Contents list available at CBIORE journal website

**International Journal of Renewable Energy Development**

Journal homepage: https://ijred.cbiore.id

Research Article

# Evaluating the performance of the Anwaralardh photovoltaic power generation plant in Jordan: Comparative analysis using artificial neural networks and multiple linear regression modeling

Suhaib Ibrahim Alma'asfa[a], Feras Younis Fraige[a,b*], Mohd Sharizal Abdul Aziz[c], Chu Yee Khor[d], Laila A. Al-khatib[e]

[a] *Mechanical Engineering Department, Faculty of Engineering, Al-Hussein Bin Talal University, Ma'an, Jordan.*

[b] *Mining & Minerals Engineering Department, Faculty of Engineering, Al-Hussein Bin Talal University, Ma'an, Jordan.*

[c] *School of Mechanical Engineering, Engineering Campus, Universiti Sains Malaysia, 14300 Nibong Tebal, Seberang Perai Selatan, Penang, Malaysia.*

[d] *Faculty of Mechanical Engineering & Technology, Universiti Malaysia Perlis (UniMAP), 02600 Arau, Perlis, Malaysia.*

[e] *Environmental Engineering Department, Faculty of Engineering, Al-Hussein Bin Talal University, Ma'an, Jordan.*

**Abstract**. The global energy demand is rising, driven by population growth, economic development, and industrialization. Shifting towards renewable energy, like solar energy, is gaining momentum worldwide because of ecological concerns and resource depletion. This paper aims to utilize Artificial Neural Networks (ANNs) and multiple linear regression (MLR) modeling techniques to evaluate the productivity of 11 MW photovoltaic (PV) solar power plant currently operational in Jordan. The case study reveals that both models can be used to predict the daily, monthly, and yearly average power produced and system efficiency with reasonable accuracy. The ANN model exhibited promising results, where the best value for the coefficient of determination ($R^2$) and mean absolute percentage error (MAPE) for training were 95.85% and 0.59%, respectively. However, $R^2$ was 93.7%, and MAPE was 1.27% for validation tests. All these results were achieved using a 7-6-1 model, with a sample ratio of 1:1 for the data allocated in training and validation. When using multiple linear regression, the $R^2$ and standard error values were 93.42% and 0.17%. On the other hand, the results showed that the yearly output power for actual and predicted by both models over the year was 24,399 MWh, 24,538 MWh, and 24,401 MWh, respectively. This research showed valuable results in the monthly output power for solar cells at the Anwaralardh PV power system project, contributing to a better understanding of solar energy generation in arid desert climates and emphasizing the potential of solar power plants to play a crucial role in achieving SDG 7 objectives.

**Keywords:** Artificial Neural Networks, photovoltaic (PV) solar power plant, multiple linear regression, solar energy



Received: 11th Feb 2024; Revised: 5th April 2024; Accepted: 25th April 2024; Available online: 1st May 2024.

## 1. Introduction

In recent decades, the world has witnessed a substantial surge in energy demand, propelled by population growth, economic expansion, shifts in lifestyle, globalization, modernization, and technological progress. In the past, electricity was heavily generated using fossil fuels such as oil, gas, and coal. These finite sources are depleted with time and cause environmental challenges (Adaramola *et al.,* 2014; Shafiee and Topal, 2009). Renewable energy sources have become a favorable alternative to traditional ones with reduced environmental and global warming challenges (Erten *et al.,* 2022; Ali *et al.,* 2019). The global movement towards renewable energy is gaining momentum, driven by growing ecological concerns, including global warming and air pollution (Erten *et al.,* 2022). Hence, choosing renewable energy, such as solar energy, is more favorable for mitigating the impacts of concerns.

The power produced from solar energy using photovoltaic systems has proven its practicality and usefulness worldwide. The new installation of power generation plants is evolving worldwide (Ahmed *et al.,* 2020; Alshafeey and Csaba, 2019; Devaraj *et al.,* 2021). For example, it increased exponentially from 15 GW in 2010 to about 250 GW in 2022, an increase of more than 16 folds in just 12 years (Wikipedia, 2023). The power production from PV systems depends on environmental and operational conditions, geographical location, and system design and configuration. Energy production strongly depends on the solar modules' irradiance (Antonanzas *et al.,* 2016; Keddouda *et al.,* 2023). Meteorological and environmental parameters, such as weather condition variations, dust, shading, and temperature, can disturb the received irradiance (Antonanzas *et al.,* 2016, Nespoli *et al.,* 2018; Keddouda *et al.,* 2023). Wind speed can affect the heat transfer from the modules and, hence, their overall efficiency. This imposed variability on

* Corresponding author
Email: ferasfraige@gmail.com (F. Fraige)

https://doi.org/10.61435/ijred.2024.60156
ISSN: 2252-4940/© 2024.The Author(s). Published by CBIORE

the energy generation from these systems represents a challenge in the integrated electricity network. Precise predictions enable grid operators to accommodate changes more efficiently (Nespoli *et al.,* 2018).

Modeling power generation from PV systems makes it necessary to consider the various variables affecting system performance and provide reasonable estimates (Ahmed *et al.,* 2020; Devaraj *et al.,* 2021). An accurate assessment of PV power generation is crucial for optimizing the economic efficiency of the power system and contributes to the stability and reliability of the electric grid (Li *et al.,* 2020; Natsheh and Samara, 2019). Predicting PV power generation can also reduce operational costs and improve electricity markets (Tsai *et al.,* 2023). It helps identify the urban form for effective PV systems deployment (Poon and Kaempf, 2019; Cavalcante *et al.,* 2021). Finally, forecasting PV power generation improves the safety and stability of the power system, especially during extreme weather conditions (Fu *et al.,* 2023).

Among several modeling techniques, Artificial Neural Network (ANN) is a valuable tool that ensures the highest accuracy when predicting PV power system performance (Mellit and Pavan, 2010; Kalogirou, 2007; Elsheikh *et al.,* 2019; Leva *et al.,* 2017). ANNs can model nonlinear relationships between input and output parameters with substantial datasets (Tu, 1996, Shehab *et al.,* 2022; Kumar *et al.,* 2016; Loutfi *et al.,* 2017). Through training, ANNs can learn from data and adapt to changing environments, making them suitable for handling complex problems that are difficult to solve using traditional methods (Ahmad *et al* 2022). However, ANN is usually considered the "black box" and it is difficult to comprehend their results (Koeppe *et al.,* 2021). Another popular statistical approach in modeling PV power generation is the multiple linear regression (MLR) model (EL-AAL *et al.,* 2023; Yadav *et al.,* 2022; Thombare *et al.,* 2022; Limouni *et al* 2022). In comparison with ANN, the MLR model is relatively simple and easy to implement. With the help of statistical tests, MLR can reveal which parameters most affect the output.

