc} is significantly reduced resulting in an overall reduction in power generated [18], [19].

Temperature prediction is a topical issue in solar PV modelling and studies have been carried out to predict solar cell temperature and hence ascertain the performance of a solar PV system. Temperature predictions have been made in literature following the need to understand the different variables that affect the performance of solar PV modules. For example, Waithiru *et al.* [20] developed statistical PV temperature prediction models for floating PVs. The statistical multiple linear regression-based models developed had an error range between 2 % and 4 %. Du *et al.* [21] developed a heat transfer model to predict the temperature performance of a PV module based on the real operating ambient conditions. This study revealed that solar irradiance was the main variable affecting PV module temperature while the current played a smaller role in increasing the PV module temperature.

Machine Learning (ML) and Artificial Intelligence (AI) are important tools explored in the prediction of solar PV cell temperature and studies have proved their importance and demonstrated the good accuracy involved in their use [22]–[24]. For example, Serrano-Lujan *et al.* [25] studied building integrated photovoltaics and developed a temperature prediction model using AI algorithms on poly crystalline PV modules namely; differential evolution and grammatical evolution. In all cases, the developed models showed relative errors of less than 4 % and there was an 11 % reduction in error compared to the Sandia model. Ceylan *et al.* [10] developed an Artificial Neural Networks (ANN) based prediction model to predict PV module temperature. The model used solar radiation data and ambient temperature as input variables during the training of the ANN model.

Temperature is thus the second most important parameter affecting solar PV performance and hence it requires greater attention. To improve the performance of PV cells from the current recorded efficiencies of 25 % [26], a closer look at the variables affecting its performance is required and these variables include temperature [27]. Studies have indicated that cell temperature is one of the most important factors affecting solar PV output yet it is also one of the least studied parameters [28]. Accurate forecasting of solar output which is dependent on solar irradiance and temperature is a critical area that needs attention for successful PV integration [29], [30]. If temperature is properly predicted, the expected performance of a solar PV power plant can be determined in advance. In as much as it is important to ascertain the temperature effect on PV module output, it is also important to predict the solar cell temperature which in turn affects PV module output.

While temperature mitigation approaches, passive or otherwise, have been proposed for solar PV, it is still important to be able to ascertain the amount of heating under varying meteorological parameters. Most empirical prediction approaches discussed in literature are mainly location dependent and hence cannot be applied to other geographical locations. On the other hand, analytical models which are generic, have a weakness in their accuracy which makes solar cell temperature prediction not as accurate as it should be. Studies have also shown that, the use of Random Forests (RF) which has been demonstrated to be a superior and high accuracy algorithm is still limited in solar cell temperature predictions. The current study therefore focuses on developing a hybrid model based on an analytical model and adapted to a Machine Learning algorithm (Random Forests) to develop a more accurate location dependent solar cell prediction model. Such hybrid models in the prediction of solar cell temperature are also still yet to be widely adopted. The rest of this manuscript is organized as follows, section 1 introduces the concept of solar cell temperature effects and its prediction. Section 2 outlines the research approach used in the study while section 3 provides the results and discussions followed by conclusions in section 4.

2. Materials and Methods

2.1. Location, Data Collection and Feature Selection

Meteorological data was downloaded from NASA Data Access Viewer (https://power.larc.nasa.gov/data-access-viewer/) for UNISA science campus located at –26.18°S, 27.85°E Roodepoort, South Africa. The downloaded data were daily averaged values for Solar Irradiance (G), Relative Humidity (R_H), Wind Speed at 2 m (W_{S2}), Wind Speed at 10 m (W_{S10}), Wind Direction at 10 m (W_{D10}), Ambient Temperature (T_a / T_{2M}), Surface Pressure (P_S), Clearness Index (K) and Precipitation (P). The downloaded data was cleaned using Python v3.0. Missing values were replaced with average values and values outside the acceptable limits e.g., $K > 1$ were deleted from the data set. Analytical prediction of solar cell temperature was done using the collected data based on the Skoplaki cell temperature model shown by Equation (1). Where; T_C is solar cell temperature, T_a is the ambient temperature, T_{ref} is the reference temperature equal to 25 °C, G is the solar irradiance (W/m^2), h_w is the *w*ind convection coefficient given by Equation (3), v_w is the wind speed close to the module surface (m/s), $h_{w,NOCT}$ is the wind convection coefficient at Nominal Operating Cell Temperature (NOCT) W/m^{2}K), τ is solar transmittance of PV module (%), α is the solar absorptance of the PV module (%) and the product $\tau\alpha$ is assumed to be 0.9. β_{ref} is the temperature coefficient of maximum power (°C^{-1}), k is the Ross coefficient (Km^2/W) and is given by Equation (2) and its value ranges between 0.02 and 0.05 while η_{ref} is the reference PV module efficiency.

$$T_C = T_a + kG\frac{h_{w,NOCT}}{h_{w,(v)}}\left[1 - \frac{\eta_{ref}}{\tau\alpha}\left(1 - \beta_{ref}T_{ref}\right)\right] \tag{1}$$

$$k = \frac{T_{c,NOCT} - T_{a,NOCT}}{G_{NOCT}} \tag{2}$$

$$h_w = 5.7 + 3.8v_w \tag{3}$$

Feature selection was done through parameter correlation analysis which was applied for dimensionality reduction. This procedure was performed using a correlation matrix to assess the relationships among the different parameters as well as their relationship to the computed cell temperature. Parameters which were correlated to each other had one of them eliminated to reduce the dimensionality of the model. This was done in preparation for the machine learning modelling phase to avoid complexity and overfitting of the model which leads to poor performance caused by too many variables in the model.

