



Predicting Azimuth and Tilt Angles of Photovoltaic Panels with Random Forest and Support Vector Machine: A Study in the Shore Area of South Papua, Indonesia

Yohanes Letsoin^{1,2,*}, David Guth³, Mohammad Khairudin², Rustam Asnawi²

¹ Department of Electrical, Faculty of Engineering, Universitas Musamus Merauke, Kamizaun 99611, Papua Selatan, Indonesia

² Department of Electrical, Faculty of Engineering, Universitas Negeri Yogyakarta, Yogyakarta 55281, Indonesia

³ Department of Mechanical Engineering, Faculty of Engineering, Czech University of Life Sciences Prague, Kamycka 129, 16500 Praha Suchdol, Czech Republic

ARTICLE INFO

Article history:

Received 14 February 2025

Received in revised form 17 March 2025

Accepted 30 June 2025

Available online 20 July 2025

ABSTRACT

Indonesia has various potential renewable energy sources such as solar energy, which can generate more than 200 GW. However, this promising energy source tends to be utilized in small numbers, i.e., less than 100 MW, compared to the potential energy. The solar photovoltaic (PV) systems significantly depend on the location, such as the optimal configuration of azimuth and tilt angles, which determine the amount of solar irradiance captured by the panels. Even though numerous studies and practical applications of PV systems have been investigated globally, unfortunately, there is a scarcity of research focusing on PV technology, specifically in the context of South Papua, where its usage remains unexplored, for example, in the optimal configuration of azimuth and tilt angles. Therefore, this research focuses on optimizing the tilt angle and azimuth of PV panels located in coastal areas of South Papua by utilizing PVSyst software and annual radiation estimation. Moreover, the elevated solar radiation levels in the area establish it as a key site for installing PV systems; on the other hand, the distinctive geographical and weather conditions require accurate positioning of panels to achieve the best possible efficiency of PV systems. To address this, we employed Random Forest (RF) and Support Vector Machine (SVM) models in a Python environment to analyse the impact of azimuth and tilt angles on the global collection plane of solar panels. The dataset included azimuth angles ranging from 0° to 60° and tilt angles from 8° to 45° representing a comprehensive range of possible configurations. Our findings indicate that an azimuth angle of 60° and a tilt angle of 8° yield the highest global collection plane values, making it the most effective configuration for maximizing solar energy capture in this region during the dry season. Conversely, in the wet season, the optimal tilt angle shifts to 30° with an azimuth angle of 0°. Furthermore, the Random Forest (RF) model outperformed Support Vector Machine (SVM) in predictive accuracy, as evidenced by a lower mean squared error. This study not only provides valuable insight into the optimization of solar field orientation but also highlights machine learning techniques in improving PV system performance. By optimizing solar energy capture, this research supports Indonesia's broader goals of transitioning to renewable energy sources, contributing to sustainable development and energy security in South Papua, Indonesia.

Keywords:

Machine learning; PV; predict; SVM; RF

* Corresponding author

E-mail address: yohanes0069ft.2023@student.uny.ac.id

<https://doi.org/10.37934/ard.137.1.167185>

1. Introduction

Indonesia encompasses several potential renewable energy resources, namely, hydropower (94.3 GW), geothermal (28.5 GW), solar (207.8 GW), wind (60.6 GW) and ocean energy (17.9 GW). Although solar energy has the highest potential, its use of energy in producing electricity is still less than 100 MW in comparison to more than 200 GW with the efficiency of available Photovoltaic (PV) technology [1]. In addition to a planned installed capacity of 3 MW (2019–2028), Indonesia's most significant solar potential is located primarily in West and East Kalimantan and South Sumatera with an estimated total of 17–20 GW. In other provinces, such as Papua Province, the achievable capacity is roughly 2 GW. Nevertheless, the Provinces' projected 18% annual population growth rate will lead to a rise in energy consumption [2]. The solar radiation in the installation area, the tilt angle and the orientation of the solar panels affect how much solar energy is absorbed by photovoltaic (PV) modules. Generally, the performance of a PV is directly proportional to its solar radiation level; as the solar radiation level rises, it correspondingly increases the specific solar energy produced by the PV [3,4]. The efficiency of these systems is influenced predominantly by the optimal configuration of solar panels, particularly the azimuth and tilt angles, which directly impact the amount of solar irradiance captured. Adjusting these angles to optimize solar energy capture can substantially enhance the performance of PV systems, as investigated in Spain. A solar PV system positioned in the south is as follows: orientation ($\gamma = 0$) and tilt angle (β) changes vary from 21° – 23° , 37° – 40° and 37° – 40° and as a result, the percentages of energy losses are 5%, 10%, 15% and 20%, respectively [5]. However, in West and Central Africa, a slight movement (up to 20°) from the optimal position and tilt is unlikely to provide a substantial impact on the incident solar radiation (irradiation decrease of less than 5%) [6]. Other prior studies, for instance [7], assessed the efficiency of a 10 kWp grid-connected copper indium selenium (CIS) PV system by simulating meteorological variables during spring, dry season and winter at a specific latitude, without consideration for the site's location being coastal or lowland. Another investigation conducted a comparative analysis of various PV technologies in diverse global regions, particularly in Asian countries like Singapore, Korea, China and Iran, while omitting Indonesia from the examination [8]. The impact of dust and shading resulting from solar PV conditions and climate was examined by Al-Ibrahim [9], yet the effects of inclination and orientation on shading patterns cast on PV panels were not investigated.

The azimuth orientation angle of a collector (α), denoted as azimuth orientation angle or, essentially, as orientation, indicates the angle measured clockwise in the horizontal plane between the true south direction (or sometimes the true north, depending on the reference) and the projection of the collector's surface onto the horizontal plane as shown in Figure 1. This angle signifies the deviation from the true south, with due south (towards the equator) regarded as an orientation of 0° as per convention. The angle of inclination (β), commonly known as the tilt angle or slope, measures the angle at which the collector plane intersects with the horizontal plane. In a scenario where the collector is in the horizontal plane facing upwards toward the sky, the tilt angle is defined as 0° .

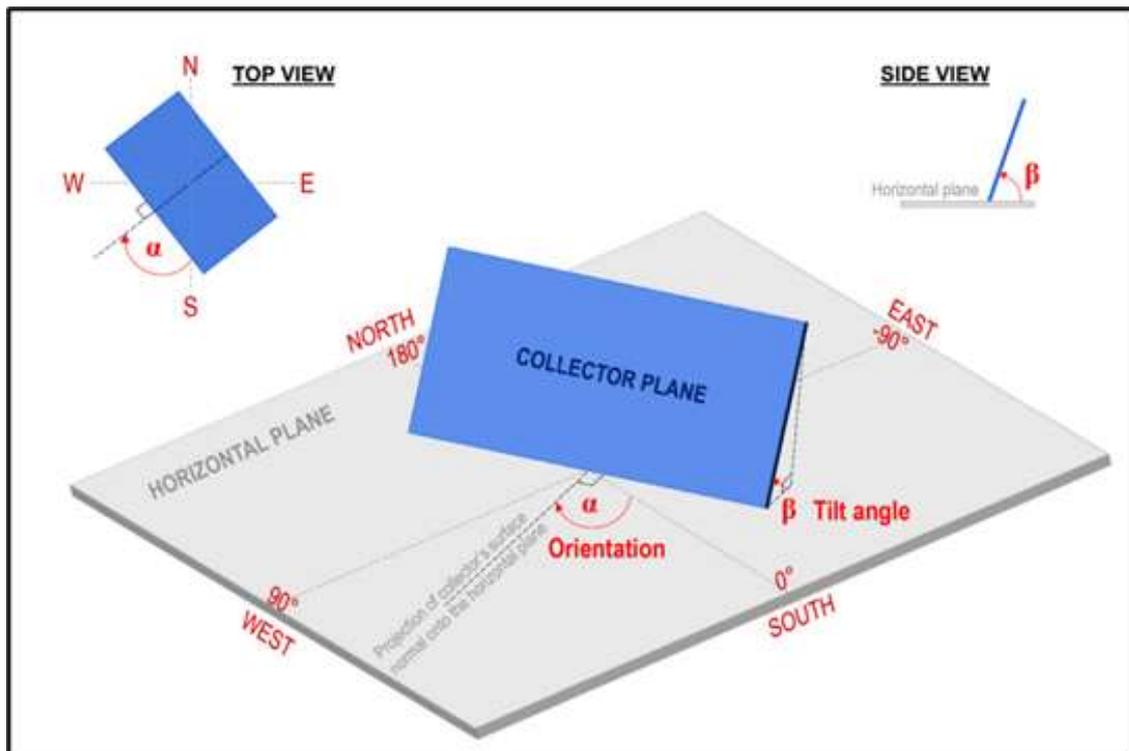


Fig. 1. Tilt and azimuth angles of the PV panel [10]

In Indonesia, a case in point involves the calculation of the specific energy yield from a rooftop PV system situated in Surabaya, Indonesia, with an assumed module tilt of approximately 30° in alignment with the typical roof incline of local residences [11]. Findings reveal that the highest specific energy output in Surabaya is attained when solar modules are oriented towards the north (0° azimuth angle), yielding approximately 1462 kWh/kWp. Conversely, solar panels facing south have the lowest specific energy yield, estimated at around 1233 kWh/kWp per month.