The primary objective of the study is to investigate the efficiency of the PV solar power system at the Anwaralardh Solar Energy Generation Company, Jordan, in 2021. Two models are employed to predict the system performance under varying operating and environmental conditions. Ambient temperature, cell temperature, horizontal and inclined irradiations, wind speed, and reference yield are monitored, and their effect on energy produced is estimated. Considering these factors and applying advanced modeling tools can make reasonably accurate predictions of the PV system's power generation. Besides, the practical implication of the current study is that by monitoring and considering these factors and applying advanced modeling tools, it is possible to make reasonably accurate predictions of a PV solar power system's energy production. This knowledge is valuable in optimizing the performance of PV solar systems, leading to more efficient energy generation and management for solar generation power plants, which aligns with the United Nation's (UN's) sustainable development goals, particularly in SDG 7, Affordable and clean energy.

## 2. Methodology

Jordan is located in the Middle East, with abundant solar resources. The country is poor in its traditional energy resources. Due to growing energy demands and plans to reduce greenhouse gas emissions and minimize dependence on fossil fuels (Kartikasari *et al.,* 2023), Jordan is expanding its renewable energy production sector in general and solar energy in particular (Al Naimat and Liang, 2023; Albaali *et al.,* 2022; Kartikasari *et al.,* 2023). Ma'an province is rich in its solar and wind energy resources. This study collected data from a solar farm belonging to 'Anwaralardh Solar Energy Generation Company' in Ma'an Solar Power Park, Jordan, with coordinates 30.152N and 35.814E, as shown in Figure 1. The location has an altitude of 1000 m, and the area is described as an arid to semiarid climate with relatively hot summers and mild winters. The site is suitable for PV power generation because of its abundant sunlight throughout the year, about 350 W/m$^2$. The average yearly meteorological and environmental parameters are shown in Table 1. In the scope of this study, the following data was collected from the project, as illustrated in Figure 2, including ambient temperature ($T_a$), cell temperature during daylight ($T_c$), horizontal irradiation ($G_h$), inclined irradiation ($G_i$), wind speed (WS), and reference yield ($Y_f$), and energy produced. Daily averaged data was collected for an entire year from 1 January 2021 to 31 December 2021. The nominal production capacity is estimated at 10.5 MW from thirty-six thousand solar modules equipped with a sun-tracking system. The rated maximum power of the module is 305 W. Module specifications used in this location are summarized in Table 2.

### *2.1 Data Calculations*

In this study, the values of daily PV energy efficiency (η) are calculated by Equation (1) (Abdallah *et al.,* 2004).

$$\eta = \frac{E}{A \times H} \tag{1}$$

**Table 1**
Average yearly values for environmental factors.

| Environment parameter | Value |
|---|---|
| Environment temperature (°C) | 18.1 |
| Incline irradiation (W/m$^2$) | 350 |
| Accumulated incline irradiation (kWh/m$^2$) | 17.1 |
| Wind speed (m/s) | 3.4 |

View A View B View C

**Fig 1.** Aerial views of Anwaralardh Solar Energy Generation Company Site in Ma'an – Jordan (Anwaralardh 2023).

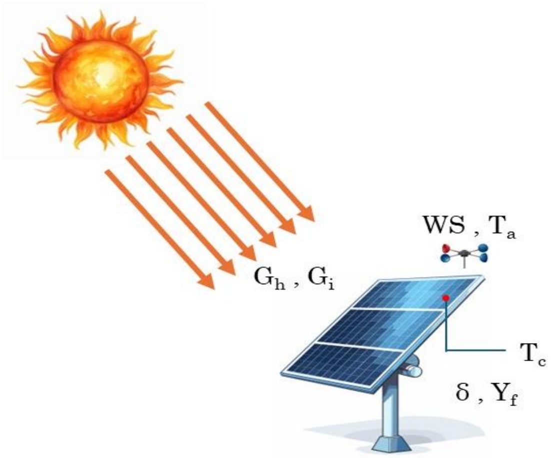


**Fig 2.** Schematic diagram showing the input variables included in this study. Ambient temperature ($T_a$), cell temperature during daylight ($T_c$), horizontal irradiation ($G_h$), inclined irradiation ($G_i$), wind speed (WS), and reference yield ($Y_f$).

**Table 2**
PV panel specifications.

| Specification | |
|---|---|
| Module Type | BYD305P6C-36 |
| Module Application Class | Class A |
| Rated Maximum Power (W) | 305 |
| Nominal Operating Cell Temperature (°C) | 45°C ± 2 |
| Weight (kg) | 22.4 |
| Dimension (mm) | 1956×992×40 |

where A is the area of the modules in $m^2$, and E is the total energy produced by PV cells per day in (kWh/day). H is the total daily irradiation impinging PV surface in (kWh/($m^2$ day)). When measured total energy ($E_m$) is used, a measured efficiency ($\eta_m$) is calculated. However, when the predicted total energy ($E_p$) is used, a predicted efficiency ($\eta_p$) is estimated.

The monthly energy produced ($E_t$) is calculated by summing the energy produced per day based on Equation (1) over a month period. Or as defined in Equation (2):

$$E_t = \sum_{month} H \times A \times \eta_p \quad (2)$$

Each hidden neuron in the network needs a specific input term, $n_j$. This term is created from a set of connection links called synapses. The network calculates the product of the weights and their respective inputs (P), and this is summed for each neuron by adding an offset value (b), also known as bias. The simplified expression for this process is represented as Equation (3) (Reyes-Téllez *et al.,* 2020):

$$n_j = Wi_{(j,1)} . P_{(1)} + Wi_{(j,2)} . P_2 + .... + Wi_{(j,r)} \, . P_r + b_{(1,j)} \quad (3)$$

where *Wi* represents the coefficients of the connection weights between the input layer and hidden layer, P is the input variables, j is the number of neurons in the hidden layer, r is the number of input neurons, and b is the bias corresponding to each neuron in the hidden layer.

The weights and bias values are updated with minimal adjustments in the backpropagation process. In each iteration, small changes are made to the weight gradients. The sigmoid function is illustrated in Equation (4) (Rasamoelina *et al.,* 2020; Rajput *et al.,* 2021; Baccelli *et al.,* 2020):

$$\sigma(x) = \frac{1}{1+e^{-n_j}} \quad (4)$$

The declination angle (δ), also known as solar declination, varies between -23.45° and 23.45° during the year and can be calculated by using Equation (5) (Duffie *et al.,* 2020):

$$\delta = 23.45° * \sin\left(\frac{360*(284+d)}{365}\right) \quad (5)$$

Here, d represents the day number in the year.

## 3. Modeling of PV Energy Production

Modeling photovoltaic (PV) energy production involves predicting the amount of electricity generated by a solar PV system based on various operating factors. There are assorted approaches to modeling the produced energy, ranging from physical models, such as PV cell models, to statistical and machine learning models. Where the former models focus on the physics and performance of the cells, the latter use statistical regression and machine learning that can capture the complex nature of the system and the dynamic performance of monitored output. The current study investigates two models based on artificial neural networks and multiple linear regression.