2.2. Experimental Set Up

An experimental set up was established on the same location to measure solar cell temperature (T_c). The experimental setup used an array of 4 ground mounted EnerSol solar panels with specifications provided in Table 1. The solar panels had an installation azimuth of 180° and a tilt of 26° as shown in Fig. 1(a). The solar cell temperature was measured at one-hour intervals from 06:30 in the morning to 19:30 in the evening using an Xi400 Optris infrared thermometer pictured in Fig. 1(b). The procedure was repeated for a period of 2 months from the 1st of January 2024 to the 29th of February 2024. The period was chosen for experimentation because it corresponds to the peak of summer in the southern hemisphere, where the study was conducted. During this time, solar irradiance levels are typically high, providing optimal conditions for assessing the performance and thermal behaviour of the solar panels. Additionally, this period is long enough to ensure that the data collected is representative of varying environmental conditions within a summer period. The two-month duration allows for the capture of sufficient data to analyse trends and draw meaningful conclusions.

(a) (b)

Fig. 1. Experimental components: (a) the solar panels; (b) Optris Xi400 Infrared Camera.

In this study, solar PV cell electrical efficiency η_c was determined using Equation (4) [10]. After considering other parameters including PV glass cover transmissivity (τ), solar cell absorptivity (α) and the packing factor (δ_c), the PV module efficiency was computed as shown in Equation (5). The values of τ, α and δ_c were respectively taken as 0.90, 0.95 and 0.90 [31], [32]. The PV module power generated was computed by equating equation 5 and equation 6 where P_{max} is the PV module power generated, A_c is the collector area and G_t is the irradiance on a tilted solar PV collector.

TABLE 1. SOLAR MODULE PARAMETERS

	Parameter	Value
1	Rated maximum power, P_{max}	335 W
2	Open circuit voltage, V_{oc}	41.56 V
3	Maximum power voltage, V_{mp}	34.04 V
4	Short circuit current, I_{sc}	10.39 A
5	Maximum Power Current, I_{mp}	9.84 A
6	Module Efficiency, η	20.4 %
7	Temperature Coefficient of maximum power, β_{ref}	–0.5 %/K
8	Nominal Operating Cell Temperature, NOCT	48±2 °C

$$\eta_c = \eta_{ref}[1 - \beta_{ref}(T_c - 25)] \tag{4}$$

$$\eta_m = \eta_c \alpha \tau \delta_c \tag{5}$$

$$\eta_m = \frac{P_{max}}{A_c G_t} \tag{6}$$

2.3. Machine Learning (ML) Model

The ML model was developed using Random Forests (RF) implemented in Python V3.5. Following dimensionality reduction, a minimum number of parameters were selected for the final modelling process using RF. RF represents a collection of decision trees and when an object needs to be classified, each tree provides a classification, and the chosen classification is that which is provided by most of the decision trees. Studies have shown that RF applies ensemble learning to boost the accuracy of classification where a problem is solved using a multiple number of models and thus this algorithm has high accuracy and superiority [33]. These multiple models are more accurate when compared to using single classifiers in the ensemble. Breiman [34] noted that a RF is thus a classifier made up of a collection of tree structured classifiers $\{h(\mathbf{x}, \Theta_k, k = 1...)\}$ in which the $\{\Theta_k\}$ are independent identically distributed random vectors where each tree has one vote to select the most popular class at input **x**. With the tree number increasing, $\Theta_1, ...PE^*$ converges to Eq. (7). Hence, as the number of trees increases, RF do not overfit but rather converges to a limiting value of the generalization error.

$$P_{X,Y}\left(P_\Theta\left(h\left(\boldsymbol{X},\Theta\right)=Y\right)-\max_{j\neq Y} P_\Theta\left(h\left(\boldsymbol{X},\Theta\right)=j\right)<0\right) \tag{7}$$

Random forests have become a favourable tool in classification problems compared to other tools such as linear regression in that linear regression assumes linearity which makes the model simple but not as flexible enough for prediction. As such RF can easily handle nonlinearities and hence have better prediction accuracy in prediction involving large data sets [35]. In this study, data was randomly split into training and testing set respectively in a ratio of 75:25.