Calculation of tilt angles is typically divided into two distinct groups: calculations based on the latitude angle [12] and calculations that are centered on the global irradiation received by the surface as implemented in different previous studies [13,14]. The estimation of the total radiation incident on an inclined plane can be determined utilizing the Hay, Davies, Klucher, Reindl model (HDKR model), which comprises the sum of direct, diffuse and reflected radiation or albedo [15,16]. Previous research has demonstrated the application of the Perez Sky model for solar radiation estimation by assessing direct normal irradiance (DNI), diffuse horizontal irradiance (DHI) values and global horizontal irradiance (GHI). Utilization of this model requires supplementary information regarding the solar zenith angle and functions based on the premise of a cloudless sky [17,18]. The Liu and Jordan model, as found in previous studies [19,20], serves as a method used to anticipate the mean solar irradiance for every month on surfaces positioned directly towards the equator, establishing the optimal angle for sloped surfaces by recognizing the point at which total daily solar radiation is optimized [21]. However, different system topologies and geographical conditions may result in significant optimization of PV performance [22].

Therefore, this study concerns how to achieve the optimal tilt and azimuth according to the PV technology installed in the shore area, while determining the optimal azimuth and tilt angles for solar panels in South Papua by employing machine learning techniques. Several earlier studies have utilized machine learning for various purposes. For instance, Alghamdi *et al.*, [23] investigated the performance prediction of a hybrid solar photovoltaic-thermoelectric (PV-TE) system utilizing various crystalline silicon solar cell types by employing deep neural networks (DNN) as an alternative to

conventional numerical methods. The research highlights the limitations of traditional numerical solvers, which are computationally expensive and time-consuming and proposes DNN as a more efficient solution for performance forecasting. Ahad *et al.*, [24] evaluated various machine learning classifiers for detecting fraudulent credit card transactions, emphasizing the limitations of conventional methods that rely on historical data. The research compared classifiers such as Decision Trees, Random Forest, K-Means, Isolation Forest, Neural Networks (NN) and the proposed Threshold Mahalanobis Distance classifier. The study concludes that TMD is a robust and efficient model for fraud detection, recommending its implementation for enhanced financial security. Also, Aburashed *et al.*, [25] propose a machine learning-based framework for SQL injection detection, leveraging algorithms such as Random Forest (RF), Gradient Boosting, Support Vector Machine (SVM) and Artificial Neural Networks (ANN). The study ranked ANN as the most effective, followed by Gradient Boosting and Random Forest. Nevertheless, the authors emphasize the importance of selecting appropriate machine learning models based on accuracy, computational efficiency and adaptability to evolving attack patterns. The study concludes that machine learning classifiers, particularly ANN, offer a promising approach to improving the detection and mitigation of SQL injection attacks in web applications. In our investigation, we utilize RF and SVM models to evaluate the impact of azimuth and tilt angles on the overall solar panel collection plane. Using these models, our objective is to identify configurations that maximize solar energy generation, thereby offering insight into optimal practices for PV system installation in South Papua. This study contributes to enhancing the efficiency of solar energy systems and aligns with Indonesia's broader initiatives towards transitioning to renewable energy sources.

2. Methodology

2.1 Study Site

The designated location for this research was situated inside the University, close to the shore area in Merauke Regency (137°38'52.9692" E–141°0'13.3233" E and 6°27'50.1456" S–9°10'1.2253" S) as presented in Figure 2. The latitude and longitude of the University are –8.49 (south) and 140.39 (east), respectively. Geographically, the Regency shares its borders directly with Papua New Guinea and Australia, featuring an average temperature of approximately 27 degrees Celsius, complemented by a humidity level of approximately 81% [26].

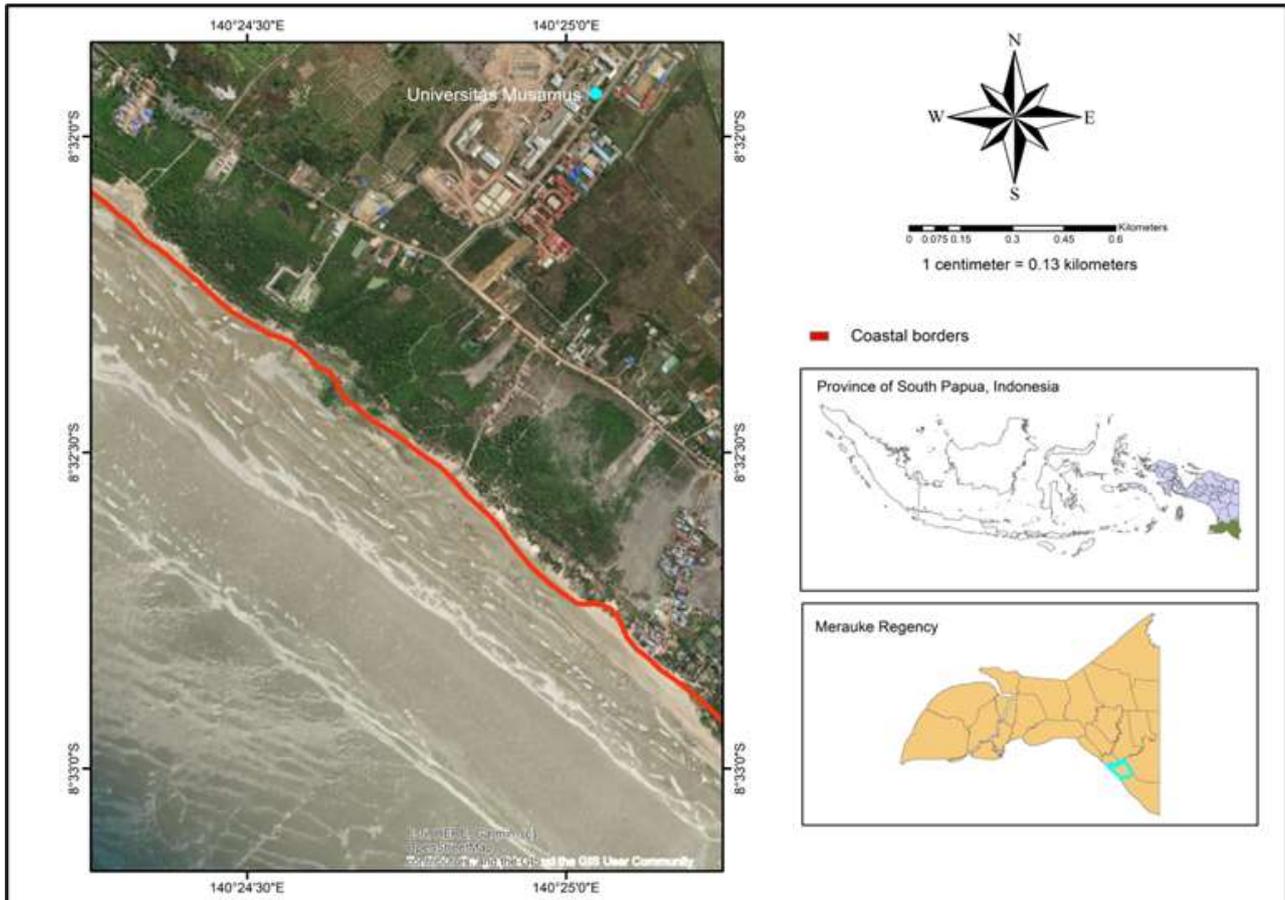


Fig. 2. Study location

2.2 Experimental Design

The Perez Sky model is employed for the estimation of solar radiation distribution on inclined surfaces featuring diverse angles and orientations. To determine the azimuth and tilt angles at a specific time and location, a comprehension of the sun's position is necessary, along with an appreciation of how the Perez Sky model accommodates diffuse radiation. Therefore, the procedures were arranged as follows:

- i. Determine location coordinates, i.e., at Universitas Musamus, positioned as displayed in Figure 2 above.
- ii. Calculate solar declination angle (δ):

$$\delta = (23.45^\circ) \times \sin \frac{360}{365} \times (d - 81) \quad (1)$$

Where, d is the day of the year (1 for January 1) as implemented in the previous study [27].

- iii. Calculate the hour angle (H) that indicates the sun's position relative to the meridian as written by Kaddoura *et al.*, [28]:

$$H = 15^\circ \times (\text{local time} - 12) \quad (2)$$

- iv. Calculate the zenith angle (θ_z), which represents the angle between the vertical and the sun's position, where L is latitude:

$$\theta_z = \sin(L) \times \sin(\delta) + \cos(L) \times \cos(\delta) \times \cos(H) \quad (3)$$

- v. Calculate solar azimuth angle (γ_s); the solar azimuth angle indicates the sun's direction relative to the north:

$$\gamma_s = \frac{-\cos(\delta) \times \sin(H)}{\cos \theta_z} \quad (4)$$

- vi. Calculate the azimuth angle for the PV panel (γ) in the southern hemisphere; PV panels typically face north for optimal exposure. The optimal azimuth angle for a PV panel can be derived as follows:

$$\gamma = 180^\circ - \frac{180^\circ}{\pi} \times \text{atan}^2(F_{11} \times \cos(\theta_z), F_{12} \times \sin(\theta_z)) \quad (5)$$

- vii. Calculate tilt angle (β). Tilt angle represents the panel's inclination from the horizontal plane. Using the zenith angle, the tilt angle is generally calculated as $\beta = 90^\circ - \theta$.

The investigation also deployed Liu Jordan models according to the procedures as follows:

- i. identify the latitude and longitude of the study field.
- ii. estimate solar declination:

$$\delta = 23.45 \times \sin\left(\frac{360}{365} \times (N + 284)\right) \quad (6)$$

Where, N is the day of the year [29].