### *3.1 Artificial Neural Network (ANN) Model*

ANN is a computational model based on the way biological neurons interact in the human brain. It is a subset of machine learning algorithms and is particularly powerful for tasks incorporating complex relationships with relatively large parameters, such as pattern recognition, classification, regression, and other complex data modeling tasks. In ANN models, the architecture is crucial as the number of layers, neurons in each layer, and the activation functions are assigned. This work employs a multi-layer feed-forward-backward propagation network. The input layer contains the seven operating parameters that influence energy generation in the output layer. Input parameters are processed through a series of weighted connections and activation functions in the hidden layers to generate the output. The input parameters in this work are ambient temperature, cell temperature during daylight, horizontal irradiation, inclined irradiation, wind speed, solar declination, and reference yield. Then, the predicted efficiency from the ANN model ($\eta_{p,\,ANN}$) is estimated. Figure 3 illustrates a schematic diagram of an ANN structure.

### *3.2 Multiple Linear Regression (MLR) Model*

MLR model is a statistical tool used extensively in science and engineering to estimate the relationship between multiple independent parameters and one dependent parameter (Freedman, 2009).

The formula for the MLR model can be expressed by Equation (6):

$$\eta_{p,\,MLR} = a_0 + a_1 X_1 + a_2 X_2 + \ldots + a_n X_n + \varepsilon \quad (6)$$

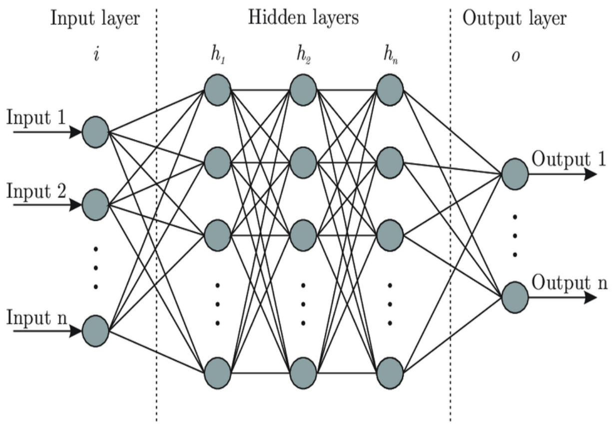


**Fig 3.** Schematic diagram of an ANN structure (Bre *et al.,* 2018).

$\eta_{p,\ MLR}$ is the predicted PV efficiency using the MLR model as a dependent variable, and $a_0$ is the intercept, which is the value of $\eta_{MLR}$ when all parameters are null. ( $a_1$ , $a_2$ , … $a_n$) are the regression coefficients of the corresponding investigated parameters ($X_1$ , $X_2$ , … $X_n$), while ε is the error term associated with the model (Freedman, 2009). Figure 4 illustrates a schematic diagram of a MLR structure.

### *3.3 Measures of Uncertainty*

In this study, two measures of uncertainty are employed. Firstly, the coefficient of determination, commonly represented as $R^2$, measures the strength of the linear relationship between two variables. The value of the correlation factor is calculated using Equation (7) (Montgomery *et al.,* 2021).

$$R^2 = 1 - (SSR / SST) \tag{7}$$

SSR is the sum of squared differences between the measured and predicted efficiencies from the model. And SST is the sum of the squared differences between the measured efficiency and the average value for measured efficiency. The value of $R^2$ is between 0 and 1, and the relationship between the investigated parameters is considered linearly strong when $R^2$ becomes closer to unity.

Secondly, the Mean Absolute Percentage Error (MAPE) is a metric commonly used to evaluate the accuracy of a predictive model. It is particularly useful in the context of time series analysis, where predictions are made for future time points. MAPE is calculated by using Equation (8) (Neville, 1978).

$$\mathrm{MAPE} = \frac{100}{n}\sum_{n=1}^{N}\left|\frac{\eta_m-\eta_p}{\eta_m}\right| \tag{8}$$

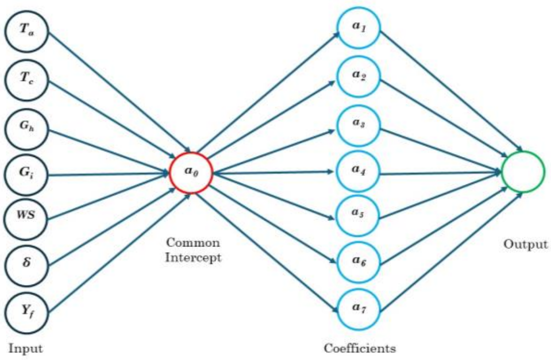


**Fig 4.** Schematic diagram of MLR structure (Choden *et al.* 2022 with modifications).

where n is each data point (1, 2, …, N).

### *3.4 Data Processing*

The collected data for the study period is inspected for any missing, irrelevant, or ambiguous information. This ensures that the dataset is clean and will not affect modelling predictions. Then, the input dataset is normalized to eliminate any scale variance, help obtain fast convergence, ensure numerical stability, and allow for generalization for new data. Input parameters are normalized to have values between 0 and 1 by using Equation (9):

$$\acute{X} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{9}$$

where X' is the normalized value of the parameter X that has minimum and maximum values ($X_{min}$ and $X_{max}$) respectively.

The dataset is then split into two parts to form a sample ratio for model training and validation. Different ratios of the dataset are incorporated to reveal its influence on model predictions, including 1:1, 2:1, and 1:2 for training and validation, respectively. Different complexity levels of ANN structures are employed to assess their performance and relative accuracy. The following primary models are used: model 7-2-1, model 7-6-1, model 7-9-1, model 7-4-3-1, and model 7-7-5-1.

## 4. Results and discussion

### *4.1 Input parameters variation.*

The seven parameters that may be attributed to energy production from the PV system in the Anwaralardh project are monitored. The variation of some of these parameters throughout the year 2021 is shown in Figure 5. Typical behaviour is noted for most parameters. Meteorological parameters such as ambient temperature and wind speed profoundly affect PV cell performance. Ambient temperature is higher in summer (June, July, and August) than in winter (December, January, and February). Wind speed is faster in winter than in summer, as illustrated in Figure 5 (a) and (b). Module temperature is critical as it affects the PV performance. It is well known that increasing cell temperature reduces PV efficiency. Module temperature follows a similar trend to ambient temperature, with the curve being shifted up, indicating greater temperature values, as illustrated in Figure 5(a).