2.4. Data Transformation

The meteorological data utilized in the study were downloaded from the NASA Data Access Viewer (https://power.larc.nasa.gov/data-access-viewer/) and included the variables outlined in Section 2.1. This data was transformed to align with the ground-measured values. To address the potential discrepancies between the predicted solar cell temperatures and the ground-measured experimental values, a regression analysis was conducted. This analysis aimed to develop a model that correlates the predicted values (T_{pred}) with the actual measured values (T_{act}). The resulting model, as presented in Eq. (8), allows for the transformation of T_{pred} into T_{act}, ensuring that the predicted values more accurately reflect the ground-measured experimental data.

$$T_c = f\left(T_{pred}, T_{act}\right) \tag{8}$$

2.5. Validation

In addition to experimental validation, the models developed in this study were validated using the HOMER software temperature model given by Eq. (9) and a comparison was also made to a model by Amiry *et al.* [36]. The prediction accuracy was analysed by means of the coefficient of determination (R^2) and the Mean Absolute Percentage Error (MAPE) as outlined by Eq. (10)–Eq. (11). Where $\tau\alpha$ is taken as 0.9. n is the number of observations; T_{act} is the actual measured cell temperature and T_{pred} is the temperature predicted by the developed model.

$$T_{\mathrm{c}} = \frac{T_{\mathrm{a}} + \left(T_{\mathrm{c,NOCT}} - T_{\mathrm{a,NOCT}}\right)\left(\frac{G}{G_{\mathrm{NOCT}}}\right)\left[1 - \frac{\eta_{\mathrm{ref}}\left(1 - \beta_{\mathrm{ref}} T_{\mathrm{ref}}\right)}{\tau\alpha}\right]}{1 + \left(T_{\mathrm{c,NOCT}} - T_{\mathrm{a,NOCT}}\right)\left(\frac{G}{G_{\mathrm{NOCT}}}\right)\left(\frac{\beta_{\mathrm{ref}}\eta_{\mathrm{ref}} T_{\mathrm{ref}}}{\tau\alpha}\right)} \tag{9}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{T_{\mathrm{act}} - T_{\mathrm{pred}}}{T_{\mathrm{act}}}\right| \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(T_{\mathrm{act}} - T_{\mathrm{pred}}\right)^2}{\sum_{i=1}^{n}\left(T_{\mathrm{act}} - \overline{T_{\mathrm{act}}}\right)^2} \tag{11}$$

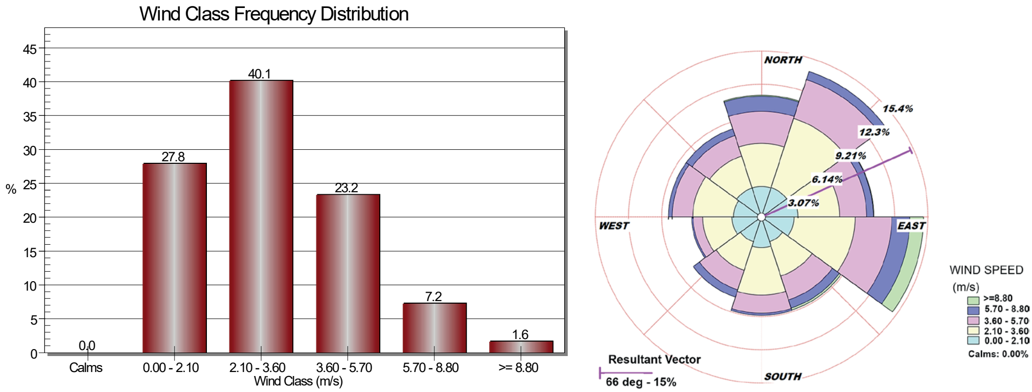
3. Results and Discussions

3.1. Parameter Analysis

The study was carried out at UNISA Florida campus in Roodepoort South Africa and Table 2 shows the average values for the meteorological parameters used. Of note, the average solar irradiance is 5.59 kW/m^2 while average values for ambient temperature, wind speed at 2 m and relative humidity were respectively 17.37 °C, 2.41 m/s and 57.92 %.

TABLE 2. AVERAGE ANNUAL VALUES OF METEOROLOGICAL PARAMETERS

	Parameter	Value
1	Global Horizontal Irradiance, G	5.59 kW/m^2
2	Clearness index, K	0.61
3	Wind speed at 2 m (W_{S2})	2.41 m/s
4	Relative humidity, R_{H}	57.92 %
5	Precipitation, mm	1.43 mm
6	Wind direction at 10 m (W_{D10})	161.97°
7	Wind speed at 10m (W_{S10})	3.60 m/s
8	Pressure, P	84.99mm (Hg)
9	Temperature at 2 m (T_{a} / T_{2M})	17.37 °C



(a) (b)

Fig. 2. (a) Wind class frequency distribution, (b) wind rose.

The analysis of the wind characteristics shows that the dominant wind speed is between 2.1 m/s and 3.6 m/s occurring 40.1 % of the time followed by wind speeds between 0 m/s and 2.1 m/s occurring 28.7 % of the time as shown in Fig. 2(a). As shown in Fig. 2(b), the resultant wind direction is 66° i.e., ENE direction. However, there is more wind blowing from the NE and the ESE directions. These winds could be playing a critical role in reducing the apparent ambient temperatures experienced in the vicinity of the solar panels used in the study. The strongest winds come from ESE where wind speeds greater than 8.80 m/s are recorded.