- iii. calculate the hour angle:

$$H = 15 \times (\text{local solar time} - 12) \quad (7)$$

- iv. calculate solar zenith and azimuth angles:

$$\cos(\theta_z) = \sin(\phi) \times \sin(\delta) + \cos(\phi) \times \cos(\delta) \times \cos(H) \quad (8)$$

- v. The solar azimuth angle γ can be derived as:

$$\sin(\gamma_s) = \frac{-\cos(\delta) \times \sin H}{\sin \theta_z} \quad (9)$$

and

$$\cos(\gamma_s) = \left(\frac{\sin(\delta) - \sin(\theta_z) \times \sin(\phi)}{\cos(\theta_z) \times \cos(\phi)} \right) \quad (10)$$

A consistent intensity referred to as the solar constant, $G_{sc} = 1367 \text{ W/m}^2$, signifies the solar energy received per unit area of a surface perpendicular to the direction of its propagation at this average

distance. Nevertheless, variations in the Sun-Earth distance caused by Earth’s movement lead to changes in the extraterrestrial radiation incident on a plane perpendicular to the radiation, denoted as G_{ext} , which is calculated as:

$$G_{ext} = G_{sc} \times \left(1 + 0.033 \times \cos\left(360 \times \frac{n}{365}\right)\right) \tag{11}$$

Where, G_{sc} is a solar constant [30].

2.3 Data Processing

The simulation datasets were acquired from Meteonorm 8.1 (2016 to 2021), which was gathered based on the geographical location of the study site, as presented in Figure 3. These datasets encompass global and diffuse irradiance measured on the horizontal plane, global irradiance recorded on tilted planes oriented at 8, 20-, 30-, 40- and 45-degrees due south, as well as global irradiance data obtained from west and east-oriented surfaces tilted at 0, 20, 45 and 60 degrees of standalone PV system. Then, the simulation data was recorded and analysed using the PV solar software named PVsyst.

Site: Mopa-Merauke (UniversitasMusamus) (Indonesia)						
Data source: Meteonorm 8.1 (2016-2021), Sat=100%						
	Global horizontal irradiation	Horizontal diffuse irradiation	Temperature	Wind Velocity	Linke turbidity	Relative humidity
	kWh/m ² /mth	kWh/m ² /mth	°C	m/s	[-]	%
January	164.1	78.0	27.8	2.40	3.495	83.3
February	164.9	78.8	27.6	2.40	3.435	84.4
March	169.2	79.5	27.5	2.19	3.311	84.0
April	160.1	66.7	27.3	2.79	3.556	84.5
May	146.1	69.1	27.0	3.29	3.712	81.9
June	129.4	65.3	25.8	3.50	3.668	84.7
July	146.7	61.0	25.5	3.70	3.604	81.9
August	162.8	77.3	25.4	3.70	3.628	80.5
September	160.4	72.7	26.0	3.80	3.774	80.6
October	188.2	83.2	27.2	3.71	3.938	79.1
November	183.6	78.7	27.9	3.20	3.818	79.6
December	172.8	82.8	28.7	2.49	3.547	78.7

Fig. 3. Monthly weather data from the study location

The dataset used in this study, as shown in Figure 3, provides crucial environmental parameters relevant to optimizing PV energy conversion in South Papua. The dataset includes monthly records of GHI, ranging from 146.1 to 188.2 kWh/m², which helps to assess solar energy potential throughout the year. Additionally, HDI values, spanning from 63.0 to 92.8 kWh/m², offer insights into atmospheric scattering effects that influence PV performance. The recorded temperature fluctuates between 25.3 °C and 28.7 °C, impacting panel efficiency due to thermal effects, while wind velocity, ranging from 2.30 to 3.70 m/s, contributes to panel cooling and energy efficiency. The dataset also includes the Linke turbidity factor, indicating atmospheric clarity, with values varying between 3.315

and 5.547 and relative humidity levels from 78.7% to 84.4%, which can affect panel performance through condensation and dirt accumulation. The preprocessing of the dataset involved key steps to ensure data quality and readiness for model training. First, data collection was conducted from the Meteonorm 8.1 database (2016–2021) and then, data cleaning addressed the missing values using interpolation and removed outliers with the Z-score method. Therefore, feature engineering entailed deriving solar energy potential indices and categorizing azimuth and tilt angles. Normalization and scaling were applied using min-max normalization for irradiation and humidity and Z-score standardization for temperature and wind speed. Finally, the dataset was split into 80% training and 20% testing. After preprocessing, the next step in this study involved applying RF and SVM models to predict the optimal azimuth and tilt angles for solar panels. The dataset is split into training (80%) and testing (20%) subsets to effectively train and validate the models. The RF model, with 100 estimators, captures nonlinear relationships in the data. The SVM model with an RBF kernel is applied to handle complex patterns in solar energy prediction. Model evaluation is executed using key performance metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and R^2 to assess the models' accuracy and generalization ability. Additionally, a CV score is performed to validate the model robustness and prevent overfitting. The results are visualized through 3D scatter plots comparing actual vs. predicted values, providing insight into the effectiveness of each model in both wet and dry seasonal conditions. Finally, the best-performing model is selected based on the evaluation results, guiding optimal solar panel positioning for enhanced energy yield.

2.4 Prediction Models

A RF represents a supervised machine-learning model that employs ensemble learning to address regression and classification tasks. Initially introduced by Leo Breiman and Adele Cutler in 2006, it was an extension of the work by Tim Kam Ho in 1995. The RF method functions by generating numerous decision trees during the training phase. It randomly selects samples from the training dataset with replacements and combines these trees using techniques such as bagging and boosting. This approach enhances diversity and reduces correlation among the ensemble of decision trees. The forecast generated by the RF model varies depending on the problem's characteristics; for regression tasks, the forecast is calculated from the average of all decision trees in the ensemble, while for classification tasks, the predicted class is determined by the majority vote of the most frequent categorical class. The primary procedures for RF can be outlined as follows:

- i. Initiate the process by considering a specified set of training data denoted as $X = x_{1...}, x_n$ along with corresponding responses $Y = y_{1...}, y_n$
- ii. A series of random samples X_b, Y_b comprising n training instances is chosen N times while ensuring replacement, where b ranges from 1 to N
- iii. Each random sample X_b, Y_b undergoes the fitting and training of a regression tree denoted as f_b
- iv. Subsequently, the final estimation of regression \hat{f} for unseen sample x' is determined by averaging the predictions from all individual regression trees on x' , using the equation as follows:

$$\hat{f} = \frac{1}{N} \sum_{b=1}^{N_1} f_b(x') \quad (12)$$

SVM effectively utilizes kernel functions to transform data from a lower to a higher dimensional feature space. Consequently, linear solutions within the higher dimensional feature space are

equivalent to nonlinear solutions in the original lower dimensional state. The Kernel function, referred to as $k(x, x')$, is a symmetrical function that effectively illustrates the similarity between observations based on their feature attributes; the variable x denotes the input space, while x' describes the vector of computed features for the training or test samples. The predictive computation of an SVM can be mathematically expressed as:

$$y = \sum_{i=1}^N (c_i k(x, x') + b) \quad (13)$$

3. Results

3.1 Optimum Tilt and Azimuth Angles

PVSyst simulation software was used to perform a PV array of 700 Wp Si-mono twin half-cell bifacial module with a Maximum Power Point Tracking (MPPT) converter. Different tilt angles were positioned from 0° – 60° in the site location, i.e., the Electrical Department of Musamus University within a latitude of -8.5344° S, longitude of 140.4152° E and an altitude of about 32 m. Two seasons were selected: dry or summer (Oct–Mar) and wet (Apr–Sept). Tilt and azimuth angles were configured for fixed tilted planes, as shown in Figure 4.

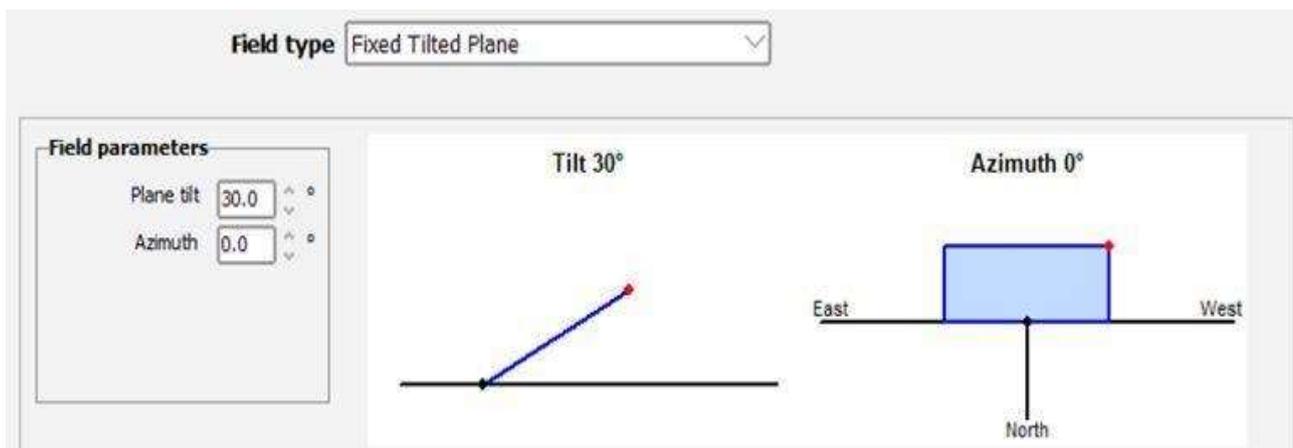


Fig. 4. Tilt and azimuth angles ($^\circ$) settings in winter

The optimal tilt (β) and azimuth angles (γ) were determined by the point at which the highest power output is achieved. Further, the power output was defined through the global collector plane (G_{cp}). This experimental work involved four ranges of tilt and azimuth angles, as demonstrated in Figure 5. In the wet season, the optimal azimuth is 0° ; the tilt angle reaches peak irradiation at 30° and continues to decrease at 45° . The 20° azimuth ranked second in collecting the amount of solar irradiance. Conversely, a low irradiation was captured through the azimuth of 60° .