The daily horizontal inclined and accumulated inclined irradiance are considered the main parameters responsible for energy production from PV systems. They are seasonally dependent, where changes in sunlight intensity and the effects of the sun-tracking system that will adjust the angle of solar cells are conditional, with higher values on hot sunny days and lower magnitudes on colder cloudy days, as indicated in Figure 5(c) and (d). They experience fluctuations and outliers, mainly due to clouds and dust affecting solar energy generation. The observed variations in the seven parameters influencing energy production from the Anwaralardh PV system in 2021 offer valuable insights for practical implications. Remarkably, the meteorological factors of ambient temperature and wind speed exhibit seasonal patterns, with higher temperatures during summer and increased wind speed in winter. These trends highlight the significance of considering seasonal variations in optimizing PV cell performance. Moreover, the critical impact of module temperature on PV efficiency is evident, with higher temperatures leading to reduced efficiency. This emphasizes

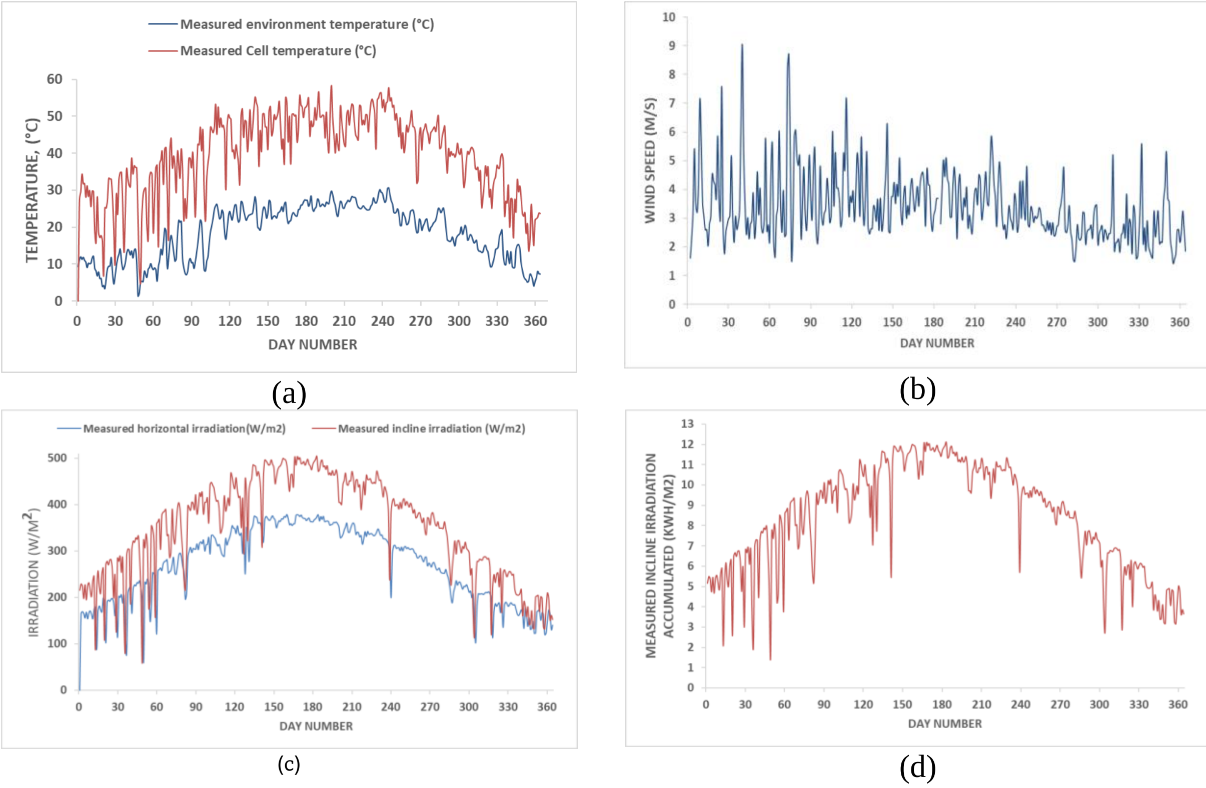


**Fig 5.** Variation of the input parameters. (a) ambient and cell temperatures, (b) wind speed, (c) horizontal and inclined irradiations, and (d) accumulated inclined irradiation.

the need for effective temperature management strategies in PV system design and operation.

*4.2 Optimum ANN Model and Sample Ratio*

The training and validation results for the different ANN architectures are presented in Table 3. The performance of ANN models is acceptable based on the calculated coefficient of determination and MAPE. Dataset division into training and validation portions according to different sample ratios shows interesting results. When using higher data portions in the training phase (sample ratio 2:1), the highest $R^2$ is observed for ANN 7-6-1, with relatively acceptable MAPE. However, the accuracy of the estimates deteriorated slightly with less values of $R^2$ and greater Mean Absolute Percentage Error (MAPE) in the validation phase. On the contrary, using fewer training data portions (sample ratio 1:2) gives higher values of $R^2$ and the lowest values of MAPE in the training phase. However, the models decline in the validation phase, indicating that the models are overfitted in the training stage and are not adequately generalized to unseen data. However, a sample ratio of 1:1 gives a better overall performance with higher accuracy and higher values of $R^2$ in the training and validation, with relatively lower MAPE for most model variations.

In terms of ANN structure suitability to predict the efficiency of the PV system, the reported results indicate that simpler models are better at predicting the investigated system. The 7-6-1 and 7-9-1 models show the best generalization for new data, with high $R^2$ values and relatively low MAPE for training and validation data.

However, models such as 7-4-3-1 and 7-7-5-1, with two hidden layers and more computing time, have not presented any improvement despite their complexity.

Dataset division according to sample ratio 1:1 gives better accuracy in terms of $R^2$ and MAPE. Among the tested ANN architectures, model 7-6-1 well predicts the monitored parameter. The testing and validation $R^2$ are 0.96 and 0.94, respectively. Accordingly, MAPE has relatively good results of 0.59 and 1.27 for testing and validation, respectively. So, it is adopted here to predict PV solar system performance.

The practical implication of the results is identifying an optimal ANN model and sample ratio for predicting the efficiency of the PV solar system. The results indicate that a sample ratio of 1:1 provides superior accuracy, achieving higher

**Table 3**
The training and validation results for the different ANN architectures with different sample ratios.

| ANN structure | Sample ratio 1:1 | | | | Sample ratio 2:1 | | | | Sample ratio 1:2 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Training | | Validation | | Training | | Validation | | Training | | Validation | |
| | $R^2$ | MAPE (%) | $R^2$ | MAPE (%) | $R^2$ | MAPE (%) | $R^2$ | MAPE (%) | $R^2$ | MAPE (%) | $R^2$ | MAPE (%) |
| 7-2-1 | 0.8850 | 1.01 | 0.8800 | 1.34 | 0.8794 | 1.15 | 0.9050 | 1.14 | 0.9470 | 0.77 | 0.8400 | 1.30 |
| 7-6-1 | 0.9585 | 0.59 | 0.9370 | 1.27 | 0.9354 | 0.89 | 0.8917 | 1.33 | 0.9610 | 0.35 | 0.8620 | 1.30 |
| 7-9-1 | 0.9407 | 0.71 | 0.9300 | 1.24 | 0.9340 | 0.99 | 0.8949 | 1.40 | 0.9690 | 0.30 | 0.8900 | 1.30 |
| 7-4-3-1 | 0.9038 | 0.88 | 0.8960 | 1.28 | 0.9090 | 1.03 | 0.8928 | 1.23 | 0.9490 | 0.38 | 0.8530 | 1.31 |
| 7-7-5-1 | 0.9190 | 0.81 | 0.9100 | 1.25 | 0.9200 | 0.93 | 0.9107 | 1.17 | 0.9530 | 0.36 | 0.8520 | 1.32 |

coefficient of determination values and lower MAPE for most model variations. The 7-6-1 ANN architecture is the most effective in predicting the monitored parameter, demonstrating high generalization for new data with $R^2$ values of 0.96 and 0.94 for testing and validation, respectively. Simplicity in ANN structure proves advantageous, as more complex models do not improve performance despite increased computational time. The practical recommendation is to use the 7-6-1 ANN model, especially with a 1:1 sample ratio, for accurate PV solar system performance predictions. It is an effective tool for accurately predicting the performance of PV solar systems.