3.2. Parameter Correlation

Parameter correlation was carried out and the relationships between the parameters used in the study were established and analysed. Parameter correlation (shown in Figs. 3 and 4) revealed that cell temperature (T_c) and ambient temperature (T_a) have a very strong positive correlation of 0.88. This indicates that in the absence of other interfering variables, as the ambient temperature increases, the solar cell temperature is expected to proportionately increase as well. On the other hand, pressure and ambient temperature (T_a) measured at 2 m above the ground have a good negative correlation of −0.63. This is an indication that the impacts of pressure on cell temperature can be fairly represented by T_a. Furthermore, as indicated in Fig. 4, there was a very strong correlation between solar irradiance, G and T_c. As solar irradiance increases, T_c is also expected to increase since upon reaching the solar cell surface, a large percentage of the energy is converted into heat. Furthermore, when G reaches the earth's surface, it is converted into heat thus raising the ambient temperature which in turn also affects the cell temperature.

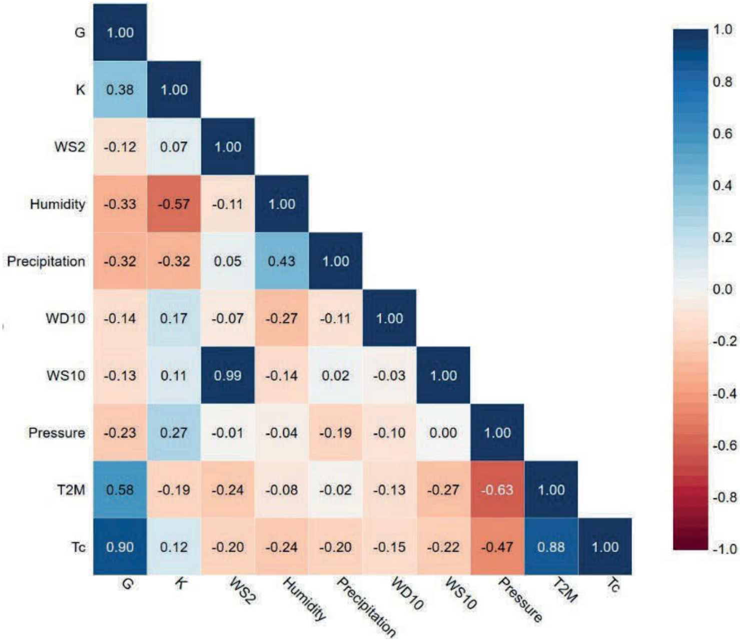


Fig. 3. Parameter correlation matrix.

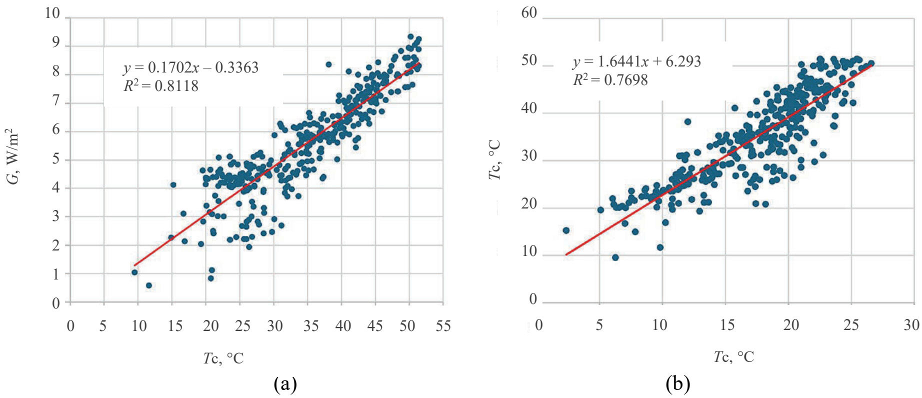


Fig. 4. Correlation analysis of (a) T_c and G, and (b) T_c and T_a.

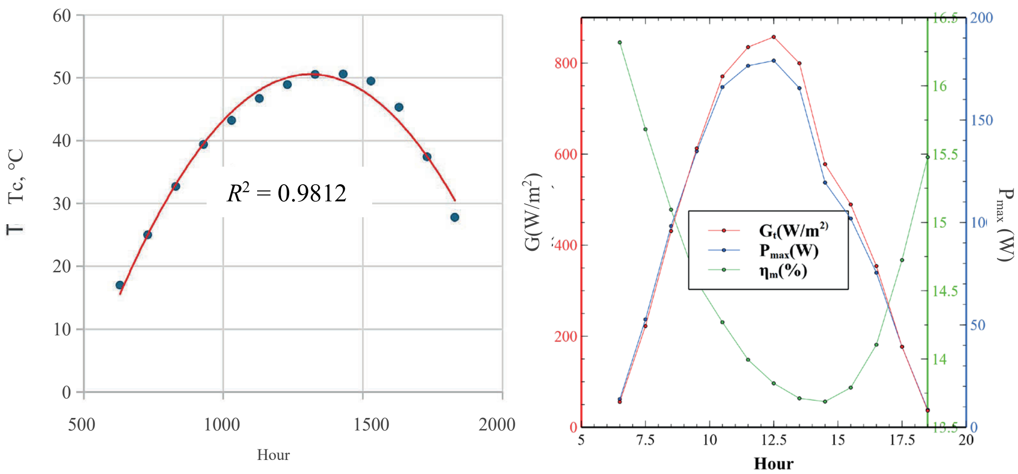
Clearness index (K) and Relative Humidity (R_H) were also found to have a moderate negative correlation of −0.57. In effect, when the clearness index increases, it shows the presents of less humidity in the environment. This is because an increase in K is an indication of less particulate matter in the atmosphere and this includes R_H. Except for G, T_a and P, T_c had a weak correlation with the other parameters analysed in the study. The study also revealed that wind speed measured at 10 m above the ground could be well represented by wind speed measured at 2 m above the ground since there is a very strong correlation of 0.99 between the two parameters. The study thus showed that solar irradiance and ambient temperature are the two most important parameters with a positive correlation with cell temperature and such conclusions were also made in literature [37], [38].