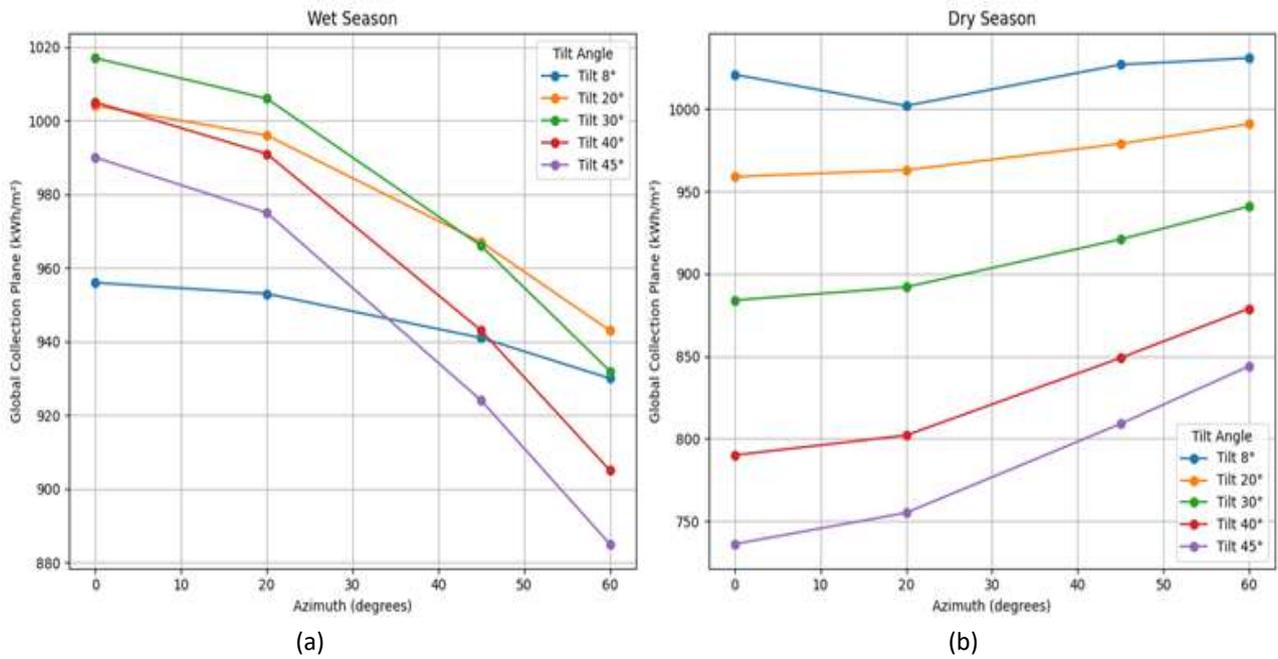


Fig. 5. Tilt and azimuth angles (°) vs max glob coll plane (kWh/m²) in (a) winter or wet season (b) summer or dry season

However, in the dry season, as displayed in Figure 5(b), lower tilts (8°) and azimuths closer to 60° optimize energy output, capturing more direct sunlight due to the sun's higher position in the sky.

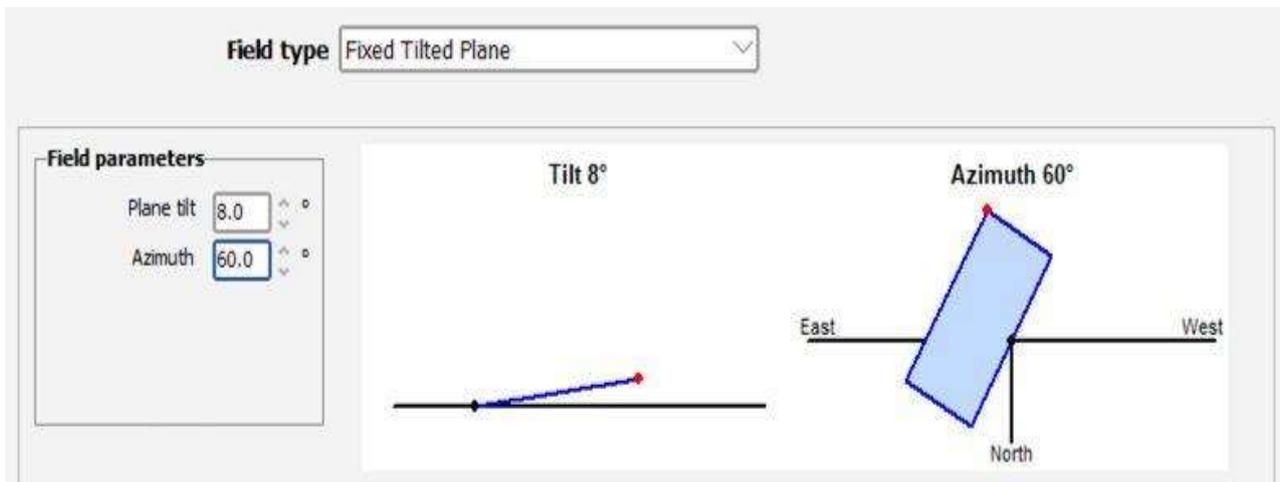


Fig. 6. Tilt and azimuth angles (°) settings in the dry season

The variability in the maximum total solar irradiance captured on an inclined surface is determined by the module's azimuth angle and latitude. This changeability highlights the importance of the orientation of solar PV modules at specific latitudes for optimizing the received solar irradiance [31,32]. This simulation work found that the tilt angle tends to be higher in the winter or wet season, i.e., about 30 degrees (Figure 5(a)), compared to the tilt angle in the summer or dry season, i.e., 8 degrees (Figure 5(b)). This is because the sun is lower in the sky during the wet season. Meanwhile, more tilt is required on the panels to maximize solar radiation. However, since the sun is higher in the sky during the summer or dry season, the panels require less tilting to maximize solar energy capture. These findings are analogous to relevant studies in China [33,34].

A previous study stated that the optimum fixed azimuth angle is determined by subtracting the absolute latitude angle from the longitude angle and the ideal fixed tilt angle is roughly equal to the location's latitude [35]. Nevertheless, this study did not investigate the circumstances that explained the limitation of the experimental work. The annual optimal tilt angle of solar panels may differ based on geographical location and the computational system utilized, for instance, optimum tilt angles vary from 0 to 59 degrees in Libya [36].

Previous studies in Türkiye indicated that the tilt angles of panels could range from -9.9 degrees to 59.2 degrees with an average annual angle of 32 degrees [37]. However, this study revealed a tilt angle of 8 degrees during the dry season and 30 degrees in the wet season, which also differs from the global tilt angle observed by numerous studies [38,39].

3.2 Performance Ratio

Figure 7 presents performance ratios (PR) at a tilt angle of 8 degrees and an azimuth of 60 degrees (Figure 7(a)) and a tilt angle of 30 degrees and an azimuth of 0 degrees (Figure 7(b)). As can be seen, PV performance with a tilt of 30 degrees and an azimuth of 0 degrees is slightly higher than a tilt angle of 8 degrees and an azimuth of 60 degrees. Nevertheless, this experimental study highlights some considerations in PR estimation, such as temperature effects. Actual energy output decreases with higher module temperatures, as investigated in several earlier studies [40,41]. Systems losses due to shading or dirt influenced actual energy output calculations [42-44]. Another concern is data quality for both actual energy output and plane of array irradiance. Therefore, some further refinement based on multiple data resources is suggested.