Comparing the performance of the 7-6-1 ANN model with the literature also highlights its improvement in predicting power generation in terms of better MAPE. The power output of a PV system in Ashland was projected by using an ANN model with a network structure of 28-20-11. The average MAPE error was 7.16% across four predicting days (Kumar and Kalavathi, 2018). Nitisanon and Hoonchareon (2017) proposed models using ANNs to predict the output power of solar cells, and the value of MAPE was 5.85%. Whereas Altan et al. (2021) used ANN to predict PV power using three different models, the results show that the ANN models have higher accuracy with MAPE = 1.95%. Rumbayan et al. (2012) reported 3.4 % MAPE when using ANN with 9 meteorological input parameters to model global solar radiation in Indonesia. Rao et al. (2022) reported 1.888 % MAPE when using the ANN model. This concludes that the 7-6-1 ANN model used in this study outperforms the other models in terms of power generation prediction.

### *4.3 MLR results*

The Multiple Linear Regression (MLR) model is also used to predict the PV solar system performance. Table 4 summarizes the regression statistics. Based on these findings, the MLR model can be defined according to Equation (10):

$$\eta_{p,\,MLR} = 0.1262 + 0.0924\, Y_f + 0.0018\, \delta - 0.0007\, WS - 0.0042\, T_a - 0.0027\, T_c + 0.013\, G_h - 0.0027\, G_i \quad (10)$$

The predictions of PV efficiency using the MLR model are quite well throughout the study period. The model is statistically significant at a 5% level as the P value of the F-test is far below 0.05. The descriptive statistics from the regression output revealed that the ($R^2$) and the adjusted coefficient of determination are about 0.94 each, with an estimated standard error of about 0.17% for all observations conducted in 2021. Thus, the MLR model provides accurate and precise results, encouraging its use in predicting PV solar efficiency. Statistically, it is observed that all predictors are significant according to the probability value of the t-test with less dependence on the declination angle and wind speed. The robust performance of the MLR model in predicting PV solar system efficiency, as evidenced by the comprehensive regression statistics, carries significant practical implications for decision-makers and researchers. The MLR model provides a dependable tool for predicting PV efficiency, enabling informed decisions on system optimization and resource allocation. This reliability is crucial for maximizing energy production and ensuring optimal system performance. The statistical significance of all predictors, with reduced dependence on declination angle and wind speed, enhances the model's applicability across varying environmental conditions. Besides, these results offer a valuable foundation for researchers to explore and refine predictive models. Understanding the model's strengths and ability to perform well across diverse conditions will guide future research in identifying variables critical to predictive accuracy. It also suggests opportunities for developing more models considering specific environmental factors and system dynamics, contributing to ongoing advancements in predictive solar energy modelling.

**Table 4**
Output data from multiple linear regression model using ANOVA for predicting PV system efficiency.

| **Multiple R** | **R Square** | **Adjusted R Square** | **Standard Error** |
|---|---|---|---|
| 0.9665 | 0.9342 | 0.9329 | 0.0017 |

**F- test**

| **Source** | **df** | **SS** | **MS** | **F** | **Significance F** |
|---|---|---|---|---|---|
| Regression | 7 | 0.0152 | 0.0022 | | |
| Residual | 356 | 0.0011 | 3.00E-06 | 722.04311 | 4.8561E-206 |
| Total | 363 | 0.0162 | | | |

**t – test**

| | **Coefficients** | **Standard Error** | **t Stat** | **P-value** | **Lower 95%** | **Upper 95%** |
|---|---|---|---|---|---|---|
| Intercept | 0.1262 | 0.0005 | 253.9767 | 0 | 0.1252 | 0.1272 |
| $Y_f$ | 0.0924 | 0.0032 | 29.0674 | 5.0905E-96 | 0.0861 | 0.0986 |
| $\delta$ | 0.0018 | 0.0011 | 1.5813 | 0.114688939 | -0.0004 | 0.0041 |
| WS | -0.0007 | 0.0006 | -1.1848 | 0.236872952 | -0.0018 | 0.0004 |
| $T_a$ | -0.0042 | 0.0010 | -4.1846 | 3.60237E-05 | -0.0062 | -0.0022 |
| $T_c$ | -0.0027 | 0.0013 | -2.0496 | 0.041135346 | -0.0053 | -0.0001 |
| $G_h$ | 0.0130 | 0.0035 | 3.6853 | 0.000264018 | 0.0061 | 0.0199 |
| $G_i$ | -0.1172 | 0.0044 | -26.7923 | 2.31477E-87 | -0.1258 | -0.1086 |

*4.4 Predicting Energy Generation Using ANN and MLR Models*

The energy generation from the solar PV system per day is estimated using ANN and MLR models. Figure 6 shows the actual energy production per day. Energy production varies from 12.41 to 95 MWh. Apparently, the produced energy follows a seasonal pattern, with higher values in the summer months and lower values in the winter months. Also, the figure illustrates the predicted energy production using ANN and MLR predictions. The predictions are close to the actual values (E_a). Interestingly, both models succeeded in predicting the magnitude of energy produced. They managed to cope with the outliers' dynamics of the system with reasonable accuracy. The success of these models in approximating the magnitude of energy produced is particularly noteworthy. Even in fluctuating conditions and unpredictable variations, the ANN and MLR models accurately predict daily energy production from the solar PV system. This emphasizes the reliability of these models and their practical utility in real-world scenarios where energy generation patterns are inherently complex and variable.

The correlation between the actual and predicted monthly output energy (Equation 2) during the investigated period is shown in Table 5. The generated energy from the project varied seasonally. It increased in the summer months and decreased in the winter months. The maximum output energy of 2658 MWh was recorded in June, while the minimum was 1124 MWh in December. Both models were able to predict the monthly energy production with a relative absolute percentage error (ep) less than 1%. The mean ep for predictions from ANN and MLR models are 0.58 and 0.48%, respectively. The annual generated energy reached about 24.4 GWh with model predictions deviations less than 0.6%. The efficiency of the solar PV system over the year is shown in Figure 7. Generally, the mean efficiency throughout the year was about 11.7%, ranging between 9 and 14%. The PV system typically performs better at lower temperatures with greater efficiencies due to the temperature coefficient of PV cells. The two models perform well in predicting PV system efficiency throughout the year with reasonable accuracy.