3.3. Solar Array Performance Analysis

The ground measured solar cell temperature was plotted against the hour of the day as shown in Fig. 5(a). The relationship between cell temperature with respect to the time of the day was found to be a quadratic polynomial shown by Eq. (12) and had an R^2 of 0.98.

$$y = -80.19 + 19.54x - 0.73x^2 \tag{12}$$

The measured solar cell temperature was used to compute solar cell efficiency (η_c) and solar module efficiency (η_m). These values were used to determine the maximum power from each of the 335 W solar modules used in the study. Fig. 5(b) shows the graphs of η_m (green), P_{max} (blue) and G_t (red) in different colors. The study indicates that the impact of temperature is more pronounced between the 10th and the 15th hour of the day during the time which both the solar cell and module efficiencies are significantly reduced due to high temperatures experienced on the solar cells. The study suggests that at high irradiance levels, temperatures are also high leading to reduced solar cell efficiencies. Fig. 5(b) also revealed that there is a lag between the maximum irradiance and the maximum reduction in solar cell efficiency. This could be attributed to temperature build up in the solar PV module leading to maximum solar cell temperatures occurring between the 14th and the 15th hour which corresponds to the period of maximum solar cell efficiency reduction.



(a) (b)

Fig. 5. (a) Hour of the day vs. solar cell temperature; (b) solar array performance analysis.

3.4. Machine Learning

A Random Forest (RF) based approach was used to develop a cell temperature T_c prediction model based on the parameters assessed in the study. As a validation for the two parameters selected through parameter correlation in section 3.2, parameter ranking was performed during the RF implementation in *Python v3.0* and the algorithm also selected T_a and G as the most important parameters as shown in Fig. 6. Variable G had an importance of 0.85 while variable T_a had an importance of 0.13, and all the other parameters had a combined importance of 0.02. This highlights the importance of solar irradiance in solar cell temperature prediction where also ambient temperature is dependent on G. This justifies the use of G and T_a as the major parameters to be considered in predictive modelling with G being the main variable and T_a the minor variable. However, their combination has a very significant contribution to the accuracy of the T_c prediction model.

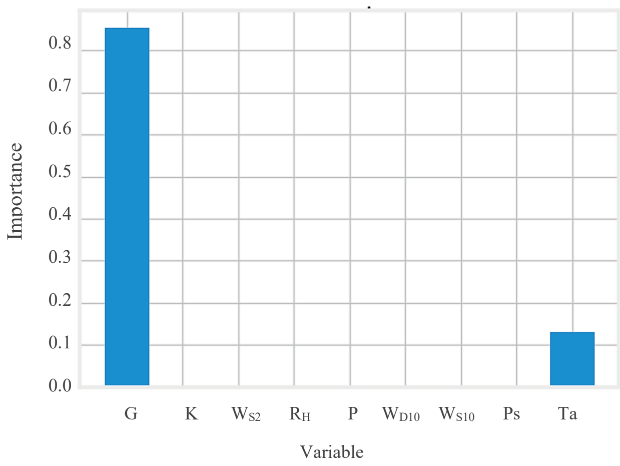
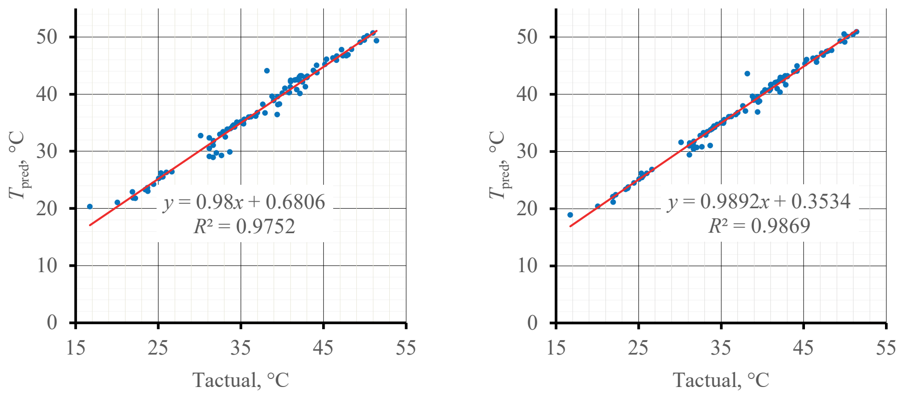


Fig. 6. Variable importance.