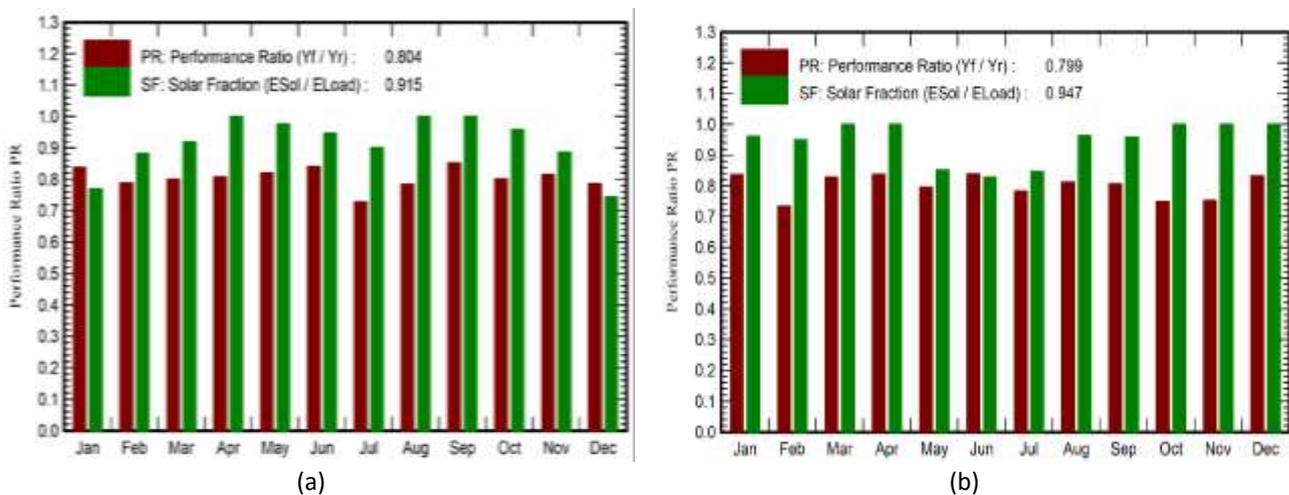


Fig. 7. Performance ratio of PV module

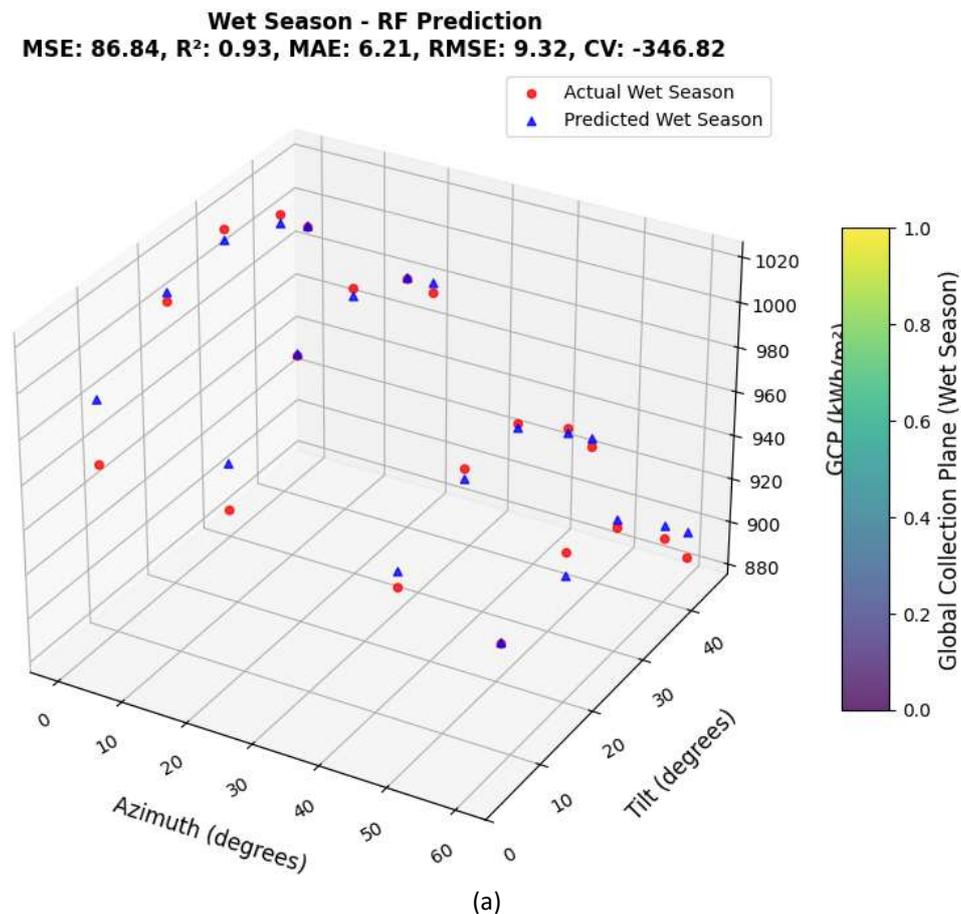
3.3 Model Performance

Implementing SVM and RF models to predict the optimal azimuth and tilt angles for solar panel installations in the shore area of South Papua, Indonesia, has provided significant insight into solar energy optimization. These models were trained using comprehensive datasets that included various combinations of tilt and azimuth angles and meteorological data to maximize the global collection plane as a measure of solar energy output. The dataset used in this study contains solar irradiance. Each performance metric serves a critical role in assessing the effectiveness of the models in the context of PV optimization (Table 1).

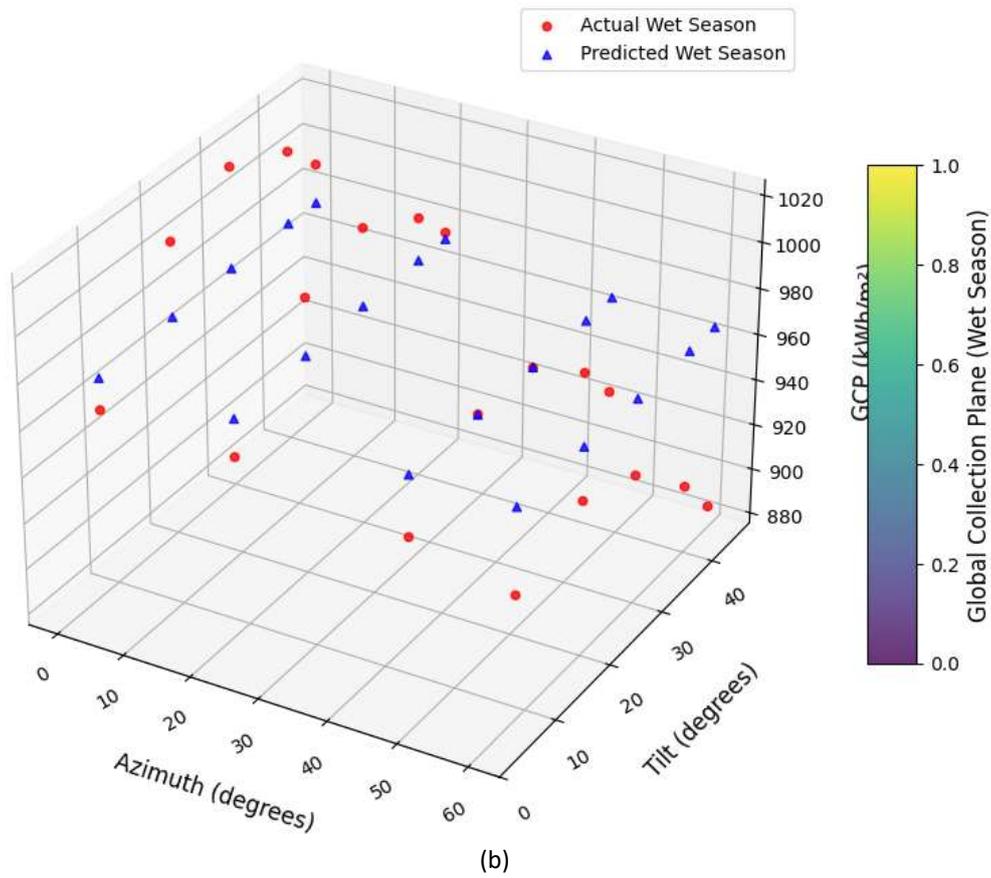
Table 1
 Metrics used in this study

Metrics	Description
Mean Squared Error (MSE)	MSE measures the average squared difference between actual and predicted values. Lower MSE values indicate better predictive accuracy. For PV optimization, minimizing MSE ensures reliable energy yield forecasts.
R ² (Coefficient of Determination)	R ² evaluates how well the model explains the variance in the data. A higher R ² value indicates a strong correlation between predicted and actual solar energy outputs, which is essential for precise system adjustments.
Mean Absolute Error (MAE)	MAE quantifies the average absolute error between predicted and actual values, providing an easy-to-interpret measure of accuracy. Lower MAE values signify more accurate energy predictions.
Root Mean Squared Error (RMSE)	RMSE measures the standard deviation of prediction errors, emphasizing larger errors more than MAE. It provides valuable insight into the reliability of PV system performance estimates.
Cross-Validation Score (CV)	CV helps assess the generalizability of the model by evaluating its performance across different subsets of data. A less negative CV score suggests better adaptability to unseen data.

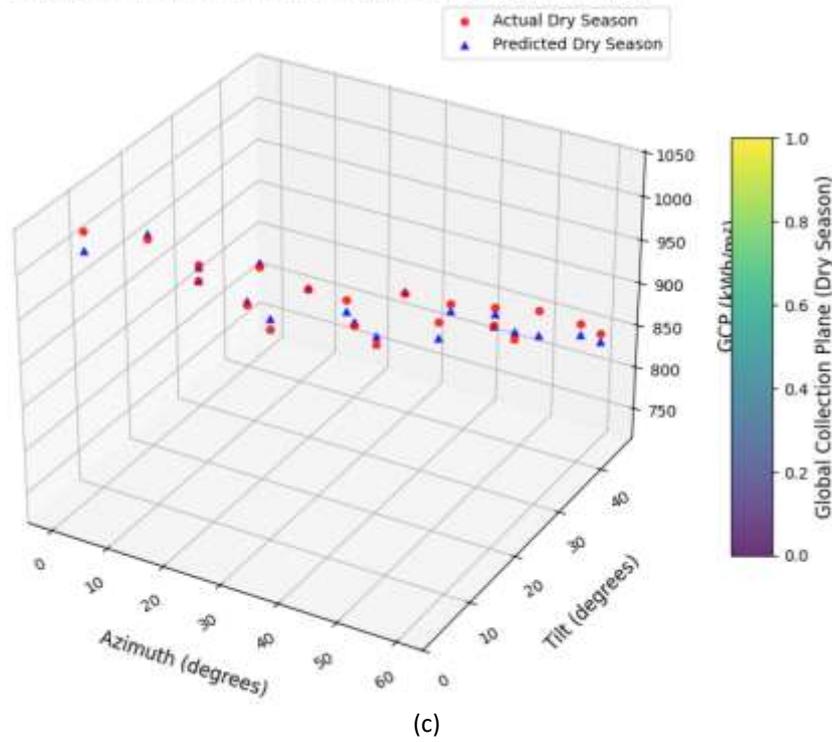
The results are shown in Figure 8. The graphs depict the performance of the RF and SVM in predicting the Global Collection Plane (Gcp) through azimuth and tilt angles during the wet and dry seasons. The horizontal axes represent the azimuth and tilt angles and the vertical axis denotes the Gcp, denoted in a color scale that ranges from dark purple (lower Gcp values) to yellow (higher Gcp values). The graphs show the actual Gcp in each season with the performance identified through mean squared error (MSE), mean absolute error (MAE), root mean squared error (RMSE) and cross-validation score (CV). A red circle explains the actual value and the prediction value is represented by a blue triangle.