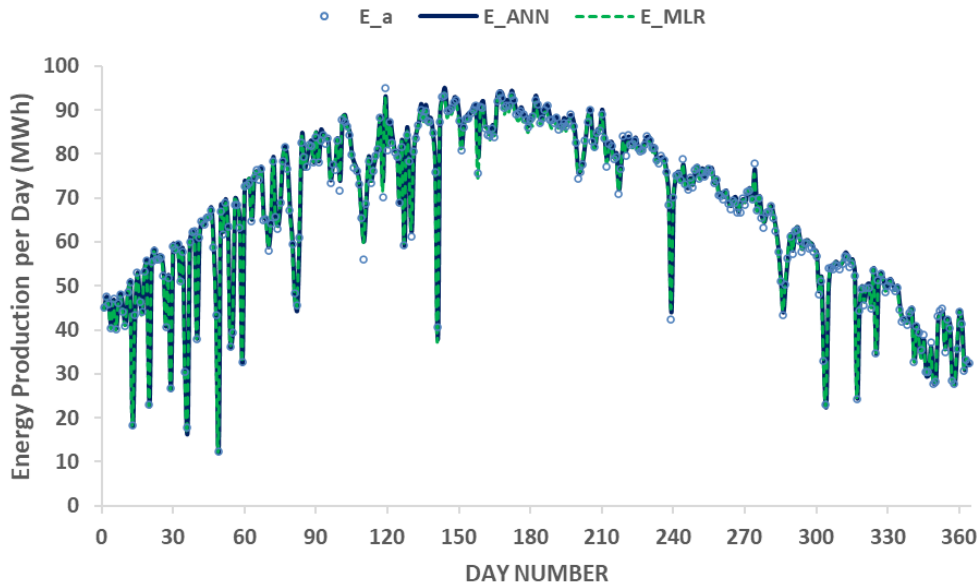


**Fig 6**. The energy production for actual (E_a) and prediction (E_ANN & E_MLR) data per day for one year.

**Table 5**
The correlation between the actual and predicted monthly output energy in 2021.

| Month | Monthly output energy (MWh) | | | Relative absolute percentage error (%) | |
|---|---|---|---|---|---|
| | Actual | ANN prediction | MLR prediction | ANN | MLR |
| January | 1453 | 1457 | 1445 | 0.31 | 0.55 |
| February | 1512 | 1518 | 1514 | 0.44 | 0.15 |
| March | 2197 | 2208 | 2207 | 0.52 | 0.44 |
| April | 2359 | 2382 | 2379 | 0.97 | 0.82 |
| May | 2555 | 2572 | 2554 | 0.70 | 0.03 |
| June | 2658 | 2678 | 2638 | 0.74 | 0.73 |
| July | 2654 | 2670 | 2643 | 0.61 | 0.41 |
| August | 2417 | 2428 | 2411 | 0.45 | 0.25 |
| September | 2167 | 2176 | 2175 | 0.41 | 0.39 |
| October | 1801 | 1813 | 1814 | 0.68 | 0.73 |
| November | 1504 | 1516 | 1509 | 0.76 | 0.33 |
| December | 1124 | 1120 | 1113 | 0.36 | 0.96 |
| Annual production (MWh) | 24399 | 24538 | 24401 | 0.57 | 0.01 |

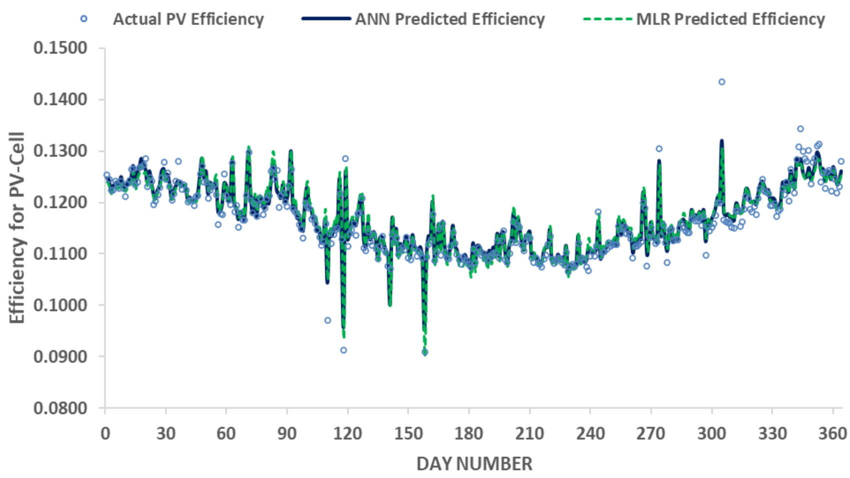


**Fig 7.** Assessment of Solar PV system efficiency throughout the year.

Estimating daily energy generation from the solar PV system using both ANN and MLR models offers several advantages. The models demonstrate a high level of accuracy, as evidenced by their ability to predict the magnitude of energy produced in proximity to actual values. This accuracy is crucial for decision-makers in the solar energy sector, providing reliable insights into daily energy generation patterns. The capability of both ANN and MLR models to capture and adapt to the seasonal pattern in energy production, with higher values in summer and lower values in winter, is advantageous. This ability allows a better understanding of the system's behaviour under varying environmental conditions, enabling better resource planning and operational adjustments. Besides, the models' success in coping with the dynamics of outliers within the system enhances their robustness. This adaptability is particularly valuable in scenarios where unexpected variations can occur, ensuring that the predictions remain reliable even in fluctuating conditions. The advantages of employing ANN and MLR models for estimating daily energy generation include high accuracy in predicting energy magnitude, the ability to capture seasonal patterns, robustness in handling outliers, collectively contributing to informed decision-making, and enhanced operational efficiency in solar PV systems.

## 5. Conclusion

This study contributed to successfully predicting the performance of the solar PV system in arid and semi-arid regions using ANN and MLR models. The relatively acceptable values of coefficients of determinations for ANN (in both training and validation) and MLR indicate their suitability in predicting the energy produced by the system. The mean absolute percentage error was adequate for both models. These are reflected in the close comparison between the actual energy produced and predicted by the two models. Applying a 1:1 sample ratio with the 7-6-1 ANN architecture made impressive $R^2$ values of 0.96 and 0.94 for testing and validation, accompanied by low MAPE. The MLR model demonstrated statistical significance at a 5% level with an $R^2$ of 0.9342. These results helped highlight the importance of considering environmental factors such as temperature and radiation to predict the efficiency of solar cells. Key energy quantities such as energy production and system efficiency are predicted with high precision. Therefore, both models adequately predict average daily, monthly, and yearly energy production from PV systems in arid and semiarid climates with minimum relative absolute percentage errors of 0.57% (ANN) and 0.01% (MLR). The outcomes affirm the adequacy and reliability of the ANN and MLR models, providing valuable insights and understanding for optimizing solar PV systems in challenging environmental conditions, particularly in desert climates. This contributes to pursuing affordable and clean energy, aligning with Sustainable Development Goal 7 (SDG 7). Future work should aim to explore real-time applications, optimize model architectures, and broaden the geographical scope to contribute to the continual improvement of solar PV system performance predictions using the ANN and MLR models.