(a) (b)

Fig. 7. Prediction model (a) with all parameters used, and (b) with G and T_a used.

For comparative purposes and to justify the two selected parameters, two models were developed in which one model used all the parameters in this study and the other used only T_a and G for the prediction of T_c. The model with all parameters, although acceptable with an R^2 of 0.98, was not as accurate as the model which considered G and T_c which had R^2 of 0.99 which was a 1.2 % performance improvement as shown in Fig. 7. This justifies the selection of the best parameters since it has an impact of reducing the time and computational resources required in the data collection and modelling. The performance improvement on the model using T_a and G could be attributed to the deliberate avoidance of overfitting and multicollinearity which come with the use of too many parameters, some of which could be correlated. The Mean Absolute Error (MAE) and the Mean Absolute Percentage Error

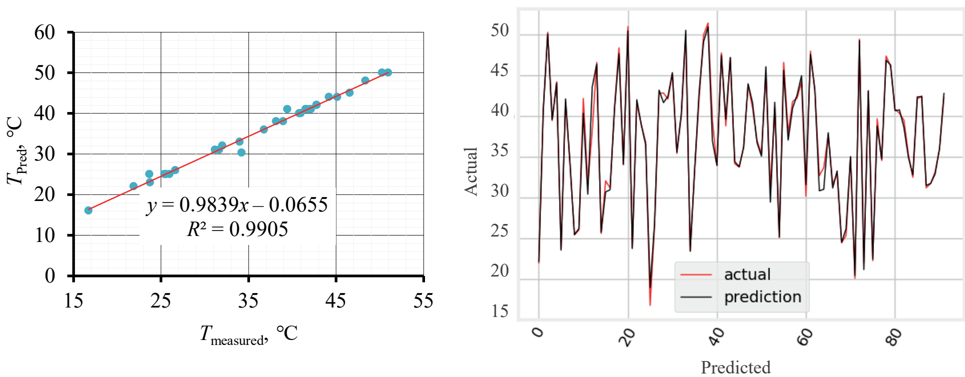
(MAPE) were respectively 0.54 °C and 2.41 % for the model using G and T_a as the input parameters.

3.5. The Hybrid Model

The RF prediction model was refined to enhance prediction accuracy. Regression analysis, illustrated in Fig. 8(a) and detailed in Eq. (13), led to the development of a modifying function. This function was applied to transform the T_c values predicted by the RF model, resulting in more accurate estimates that closely align with ground-based T_c measurements. Results from RF were regressed with measured values of T_c to ascertain the expression needed for transforming the RF based results to mimic measured values. The resulting transformed model was almost a simple linear translation of −0.1 such that $y(x) \rightarrow y(x) - 0.10025$. The results showed a performance improvement of 0.13 % from the RF model to the hybrid model. Due to the relatively smaller improvement in the accuracy with R^2 equal to 99.05 % for the hybrid model, the study has demonstrated that RFs are highly accurate in predictive modelling. The resulting hybrid model had a Mean Absolute Percentage Error (MAPE) of 0.08 % and an R^2 of 99.05 % which, however, shows a very significant reduction in the MAPE from 0.54 % to 0.08 % indicating an improvement of 85.2 %. The study revealed that although there was no significant improvement in the R^2, there was a significant reduction in the prediction error which made the model even more accurate. This thus implies that, when the hybrid model is used, prediction errors are very minimal when compared to the stand-alone RF model.

The resulting expression to transform RFs output to ground measured output is represented by Eq. (13) and Fig. 8(b), which is an illustration of the performance of the hybrid model in which the actual ground measured values were compared to the hybrid Random Forest – Regression model used in this study. The results have indicated that there is no significant difference between the actual ground measured values and the predicted values using the hybrid model. These results were compared to the HOMER model, and it was found to perform better by 3.5 %.

$$y = 0.9839x - 0.0655 \tag{13}$$



(a) (b)

Fig. 8. (a) Modified prediction model, (b) actual vs predicted temperature.

4. Conclusions

- In this study, a hybrid photovoltaic cell temperature prediction model was developed based on random forests and regression analysis. Meteorological parameters were selected using parameter correlation and the random forests algorithm. The following conclusions were made from this study.
- Ambient temperature and solar irradiance are the most important parameters in predicting solar photovoltaic cell temperature.
- Elimination of parameters such as wind speed, relative humidity and surface pressure do not significantly affect the model developed.
- Random forests are an important tool in predictive modelling and can provide models with high accuracy.
- Hybrid models are more accurate than stand-alone models with prediction accuracy as high as 99.05 % and are thus recommended for use in situations demanding higher prediction accuracies.