Wet Season - SVM Prediction
MSE: 1145.16, R²: 0.10, MAE: 28.06, RMSE: 33.84, CV: -1776.01



Dry Season - RF Prediction
MSE: 128.36, R²: 0.98, MAE: 8.82, RMSE: 11.33, CV: -368.84



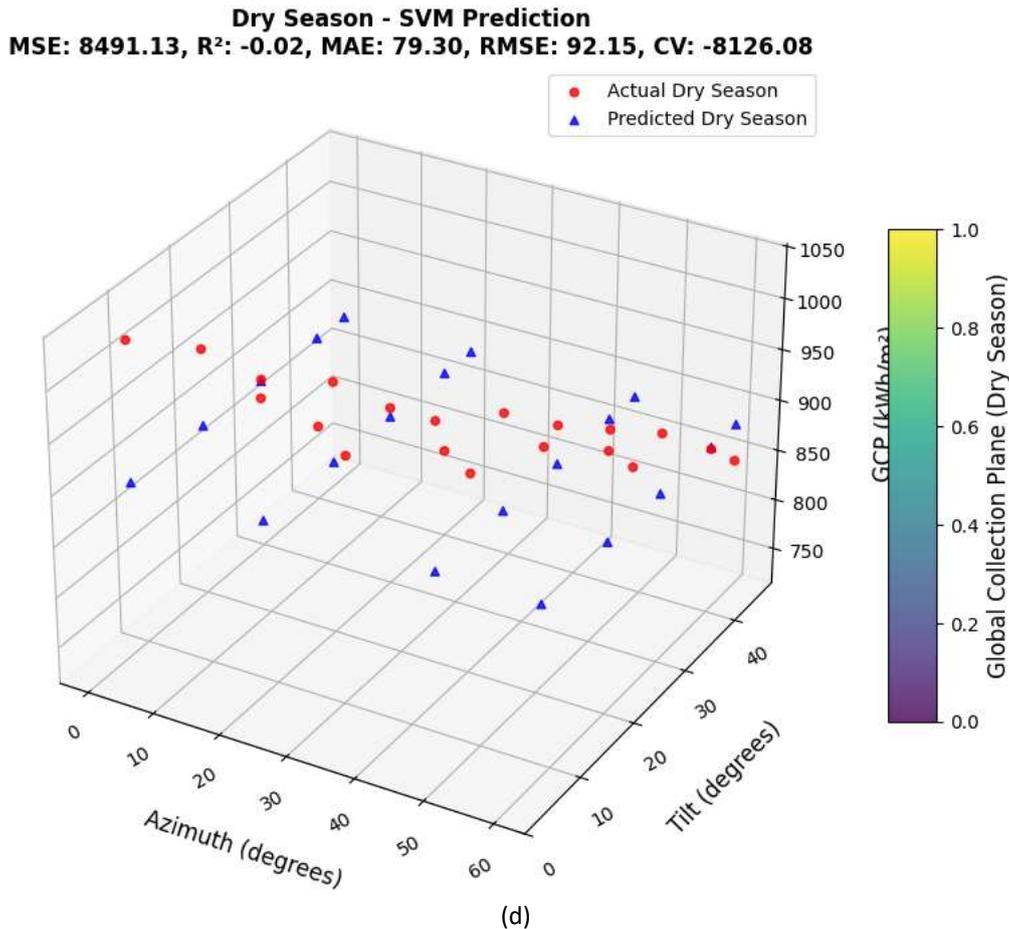


Fig. 8: RF and SVM models for analysing azimuth and tilt angles in (a-b) wet season, (c-d) dry season

In the wet season, the RF prediction graph describes the actual Gcp values as closely aligned with the predicted values across most azimuth and tilt angle combinations, as shown in Figure 8(a). This indicates the RF model's ability to generalize effectively and capture underlying patterns in the data. The RF model achieved a mean squared error (MSE) of 86.84, mean absolute error (MAE) of 6.21, root mean squared error (RMSE) of 9.32, while a cross-validation (CV) score of -346.82 and R^2 of 0.93 reflecting the model's effectiveness in predicting the Gcp during the wet season. The relatively high R^2 value suggests that the RF model explains a significant portion of the variance in the Gcp. However, the SVM model reveals a less accurate prediction, as indicated by a MSE of 1145.16, MAE of 28.06 and a RMSE of 33.84, CV score of -1776.01 and an R^2 of 0.10, which suggests that the SVM model was unable to capture underlying patterns in the data during the wet season. The superior performance of RF can be attributed to its ensemble learning approach, which effectively captures the complex interactions between input variables, i.e., tilt, azimuth and weather conditions. Nevertheless, the SVM models struggle to accurately predict the Gcp across different azimuth and tilt angles, specifically in the presence of nonlinear relationships in the data during the wet season. It can also be explained through CV results for each model across seasons. For example, a CV score of -1776 , the significantly larger negative CV score suggests that the SVM model struggles to generalize effectively for wet season data. This poor performance may be due to the model's inability to efficiently capture the nonlinear relationships between azimuth, tilt and solar irradiance. A high negative score implies overfitting training data, resulting in poor adaptability to new data.

During the dry season, the RF model outperformed the SVM model, throughout the metric evaluation. The Random Forest model obtained an MSE of 128.36, MAE of 8.82, RMSE of 11.33, CV of -368.84 and an R^2 of 0.98, which is a slight drop in performance compared to the wet season but still indicates a reasonably good fit. The SVM model's performance worsened in the dry season, with an MSE of 8491.13, MAE of 79.30, RMSE of 92.15, CV of -8126.08 and an R^2 of -0.02 . The substantial increase in MSE and the highly negative R^2 value for the SVM model suggest it was unable to model the relationship between the input features and the global collection plane effectively during the dry season. The graphs for the dry season show that the RF model's predictions still maintain a close alignment with actual data points. However, the SVM model displays a greater discrepancy between predicted and actual values, particularly in lower global collection planes, leading to its poor performance metrics. While potent in certain scenarios, SVM may struggle with larger datasets or more complex relationships without appropriate kernel selection and parameter tuning. This limitation was evident in this study's higher MSE observed during dry and wet seasons.

The primary difference between the RF and SVM models in both seasons is their capacity to generalize and capture the relationship between the input variables, namely azimuth and tilt angles and the target variable, i.e., the global collection plane. RF, being an ensemble learning method, leverages multiple decision trees to model complex interactions between variables. This inherent flexibility allows RF to perform well in varying seasonal conditions [45]. In contrast, SVM's reliance on finding the optimal hyperplane in a transformed feature space may struggle in nonlinear and complex datasets, as evidenced by its poor performance, particularly in the dry season.

Numerous earlier studies have employed various MLAs, as explained in the section introduction, for instance, gradient boosting machines (GBM) and neural networks (NN) offer competitive performance which often requires extensive hyperparameter tuning, larger datasets and the need for interpretability. Moreover, Neural networks, although powerful, may suffer from overfitting with smaller datasets and require significant training time and computational power, as experimented in [46,47], which is not always practical in resource-limited settings.

Our findings showed that RF and SVM are well-suited for capturing complex, nonlinear relationships between input variables, such as azimuth and tilt angles and the output variable, i.e., the global collection plane of solar panels. RF, in particular, is an ensemble learning method that reduces overfitting through averaging multiple decision trees, making it robust against noise and variability in meteorological data. SVM is effective in high-dimensional spaces and provides a clear boundary for classification tasks, making it suitable for analysing diverse environmental conditions. Nonetheless, as presented in our study, RF demonstrated superior predictive accuracy, with a lower MSE compared to SVM. This further validated the choice of RF as an effective model for predicting optimal solar panel configurations. The SVM model provided valuable insight into the general trends but showed limitations in handling highly nonlinear patterns in the data, reinforcing the advantage of ensemble approaches like RF. The comparison of model performance is presented in Figure 9 below.

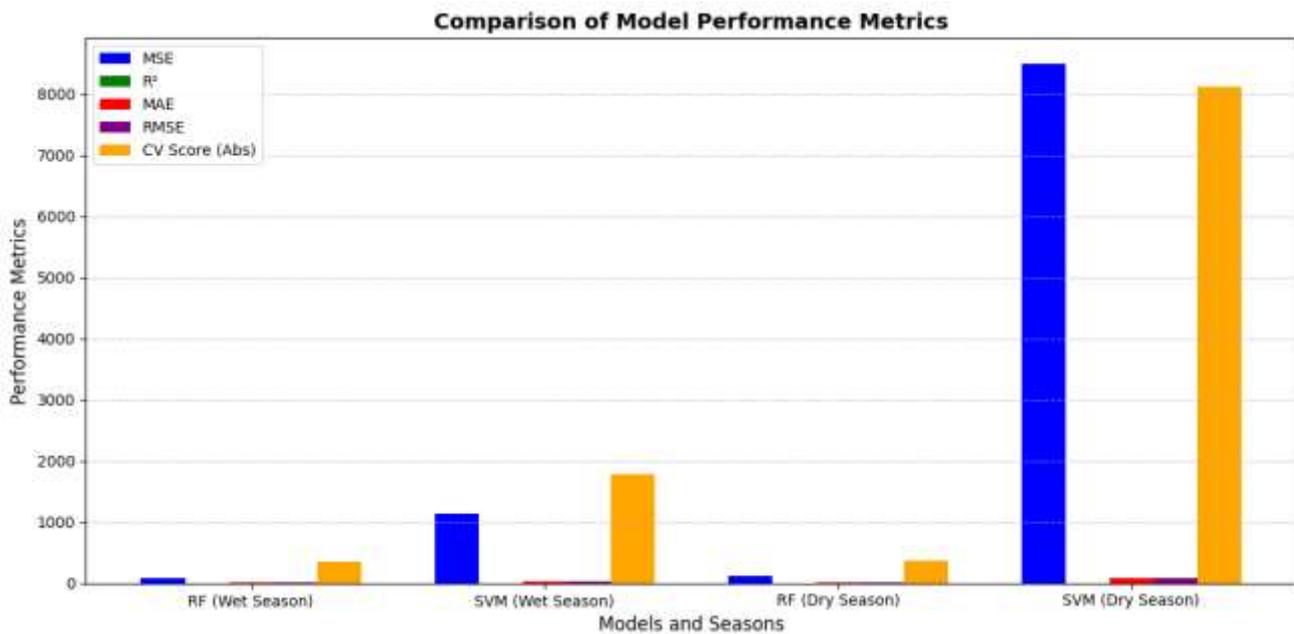


Fig. 9. The MSE, R², MAE, RMSE and CV score of each predictor

The findings in our study align with the results by Scott *et al.*, [48], whose study also investigated the use of machine learning to predict PV energy generation in an operational university campus in Manchester, UK. The result demonstrated that Random Forest achieved the lowest average RMSE of 32, outperforming Support Vector Machine (32.3 RMSE), Linear Regression (36.5 RMSE) and Neural Network (38.9 RMSE), even with a small dataset. Our findings also reveal that the choice of MLAs depends on the dataset characteristics and forecast requirements, with RF proving to be a reliable choice for PV forecasting due to its lower computational requirements and higher accuracy when compared to SVM. The findings emphasize the importance of selecting an appropriate MLA based on accuracy, usability and computational efficiency for real-world applications in energy management systems.