## References

Abdallah, S., Nijmeh, S. (2004). Two axes sun tracking system with PLC control. *Energy Conversion and Management*, 45 (11-12), 1931-1939. https://doi.org/10.1016/j.enconman.2003.10.007

Adaramola, M. S., Paul, S. S. & Oyewola, O. M. (2014). Assessment of decentralized hybrid PV solar-diesel power system for applications in Northern part of Nigeria. *Energy for Sustainable Development,* 19, 72-82; https://doi.org/10.1016/j.esd.2013.12.007

Ahmad, E. Z., Jarimi, H., & Razak, T. R. (2022). Artificial Neural Network Prediction Model of Dust Effect on Photovoltaic Performance for Residential applications: Malaysia Case Study. *International Journal of Renewable Energy Development*, 11(2), 365-373. https://doi.org/10.14710/ijred.2022.42195

Ahmed, R., Sreeram, V., Mishra, Y., Arif, M. J. R. & Reviews, S. E. (2020). A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. *Renewable and Sustainable Energy Reviews,* 124, 109792. https://doi.org/10.1016/j.rser.2020.109792

Al Naimat, A. & Liang, D. (2023). Substantial gains of renewable energy adoption and implementation in Maan, Jordan: A critical review. *Results in Engineering* 101367.https://doi.org/10.1016/j.rineng.2023.101367

Albaali, A. G., Shahateet, M. (2022). Energy Applications in Green Building to Fulfil the Goals of Sustainable Development: The Case of Jordan. *International Journal of Energy Economics and Policy*, 12 (6), 188-193. https://doi:10.32479/ijeep.13547

Ali, T., Ma, H. & Nahian, A. J. (2019). An analysis of the renewable energy technology selection in the southern region of Bangladesh using a hybrid multi-criteria decision making (MCDM) method. *International Journal of Renewable Energy Research-IJRER*, 9, 1838-1848.

Alshafeey, M. & Csaba, C. (2019, August). A case study of grid-connected solar farm control using artificial intelligence genetic