One limitation of this study is the use of data from a single location (UNISA science campus) and a specific two-month experimental period. Despite this, the results are robust due to the comprehensive data collection and rigorous analysis methods employed. The chosen period coincides with peak summer conditions, which provides a representative snapshot of high solar irradiance and temperature variations. Additionally, the findings are supported by strong correlations between key parameters and validated through multiple models, including a refined hybrid model with improved accuracy. This approach ensures that the results offer reliable insights into solar cell temperature behaviour, even within the study's constraints.

Further research could explore extending the study over multiple seasons and locations to enhance the generalizability of the predictive models. Incorporating high-resolution data collection and additional environmental factors, such as pollutants and shading effects, could refine prediction accuracy. Investigating alternative machine learning algorithms and real-time monitoring systems could also improve model performance. Additionally, studying the economic impact of these models on solar energy systems and examining the effects of emerging technologies on solar cell temperature could provide valuable insights for optimizing performance and efficiency.

References

[1] Ding Q., Huang J., Chen J., Luo X. Climate warming, renewable energy consumption and rare earth market: Evidence from the United States. *Energy* 2024:290:130276. https://doi.org/10.1016/j.energy.2024.130276

[2] Blumberga D., Chen B., Ozarska A., Indzere Z. Lauka D. Energy, Bioeconomy, Climate Changes and Environment Nexus. *Environmental and Climate Technologies* 2019:23(3):370–392. https://doi.org/10.2478/rtuect-2019-0102

[3] Kostevica V., Dzikevics M. Bibliometric Analysis of the Climate Change Impact on Energy Systems. *Environmental and Climate Technologies* 2023:27(1):950–963. https://doi.org/10.2478/rtuect-2023-0069

[4] Khamisani A. A., Liu D. P. P., Cloward D. J., Bai D. R. Design Methodology of Off-Grid PV Solar Powered System (A Case Study of Solar Powered Bus Shelter).

[5] Liu Q., Yu G., Liu J. J. Solar Radiation as Large-Scale Resource for Energy-Short World. *Energy Environ.* 2009:20(3):319–329. https://doi.org/10.1260/095830509788066466

[6] Iheanetu K. J. Solar Photovoltaic Power Forecasting: A Review. *Sustainability* 2022:14(24):17005. https://doi.org/10.3390/su142417005

[7] Bodnár I., Matusz-Kalász D., Koós D. Experimental and numerical analysis of solar cell temperature transients. *Pollack Periodica* 2021:16(2):104–109. https://doi.org/10.1556/606.2020.00260

[8] Shaik F., Lingala S. S., Veeraboina P. Effect of various parameters on the performance of solar PV power plant: a review and the experimental study. *Sustainable Energy Research* 2023:10(1):6. https://doi.org/10.1186/s40807-023-00076-x

