



Comprehensive Evaluation of Machine Learning MPPT Algorithms for a PV System Under Different Weather Conditions

Mpho Sam Nkambule¹ · Ali N. Hasan² · Ahmed Ali¹ · Junhee Hong³ · Zong Woo Geem³

Received: 16 May 2020 / Revised: 16 October 2020 / Accepted: 23 October 2020
© The Korean Institute of Electrical Engineers 2020

Abstract

The rapid growth of demand for electrical energy and the depletion of fossil fuels opened the door for renewable energy; with solar energy being one of the most popular sources, as it is considered pollution free, freely available and requires minimal maintenance. This paper investigates the feasibility of using machine learning (ML) based MPPT techniques, to harness maximum power on a PV system under PSC. In this study, certain contributions to the field of PV systems and ML based systems were made by introducing nine (9) ML based MPPT techniques, by presenting three (3) experiments under different weather conditions. Decision Tree (DT), Multivariate Linear Regression (MLR), Gaussian Process Regression (GPR), Weighted K-Nearest Neighbors (WK-NN), Linear Discriminant Analysis (LDA), Bagged Tree (BT), Naïve Bayes classifier (NBC), Support Vector Machine (SVM) and Recurrent Neural Network (RNN) performances are validated and proved using MATLAB SIMULINK simulation software. The experimental results demonstrated that WK-NN performs significantly better when compared with other proposed ML based algorithms.

Keywords Maximum power point tracking (MPPT) · Partial shading conditions (PSC) · Machine learning (ML) · DC–DC boost converter

1 Introduction

In 1839, the photovoltaic effect was discovered when using sunlight to generate electrical energy by French scientist Edmond Becquerel. He was only 19 years old when he discovered that placing silver chloride in an acidic solution

would result in generating current and voltage. The Solar Energy Conversion System has been attracting wide attention as a renewable energy source, due to the increased cost of fossil fuel and its environmental impact. Solar renewable resources are regarded as a cleaner source of energy and they reduce greenhouse gases [1, 2].

In between 1990 and 2019, the Photovoltaic Solar System has grown exponentially worldwide. In 1996, the United States (US) became the first country in history to produce 77 MW of the power through a PV solar system. In about the year 1999, Japan became the world's leader, producing more solar power than any other country. In the mid-2000s, Germany surpassed Japan to produce over 41GW of solar power. China became the world's biggest solar power producer to date. It made history by producing more than 100GW of solar power more than any other country in the world. In 2018, PV systems were used to generate a cumulative global amount of power of 512GW [3]. The subject of the energy catastrophe and the environmental study has attracted further attention in this era. Solar PV research expanded due to its merits such as being a green energy, its global availability, inexpensive maintenance costs and being noise-free [4].

✉ Zong Woo Geem
geem@gachon.ac.kr

Mpho Sam Nkambule
nkambule833@gmail.com

Ali N. Hasan
ahasan1@hct.ac.ae

Ahmed Ali
aali@uj.ac.za

Junhee Hong
hongpa@gachon.ac.kr

¹ Department of Electrical Engineering, University of Johannesburg, Johannesburg, South Africa

² Department of Electrical Engineering, Higher Colleges of Technology, Abu Dhabi, United Arab Emirates

³ Department of Energy IT, Gachon University, Seongnam 13120, South Korea

Despite the advantages rendered by the PV system, it still experiences some restraints. A PV system without an MPPT controller operates at a very low efficiency level. Consequently, the necessity arose for an MPPT controller to harvest maximum power out of a PV system. Most MPPT algorithms have been proposed to track MPP in a photovoltaic system [4]. Nevertheless, some of the MPPT techniques failed to track GMPP under PSC. This resulted in PV system power losses and thus low efficiency operations [5, 6].

Good examples of conventional methods are introduced in [7], where Perturb & Observe (P&O) and Incremental Conductance (INC) are studied. These methods have merits such as being easy to implement and program. However, they cannot assure a desired GMPP under changing environmental conditions [8]. They fail to distinguish the global power amongst all other points, which results in power loss.