algorithm to accommodate peak demand In *Journal of Physics: Conference Series,* 1304 (1), p. 012017). https://doi.org/10.1088/1742-6596/1304/1/012017
Altan, A. D., Diken, B. & Kayişoğlu, B. (2021). Prediction of Photovoltaic Panel Power Outputs Using Time Series and Artificial Neural Network Methods. *Journal of Tekirdag Agricultural Faculty* 18 (3), 457-469. https://doi.org/10.33462/jotaf.837446
Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martinez-De-Pison, F. J. & Antonanzas-Torres, F. (2016). Review of photovoltaic power forecasting. *Solar Energy,* 136, 78-111. http://dx.doi.org/10.1016/j.solener.2016.06.069
Anwaralardh, (2023), Anwaralardh company website, https://anwaralardh.com/ [Online]. [Accessed on 2-12-2023].
Baccelli, G., Stathis, D., Hemani, A. & Martina, M. (2020). a non-linear arithmetic unit for neural networks. ACM/IEEE Design Automation Conference (DAC), 2020. IEEE, 1-6.
Bre, F., Gimenez, J. M., Fachinotti, V. D (2018). Prediction of wind pressure coefficients on building surfaces using artificial neural networks. *Energy and Buildings*, 158, 1429-1441. https://doi.org/10.1016/j.enbuild.2017.11.045
Cavalcante, R. L., Costa, T. O., Almeida, M. P., Williamson, S., Galhardo, M. A. B., Macêdo, W. N. (2021). Photovoltaic penetration in isolated thermoelectric power plants in Brazil: Power regulation and specific consumption aspects. *International Journal of Electrical Power & Energy Systems,* 129, 106648. https://doi.org/10.1016/j.ijepes.2020.106648
Choden, Y., Chokden, S., Rabten, T., Chhetri, N., Aryan, K., Al Abdouli, K., (2022). Performance assessment of data driven water models using water quality parameters of Wangchu river, Bhutan. *SN Appl. Sci.* 4, 290. https://doi.org/10.1007/s42452-022-05181-y.
Devaraj, J., Madurai Elavarasan, R., Shafiullah, G., Jamal, T. & Khan, I. (2021). A holistic review on energy forecasting using big data and deep learning models. *International Journal of Energy Research*, 45 (9), 13489-13530. https://doi.org/10.1002/er.6679
Duffie, J. A., Beckman, W. A. & Blair, N. (2020). Solar engineering of thermal processes, photovoltaics and wind, John Wiley & Sons.
El-Aal, S. A., Alqabli, M. A., Naim, A. A. (2023). Forecasting solar photovoltaic energy production using linear regression-based techniques. *Journal of Theoretical and Applied Information Technology*, 101(9), http://www.jatit.org/volumes/Vol101No9/9Vol101No9.pdf
Elsheikh, A. H., Sharshir, S. W., Abd Elaziz, M., Kabeel, A. E., Guilan, W. & Haiou, Z. (2019). Modeling of solar energy systems using artificial neural network: A comprehensive review. *Solar Energy*, 180, 622-639. https://doi.org/10.1016/j.solener.2019.01.037
Erten, M. Y., Aydilek, H. J. (2022). Solar Power Prediction using Regression Models. *International Journal of Engineering Research and Development*, 14 (3), 333-342. https://doi.org/10.29137/umagd.1100957
Freedman, D. A. (2009). Statistical models: theory and practice, cambridge university press.
Fu, X., Wang, X., Gong, Y., Wang, Y. & Zhang, Y. (2023). Impact of Snow Weather on PV Power Generation and Improvement of Power Forecasting. 2023 *International Conference on Power Energy Systems and Applications (ICoPESA),* 2023. IEEE, 448-453. https://doi: 10.1109/ICoPESA56898.2023.10140199
Kalogirou, S. (2007). Artificial intelligence in energy and renewable energy systems, Nova Publishers.
Kartikasari, F. D., Tarigan, E., Irawati, F., Louk, M. H. L., Limanto, S., Asmawati, E. (2023). Optimal solar panel tilt angle calculation and simulation in Indonesia: A Liu and Jordan sky isotropic model-based approach. *International Journal of Science and Research Archive,* 9 (2), 116-121. https://doi.org/10.30574/ijsra.2023.9.2.0517
Keddouda, A., Ihaddadene, R., Boukhari, A., Atia, A., Arıcı, M., Lebbihiat, N., & Ihaddadene, N. (2023). Solar photovoltaic power prediction using artificial neural network and multiple regression considering ambient and operating conditions. *Energy Conversion and Management, 288,* 117186. https://doi.org/10.1016/j.enconman.2023.117186
Koeppe, A., Bamer, F., Selzer, M., Nestler, B. & Markert, B. (2021). Explainable artificial intelligence for mechanics: physics-informing neural networks for constitutive models. *Front. Mater.* 8:824958. https://doi.org/10.48550/arXiv.2104.10683
Kumar, K. R. & Kalavathi, M. S. (2018). Artificial intelligence based forecast models for predicting solar power generation. *Materials Today: Proceedings*. 5 (1), 796-802. https://doi.org/10.1016/j.matpr.2017.11.149
Kumar, N., Sharma, S. P., Sinha, U. K. & Nayak, Y. K. J. I. (2016). Prediction of solar energy based on intelligent ANN modeling. *International Journal of Renewable Energy Research*, 6 (1), 183-188. https://doi.org/10.20508/ijrer.v6i1.3307.g6773
Leva, S., Dolara, A., Grimaccia, F., Mussetta, M., Ogliari, E. J. M. & Simulation, C. (2017). Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power. *Mathematics and Computers in Simulation*, 131, 88-100. http://dx.doi.org/10.1016/j.matcom.2015.05.010
Li, G., Xie, S., Wang, B., Xin, J., Li, Y. & Du, S. (2020). Photovoltaic power forecasting with a hybrid deep learning approach. *IEEE access*, 8, 175871-175880. https://doi: 10.1109/ACCESS.2020.3025860
Limouni, T., Yaagoubi, R., Bouziane, K., Guissi, K., & Baali, E. H. (2022). Univariate and Multivariate LSTM Models for One Step and Multistep PV Power Forecasting. *International Journal of Renewable Energy Development,* 11(3), 815-828. https://doi.org/10.14710/ijred.2022.43953
Loutfi, H., Bernatchou, A. & Tadili, R. (2017). Generation of horizontal hourly global solar radiation from exogenous variables using an artificial neural network in Fes (Morocco). *International Journal of Renewable Energy Research,* 7 (3), 11. https://doi.org/10.20508/ijrer.v7i3.5852.g7140
Mellit, A. & Pavan, A. M. J. S. E. (2010). A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar Energy,* 84 (5), 807-821. https://dx.doi.org/10.1016/j.solener.2010.02.006
Montgomery, D. C., Peck, E. A. & Vining, G. (2021). Introduction to linear regression analysis, John Wiley & Sons.
Natsheh, E. & Samara, S. (2019). Toward Better PV Panel's Output Power Prediction; a Module Based on Nonlinear Autoregressive Neural Network with Exogenous Inputs. *Applied Sciences*, 9 (18), 3670.https://doi.org/10.3390/app9183670
Neville, R. C. (1978). Solar energy collector orientation and tracking mode. *Solar Energy*, 20 (1), 7-11. https://doi.org/10.1016/0038-092X(78)90134-2
Nitisanon, S., & Hoonchareon, N. (2017, July). Solar power forecast with weather classification using self-organized map. In *2017 IEEE power & energy society general meeting* (pp. 1-5). IEEE. https://doi.org/10.1109/PESGM.2017.8274548
Tsai, W.-C., *Et Al.,* A Review Of State-Ofnespoli, A., Ogliari, E., Dolara, A., Grimaccia, F., Leva, S., & Mussetta, M. (2018, July). Validation of ANN training approaches for day-ahead photovoltaic forecasts. *In 2018 International Joint Conference on Neural Networks (IJCNN)* (pp.1-6). IEEE. https://doi.org/10.1109/IJCNN.2018.8489451
Poon, K. H. & Kaempf, J. (2019, November). A morphological based PV generation and energy consumption predictive model for Singapore neighbourhood. *Journal of Physics: Conference Series*, 2019. IOP Publishing, 012033. https://doi.org/10.1088/1742-6596/1343/1/012033
Rajput, G., Raut, G., Chandra, M., Vishvakarma, S. (2021). VLSI implementation of transcendental function hyperbolic tangent for deep neural network accelerators. *Microprocessors and Microsystems,* 84, 104270. https://doi.org/10.1016/j.micpro.2021.104270.
Rao, D. V. S. K. K., Prusty, B. R., & Myneni, H. (2022, July). Bright Sunshine Duration Index-Based Prediction of Solar PV Power Using ANN Approach. In *2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP)* (pp. 1-5). IEEE. https://doi.org/10.1109/ICICCSP53532.2022.9862452
Rasamoelina, A. D. , Adjailia, F. and Sinčák,P. (2020) A Review of Activation Function for Artificial Neural Network, *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI),* Herlany, Slovakia, 2020, pp. 281-286, doi: https://doi.org/10.1109/SAMI48414.2020.9108717 .
Reyes-Téllez, E., Parrales, A., Ramírez-Ramos, G., Hernández, J., Urquiza, G., Heredia, M. & Sierra, F. (2020). Analysis of transfer functions and normalizations in an ANN model that predicts the transport of energy in a parabolic trough solar collector. *Desalination and Water Treatment,* 200, 23-41. https://doi.org/10.5004/dwt.2020.26063.

Rumbayan, M., Abudureyimu, A., & Nagasaka, K. (2012). Mapping of solar energy potential in Indonesia using artificial neural network and geographical information system. *Renewable and Sustainable Energy Reviews*, *16*(3), 1437-1449. https://doi.org/10.1016/j.rser.2011.11.024

Shafiee, S. & Topal, E., (2009). When will fossil fuel reserves be diminished? *Energy Policy,* 37 (1), 181-189; https://doi.org/10.1016/j.enpol.2008.08.016

Shehab, M., Abualigah, L., Omari, M., Shambour, M. K. Y., Alshinwan, M., Abuaddous, H. Y., & Khasawneh, A. M. (2022). Artificial neural networks for engineering applications: A review. *Artificial Neural Networks for Renewable Energy Systems and Real-World Applications*, 189-206. https://doi.org/10.1016/B978-0-12-820793-2.00003-3

Thombare, S., Pande, V., Kulkarni, R. & Kakade, S. (2022). Prediction of Solar Power Using Linear Regression. Sustainable Energy and Technological Advancements: *Proceedings of ISSETA 2021*. Springer. https://doi.org/10.1007/978-981-16-9033-4_53

Tsai, W.-C., Tu, C.-S., Hong, C.-M. & Lin, W.-M. (2023). A Review of State-of-the-art and Short-Term Forecasting Models for Solar PV Power Generation. *Energies,* 16(14), 5436; https://doi.org/10.3390/en16145436

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology* , 49 (11), 1225-1231. https://doi.org/10.1016/S0895-4356(96)00002-9

Yadav, O., Kannan, R., Meraj, S. T. & Masaoud, A. (2022). Machine Learning Based Prediction of Output PV Power in India and Malaysia with the Use of Statistical Regression. *Mathematical Problems in Engineering,* 2022, Article ID 5680635. https://doi.org/10.1155/2022/5680635