[9] Jathar L. D., *et al.* Comprehensive review of environmental factors influencing the performance of photovoltaic panels: Concern over emissions at various phases throughout the lifecycle. *Environ. Pollution* 2023:326:121474. https://doi.org/10.1016/j.envpol.2023.121474
[10] Ceylan İ., Erkaymaz O., Gedik E., Gürel A. E. The prediction of photovoltaic module temperature with artificial neural networks. *Case Stud. Therm. Eng.* 2014:3:11–20. https://doi.org/10.1016/j.csite.2014.02.001
[11] Schiro F., Benato A., Stoppato A., Destro N. Improving photovoltaics efficiency by water cooling: Modelling and experimental approach. *Energy* 2017:137:798–810. https://doi.org/10.1016/j.energy.2017.04.164
[12] Ansari E., Akhtar M. N., Othman W. A. F. W., Abu Bakar E., Alhady S. S. N. Numerical Investigation of Thermal Efficiency of a Solar Cell. *Applied Sciences* 2022:12(21):10887. https://doi.org/10.3390/app122110887
[13] Parthiban R., Ponnambalam P. An Enhancement of the Solar Panel Efficiency: A Comprehensive Review. *Front. Energy Res.* 2022:10. https://doi.org/10.3389/fenrg.2022.937155
[14] Cheraghizade M., Jamali-Sheini F. Photovoltaic behavior of SnS solar cells under temperature variations. *Optik* 2022:254:168635. https://doi.org/10.1016/j.ijleo.2022.168635
[15] Wei Z. *et al.* Understanding the temperature sensitivity of the photovoltaic parameters of perovskite solar cells. *Solar Energy* 2023:264:112040. https://doi.org/10.1016/j.solener.2023.112040
[16] Piotrowski L. J., Simões M. G., Farret F. A. Feasibility of water-cooled photovoltaic panels under the efficiency and durability aspects. *Solar Energy* 2020:207:103–109. https://doi.org/10.1016/j.solener.2020.06.087
[17] Kersten F. *et al.* Degradation of multicrystalline silicon solar cells and modules after illumination at elevated temperature. *Solar Energy Materials and Solar Cells* 2015:142:83–86. https://doi.org/10.1016/j.solmat.2015.06.015
[18] Taghinia A., Yazdi F., Fazel P., Anousheh S. N., Davoudi K. G. Comparison of single junction GaAs and In0.2Ga0.8N based solar cells at various temperatures. *Energy Procedia* 2012:14:919–924. https://doi.org/10.1016/j.egypro.2011.12.1033
[19] Liao W., Heo Y., Xu S. Evaluation of Temperature Dependent Models for PV Yield Prediction. [Online]. [Accessed 18.09.2021]. Available: https://www.semanticscholar.org/paper/Evaluation-of-Temperature-Dependent-Models-for-PV-Liao-Heo/a232d5d270cfdbef9feef9e603e64ba3c314d59d
[20] Kamuyu W. C. L., J. Won L. C., Ahn H. Prediction Model of Photovoltaic Module Temperature for Power Performance of Floating PVs. *Energies* 2018:11(2):447. https://doi.org/10.3390/en11020447
[21] Du Y., Tao W., Liu Y., Jiang J., Huang H. Heat transfer modeling and temperature experiments of crystalline silicon photovoltaic modules. *Solar Energy* 2017:146:257–263. https://doi.org/10.1016/j.solener.2017.02.049
[22] Vijaykumar R., Rudramoorthy R., Mangalore A. R. Prediction of Solar PV Panel Temperature Using Mathematical Models and Artificial Neural Networks. *J. Comput. Theor. Nanosci.* 2017:14(10):4986–4997. https://doi.org/10.1166/jctn.2017.6909
[23] Coskun C., Koçyiğit N., Oktay Z. Estimation of pv module surface temperature using artificial neural networks. *Mugla J. Sci. Technol.* 2016:2(2). https://doi.org/10.22531/muglajsci.283611
[24] Motuzienė V., Bielskus J., Lapinskienė V., Rynkun G. Office Building's Occupancy Prediction Using Extreme Learning Machine Model with Different Optimization Algorithms. *Environ. Clim. Technol.* 2021:25(1):525–536. https://doi.org/10.2478/rtuect-2021-0038
[25] Serrano-Luján L., Toledo C., Colmenar J. M., Abad J., Urbina A. Accurate thermal prediction model for building-integrated photovoltaics systems using guided artificial intelligence algorithms. *Applied Energy* 2022:315:119015. https://doi.org/10.1016/j.apenergy.2022.119015
[26] Jošt M. *et al.* Perovskite Solar Cells go Outdoors: Field Testing and Temperature Effects on Energy Yield. *Adv. Energy Mater.* 2020:10(25):2000454. https://doi.org/10.1002/aenm.202000454
[27] Meng Q. *et al.* Effect of temperature on the performance of perovskite solar cells. *J. Mater. Sci. Mater. Electron.* 2020:32:12784–12792. https://doi.org/10.1007/s10854-020-03029-y
[28] Khaledi P., Behboodnia M., Karimi M. Simulation and Optimization of Temperature Effect in Solar Cells CdTe with Back Connection Cu_2O. *Int. J. Opt.* 2022:e1207082. https://doi.org/10.1155/2022/1207082
[29] Zhang C., Zhang Y., Su J., Gu T., Yang M. Performance prediction of PV modules based on artificial neural network and explicit analytical model. *J. Renew. Sustain. Energy* 2020:12(1):013501. https://doi.org/10.1063/1.5131432
[30] Paulescu M. *et al.* Online Forecasting of the Solar Energy Production. *Ann. West Univ. Timisoara – Phys.* 2018:60(1):104–110. https://doi.org/10.2478/awutp-2018-0011
[31] Mishra R., Tiwari G. Energy and exergy analysis of hybrid photovoltaic thermal water collector for constant collection temperature mode. *Solar Energy* 2013:90:58–67. https://doi.org/10.1016/j.solener.2012.12.022
[32] Dubey S., Solanki S. C., Tiwari A. Energy and exergy analysis of PV/T air collectors connected in series. *Energy Build.* 2009:41. https://doi.org/10.1016/j.enbuild.2009.03.010
[33] Fawagreh K., Gaber M. M., Elyan E. Random forests: from early developments to recent advancements. *Syst. Sci. Control Eng.* 2014:2(1):602–609. https://doi.org/10.1080/21642583.2014.956265
[34] Breiman L. Random Forests. *Mach. Learn.* 2001:45(1):5–32. https://doi.org/10.1023/A:1010933404324
[35] Schonlau M., Zou R. Y. The random forest algorithm for statistical learning. *Stata J. Promot. Commun. Stat. Stata* 2020:20(1):3–29. https://doi.org/10.1177/1536867X20909688

[36] Amiry H. *et al.* Assessment of improved models for predicting PV module temperature and their electrical performance in a semi-arid coastal region. *Int. J. Green Energy* 2023:20(14):1584–1596. https://doi.org/10.1080/15435075.2023.2166788

[37] Gholami A. *et al.* Impact of harsh weather conditions on solar photovoltaic cell temperature: Experimental analysis and thermal-optical modeling. *Solar Energy* 2023:252:176–194. https://doi.org/10.1016/j.solener.2023.01.039

[38] Du Y. *et al.* Evaluation of photovoltaic panel temperature in realistic scenarios. *Energy Convers. Manag.* 2016:108:60–67. https://doi.org/10.1016/j.enconman.2015.10.065