4. Conclusions

This study investigated the prediction of PV energy output using RF and SVM models, focusing on varying azimuth and tilt angles during wet and dry seasons. The Random Forest model identified that a tilt angle of 8° combined with an azimuth angle of 60° in the dry season provided the highest energy output, aligning well with the sun’s trajectory during these months. This configuration maximizes direct solar irradiance capture, as shown by the close match between actual and predicted values. During the wet season, the RF model indicated that a tilt angle of 30° and an azimuth of 0° were optimal, reflecting the need to capture diffuse solar radiation more effectively due to increased cloud cover. The SVM model also demonstrated its effectiveness in capturing the complex, nonlinear relationships between azimuth, tilt angles and solar energy output. The model confirmed the Random Forest results, with the dry season’s optimal configuration being a tilt angle of 8° and an azimuth of 60° and the wet season’s optimal being a tilt angle of 30° and an azimuth of 0°.

The RF achieved lower MSE and higher R² values, indicating a better fit and more accurate predictions. In the wet season, RF’s performance was notably strong, with an MSE of 86.84 and R² of 0.93, compared to SVM’s MSE of 1145.16 and a small R² value of 0.10, indicating poor model performance. Similarly, in the dry season, Random Forest achieved an MSE of 128.36 and an R² of 0.98, while SVM showed significantly reduced performance with an MSE of 8491.13 and an R² of

-0.02. The study revealed that RF, as an ensemble method, is better suited to handle the complex interactions between azimuth, tilt angles and Gcp, specifically under varying seasonal conditions. Conversely, SVM, based on finding the optimal hyperplane in a transformed feature space, struggles to model these relationships effectively, particularly in the dry season.

Our findings underscore the potential of machine learning models to improve solar energy optimization, supporting the broader adoption of renewable energy technologies in diverse environmental contexts. Therefore, various environmental factors, cloud cover, albedo and shading in different climatic and geographic areas in South Papua need to be taken into account, as well as providing a comprehensive and relevant dataset for further study.

Acknowledgment

The authors express our appreciation for the support and assistance extended by Universitas Musamus and Universitas Negeri Yogyakarta during the research process. The data presented in this article, depicted through figures and tables, constitute a portion of the ongoing research carried out for the doctoral thesis of the primary author. The field data, along with the study map, was collected before the announcement of the new administrative framework of the region.

References

- [1] Institute for Essential Services Reform (IESR). "Laporan Status Energi Bersih Indonesia: Potensi, Kapasitas Terpasang, dan Rencana Pembangunan Pembangkit Listrik Energi Terbarukan 2019." *IESR*, (2019). Available: <https://iesr.or.id/infografis/laporan-status-energi-terbarukan-indonesia>
- [2] Letsoin, Sri Murniani Angelina, David Herak and Ratna Chrismiari Purwestri. "Evaluation land use cover changes over 29 years in papua province of indonesia using remote sensing data." In *IOP Conference Series: Earth and Environmental Science*, vol. 1034, no. 1, p. 012013. IOP Publishing, 2022. <https://doi.org/10.1088/1755-1315/1034/1/012013>
- [3] Ashetehe, Ahunim Abebe, Belachew Bantayirga Gessesse and Fekadu Shewarega. "A generalized approach for the determination of optimum tilt angle for solar photovoltaic modules with selected locations in Ethiopia as illustration examples." *Scientific African* 18 (2022): e01433. <https://doi.org/10.1016/j.sciaf.2022.e01433>
- [4] Babatunde, A. A., Sij Abbasoglu and M. Senol. "Analysis of the impact of dust, tilt angle and orientation on performance of PV Plants." *Renewable and Sustainable Energy Reviews* 90 (2018): 1017-1026. <https://doi.org/10.1016/j.rser.2018.03.102>
- [5] Barbón, Arsenio, Covadonga Bayón-Cueli, Luis Bayón and C. Rodríguez-Suanzes. "Analysis of the tilt and azimuth angles of photovoltaic systems in non-ideal positions for urban applications." *Applied Energy* 305 (2022): 117802. <https://doi.org/10.1016/j.apenergy.2021.117802>
- [6] N'Tsoukpoe, Kokouvi Edem. "Effect of orientation and tilt angles of solar collectors on their performance: Analysis of the relevance of general recommendations in the West and Central African context." *Scientific African* 15 (2022): e01069. <https://doi.org/10.1016/j.sciaf.2021.e01069>
- [7] Elhamaoui, Said, Aboubakr Benazzouz, Abderrazzak Elamim, Ibtihal Ait Abdelmoula, Khalil Tijani, Abdellatif Ghennioui and Mohamed El Khaili. "Long-term outdoor performance and degradation evaluation of CIS PV plant under the semi-arid climate of Benguerir Morocco." *Energy Reports* 9 (2023): 322-339. <https://doi.org/10.1016/j.egy.2023.05.272>
- [8] Srivastava, Rachit, A. N. Tiwari and V. K. Giri. "An overview on performance of PV plants commissioned at different places in the world." *Energy for Sustainable Development* 54 (2020): 51-59. <https://doi.org/10.1016/j.esd.2019.10.004>
- [9] Al-Ibrahim, Ehab. "Impact of Dust and Shade on Solar Panel Efficiency and Development of a Simple Method for Measuring the Impact of Dust in any Location." *Journal of Sustainable Development of Energy, Water and Environment Systems* 11, no. 2 (2023): 1-14. <https://doi.org/10.13044/j.sdewes.d11.0448>
- [10] Jamal, Jamal, Ilyas Mansur, Adam Rasid, Musrady Mulyadi, Muhammad Dihyah Marwan and Marwan Marwan. "Evaluating the shading effect of photovoltaic panels to optimize the performance ratio of a solar power system." *Results in Engineering* 21 (2024). <https://doi.org/10.1016/j.rineng.2024.101878>
- [11] Tarigan, Elieser. "Azimuth Angle Impact on Specific Energy Output of Rooftops PV System in Surabaya, Indonesia." In *2022 9th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, pp. 25-28. IEEE, 2022. <https://doi.org/10.1109/ICITACEE55701.2022.9923975>

- [12] Nicolás-Martín, Carolina, David Santos-Martín, Mónica Chinchilla-Sánchez and Scott Lemon. "A global annual optimum tilt angle model for photovoltaic generation to use in the absence of local meteorological data." *Renewable Energy* 161 (2020): 722-735. <https://doi.org/10.1016/j.renene.2020.07.098>
- [13] Yakup, Mohd Azmi bin Hj Mohd and A. Q. Malik. "Optimum tilt angle and orientation for solar collector in Brunei Darussalam." *Renewable energy* 24, no. 2 (2001): 223-234. [https://doi.org/10.1016/S0960-1481\(00\)00168-3](https://doi.org/10.1016/S0960-1481(00)00168-3)
- [14] Elsharif, Anas MH, Tarek AM Hamad, Mohamed Almakhtar, A. M. Elbrek and Fasial Mohamed. "Optimization and Influence of Tilt Angle on PV Module Performance: Experimental Study." In *2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*, pp. 765-770. IEEE, 2023. <https://doi.org/10.1109/MI-STA57575.2023.10169285>
- [15] Liu, Xianping. "Calculation and analysis of optimal tilt angle for PV/T hybrid collector." In *2012 Second International Conference on Intelligent System Design and Engineering Application*, pp. 791-795. IEEE, 2012. <https://doi.org/10.1109/ISdea.2012.702>
- [16] Xu, Ruidong, Kai Ni, Yihua Hu, Jikai Si, Huiqing Wen and Dongsheng Yu. "Analysis of the optimum tilt angle for a soiled PV panel." *Energy Conversion and Management* 148 (2017): 100-109. <https://doi.org/10.1016/j.enconman.2017.05.058>
- [17] Rowlands, Ian H., Briana Paige Kemery and Ian Beausoleil-Morrison. "Optimal solar-PV tilt angle and azimuth: An Ontario (Canada) case-study." *Energy Policy* 39, no. 3 (2011): 1397-1409. <https://doi.org/10.1016/j.enpol.2010.12.012>
- [18] Chen, X. M., Y. Li, Z. G. Zhao, T. Ma and R. Z. Wang. "General method to obtain recommended tilt and azimuth angles for photovoltaic systems worldwide." *Solar Energy* 172 (2018): 46-57. <https://doi.org/10.1016/j.solener.2018.06.045>
- [19] Morcos, V. H. "Optimum tilt angle and orientation for solar collectors in Assiut, Egypt." *Renewable energy* 4, no. 3 (1994): 291-298. [https://doi.org/10.1016/0960-1481\(94\)90032-9](https://doi.org/10.1016/0960-1481(94)90032-9)
- [20] Kallioğlu, Mehmet Ali, Ali Serkan Avcı, Ashutosh Sharma, Rohit Khargotra and Tej Singh. "Solar collector tilt angle optimization for agrivoltaic systems." *Case Studies in Thermal Engineering* 54 (2024). <https://doi.org/10.1016/j.csite.2024.103998>
- [21] Bakirci, Kadir. "General models for optimum tilt angles of solar panels: Turkey case study." *Renewable and Sustainable Energy Reviews* 16, no. 8 (2012): 6149-6159. <https://doi.org/10.1016/j.rser.2012.07.009>
- [22] Abdeen, Eltaib, Mohamed Orabi and El-Sayed Hasaneen. "Optimum tilt angle for photovoltaic system in desert environment." *Solar energy* 155 (2017): 267-280. <https://doi.org/10.1016/j.solener.2017.06.031>
- [23] Alghamdi, Hisham, Chika Maduabuchi, Abdullah Albaker, Abdulaziz Almalaq, Turki Alsuwian and Ibrahim Alatawi. "Machine Learning Performance Prediction of a Solar Photovoltaic-Thermoelectric System with Various Crystalline Silicon Cell Types." *International Journal of Energy Research* 2023, no. 1 (2023): 1990593. <https://doi.org/10.1155/2023/1990593>
- [24] Ahad, Nor Aishah, Friday Zinzendoff Okwonu, Yik Siong Pang and Olimjon Shukurovich Sharipov. "A Comparative Performance Analysis of Several Machine Learning Classifiers on the Credit Card Data." *Journal of Advanced Research in Computing and Applications* 37, no. 1 (2024): 50-64. <https://doi.org/10.37934/arca.37.1.5064>
- [25] Aburashed, Laila, Marah AL Amoush and Wardeh Alrefai. "SQL injection attack detection using machine learning algorithms." *Semarak International Journal of Machine Learning* 2, no. 1 (2024): 1-12. <https://doi.org/10.37934/sijml.2.1.112>
- [26] Letsoin, Sri Murniani Angelina, David Herak, Fajar Rahmawan and Ratna Chrismiari Purwestri. "Land cover changes from 1990 to 2019 in Papua, Indonesia: Results of the remote sensing imagery." *Sustainability* 12, no. 16 (2020): 6623. <https://doi.org/10.3390/su12166623>
- [27] Osmani, Khaled, Mohamad Ramadan, Thierry Lemenand, Bruno Castanier and Ahmad Haddad. "Optimization of PV array tilt angle for minimum levelized cost of energy." *Computers & Electrical Engineering* 96 (2021): 107474. <https://doi.org/10.1016/j.compeleceng.2021.107474>
- [28] Kaddoura, Tarek O., Makbul AM Ramli and Yusuf A. Al-Turki. "On the estimation of the optimum tilt angle of PV panel in Saudi Arabia." *Renewable and Sustainable Energy Reviews* 65 (2016): 626-634. <https://doi.org/10.1016/j.rser.2016.07.032>
- [29] Obiwulu, Anthony Umunnakwe, Nald Erusiafe, Muteeu Abayomi Olopade and Samuel Chukwujindu Nwokolo. "Modeling and estimation of the optimal tilt angle, maximum incident solar radiation and global radiation index of the photovoltaic system." *Heliyon* 8, no. 6 (2022). <https://doi.org/10.1016/j.heliyon.2022.e09598>
- [30] Stanciu, Dorin, Camelia Stanciu and Ioana Paraschiv. "Mathematical links between optimum solar collector tilts in isotropic sky for intercepting maximum solar irradiance." *Journal of Atmospheric and Solar-Terrestrial Physics* 137 (2016): 58-65. <https://doi.org/10.1016/j.jastp.2015.11.020>

- [31] Masrur, Hasan, Mohammad Lutfi Othman, Ahmed Amirul Arefin, Noor Izzri Abd Wahab, Syed Zahurul Islam and Tomonobu Senju. "Determining optimal tilt angle to maximize the PV yield." In *2020 IEEE International Conference on Power and Energy (PECon)*, pp. 219-223. IEEE, 2020. <https://doi.org/10.1109/PECon48942.2020.9314455>
- [32] Mukisa, Nicholas and Ramon Zamora. "Optimal tilt angle for solar photovoltaic modules on pitched rooftops: A case of low latitude equatorial region." *Sustainable Energy Technologies and Assessments* 50 (2022): 101821. <https://doi.org/10.1016/j.seta.2021.101821>
- [33] Lu, Ning and Jun Qin. "Optimization of tilt angle for PV in China with long-term hourly surface solar radiation." *Renewable Energy* 229 (2024): 120741. <https://doi.org/10.1016/j.renene.2024.120741>
- [34] Wei, Donghui, Ali Basem, As' ad Alizadeh, Dheyaa J. Jasim, Haydar AS Aljaafari, Mohammadali Fazilati, Babak Mehmandoust and Soheil Salahshour. "Optimum tilt and azimuth angles of heat pipe solar collector, an experimental approach." *Case Studies in Thermal Engineering* 55 (2024): 104083. <https://doi.org/10.1016/j.csite.2024.104083>
- [35] Le Roux, Willem Gabriel. "Optimum tilt and azimuth angles for fixed solar collectors in South Africa using measured data." *Renewable Energy* 96 (2016): 603-612. <https://doi.org/10.1016/j.renene.2016.05.003>
- [36] Teyabeen, Alhassan Ali and Faisal Mohamed. "Modelling and Estimation of The Optimum Tilt Angle of Photovoltaic Systems, Case Study: Libya." In *2022 13th International Renewable Energy Congress (IREC)*, pp. 1-7. IEEE, 2022. <https://doi.org/10.1109/IREC56325.2022.10002019>
- [37] Kallioğlu, Mehmet Ali, Aydın Durmuş, Hakan Karakaya and Adem Yılmaz. "Empirical calculation of the optimal tilt angle for solar collectors in northern hemisphere." *Energy Sources, Part A: Recovery, Utilization and Environmental Effects* 42, no. 11 (2020): 1335-1358. <https://doi.org/10.1080/15567036.2019.1663315>
- [38] Chinchilla, Monica, David Santos-Martín, Miguel Carpintero-Rentería and Scott Lemon. "Worldwide annual optimum tilt angle model for solar collectors and photovoltaic systems in the absence of site meteorological data." *Applied Energy* 281 (2021): 116056. <https://doi.org/10.1016/j.apenergy.2020.116056>
- [39] Sharma, Ashutosh, Mehmet Ali Kallioğlu, Anchal Awasthi, Ranchan Chauhan, Gusztáv Fekete and Tej Singh. "Correlation formulation for optimum tilt angle for maximizing the solar radiation on solar collector in the Western Himalayan region." *Case Studies in Thermal Engineering* 26 (2021): 101185. <https://doi.org/10.1016/j.csite.2021.101185>
- [40] Al-Ghussain, Loiy, Onur Taylan, Mohammad Abujubbeh and Muhammed A. Hassan. "Optimizing the orientation of solar photovoltaic systems considering the effects of irradiation and cell temperature models with dust accumulation." *Solar Energy* 249 (2023): 67-80. <https://doi.org/10.1016/j.solener.2022.11.029>
- [41] Jatoi, Abdul Rehman, Saleem Raza Samo and Abdul Qayoom Jakhrani. "Performance evaluation of various photovoltaic module technologies at Nawabshah Pakistan." *International Journal of Renewable Energy Development* 10, no. 1 (2021): 97. <https://doi.org/10.14710/ijred.2021.32352>
- [42] Rusiana, Iskandar Handoko, Zainal Yuda Bakti and Susanto Sambasri. "Study and analysis of shading effects on photovoltaic application system." In *MATEC Web of Conferences*, vol. 218, p. 02004. EDP Sciences, 2018. <https://doi.org/10.1051/mateconf/201821802004>
- [43] Lee, Hyomun, Minjoo Choi, Jaewon Kim, Dongsu Kim, Sangsoo Bae and Jongho Yoon. "Partial shading losses mitigation of a rooftop PV system with DC power optimizers based on operation and simulation-based evaluation." *Solar Energy* 265 (2023): 112053. <https://doi.org/10.1016/j.solener.2023.112053>
- [44] Trzmiel, G., D. Głuchy and D. Kurz. "The impact of shading on the exploitation of photovoltaic installations." *Renewable Energy* 153 (2020): 480-498. <https://doi.org/10.1016/j.renene.2020.02.010>
- [45] Saputra, Dimas Chaerul Ekty, Alfian Ma'arif and Khamron Sunat. "Optimizing predictive performance: Hyperparameter tuning in stacked multi-kernel support vector machine random Forest models for diabetes identification." *Journal of Robotics and Control (JRC)* 4, no. 6 (2023): 896-904. <https://doi.org/10.18196/jrc.v4i6.20898>
- [46] Letsoin, Sri Murniani Angelina, Ratna Chrismiari Purwestri, Fajar Rahmawan and David Herak. "Recognition of sago palm trees based on transfer learning." *Remote Sensing* 14, no. 19 (2022): 4932. <https://doi.org/10.3390/rs14194932>
- [47] Letsoin, Sri Murniani Angelina and David Herak. "The effect of parameter adjustment in sago palm classification-based convolutional neural network (CNN) model." *Research in Agricultural Engineering* 70, no. 3 (2024). <https://doi.org/10.17221/65/2023-RAE>
- [48] Scott, Connor, Mominul Ahsan and Alhussein Albarbar. "Machine learning for forecasting a photovoltaic (PV) generation system." *Energy* 278 (2023): 127807. <https://doi.org/10.1016/j.energy.2023.127807>