The researcher [9] used Fractional open circuit voltage (FVoc) method to ensure that the PV system operated at MPP. FVoc is one of the most widely used techniques in the previous years. This algorithm utilizes the concept of a PV array voltage that should be in proportion to Voc at MPP. This depends on the solar cell's material, fabrication and weather conditions. It is based on $V_{mpp} = k1 \times V_o$. The V_{mpp} is determined by measuring the Voc by switching off the power.

The Short Circuit Current algorithm was initiated in previous years by various investigators. This technique works identically to Voc algorithm. It utilizes the concept of the linear relationship between a PV array (I_{mpp}) and short circuit current (I_{sc}) [10–12]. It is based on $I_{mpp} = k2 \times I_{sc}$. Both techniques are inessential under PSC. Both algorithms suffer from a temporary power loss, have a low convergence speed and are not precise under PSC. Under PSC, they both fail to operate at the maximum power due to the P–V curve characteristics and are not able to identify the true power peaks.

In [13], the authors proposed Artificial Intelligence (AI) MPPT techniques to overcome the shortcomings of the conventional methods. The author used a Feedforward Neural Network (FNN). It assesses the inputs by utilizing the feedforward method, differentiates its sum by means of activation functions, and sends the results to the other neurons. The author used the irradiance and temperature as inputs and voltage at MPP as the output. The algorithm has high tracking speed, better performance and a fast settling time around MPP [14]. Despite its good performance, FNN suffers from not having an internal state memory, so it cannot process the sequence of input and requires a lot of computational time.

In [15], the researchers proposed and implemented the Artificial Neuro-Fuzzy Inference System (ANFIS). It is one of the hybrid methods, which combines the Fuzzy Logic Control (FLC) with the ANN. It obviously has an advantage over separate NN and FLC. It uses the Sugeno Fuzzy

Inference System (FIS), which is computational, powerful and performs effectively for non-linear P–V and I–V characteristics and improves the efficiency of the PV panel. The ANFIS data training also requires historical data, which can be real time process data and dynamic PV model simulated data. A Power Integral Derivative (PID) is employed on the ANFIS as error signal control. This reduces the error. However, ANFIS has a vigorous computational complexity limitation. It also has surface oscillations around points caused by high partition numbers.

Other researchers in [16] proposed the Optimization MPPT technique to overcome the demerits of the conventional methods. A metaheuristic technique such as Particle Swarm Optimization (PSO) is one of the powerful algorithms that is inspired by research on bird flocking behavior in nature [17]. It is a simple computational structure method that sustains a flock of particles searching over solutions carried by their particles. The individual particles compete with each other's success. Each particle's position defends the best experience of the nearest particle, while they interrelate with each other to share their understanding. PSO is based on Global Best position (GB) and Local Best position (LB). The GB and the LB represent the generated power by an exact operating voltage. The authors proved that the method could identify the GMPP. However, this method is complex to implement, and it takes time to track GMPP. In fact, the robustness, and the ability to discern GMPP under expeditious change of insolation and the ease/low cost of implementation, must be part of the MPPT algorithm specifications.

In [18], the researcher introduced Ant Colony Optimization (ACO). This method is inspired by the deportment or the conduct of ants and their way of communication. This optimization method works well on non-linear studies. Ants usually find the shortest route when doing a random search. This involves pheromones [19]. It is based on probability studies; the thickness of the pheromones grows as the number of ants that use the same route increases; hence, the probability that other ants will use the same route becomes greater. Eventually, by means of conveying information, a gradient showing the number of ants follows the route until the shortest route is tracked [20]. Regardless of its good performance, this technique has some disadvantages such as the complexity of implementation, the high cost and the need for comprehensive computation.

In [21], the author implemented the Genetic Algorithm (GA) strategy. That is an optimization method initially established by Holland [22]. It is based on solving natural and unnatural optimization problems, based on a biological evolutionary process. It constantly improves the population density into a single pattern. It nominates a random recent population to be the mother and it utilizes it to make children for the next generation. Over consecutive generations,

the population drifts towards the ideal solution. The GA MPPT algorithm has three steps; selection, crossover and mutation. In the selection step, the ideal number of candidates (chromosomes) are selected, based on their level of fitness for insertion into the upcoming generation. However, this method is very expensive to implement, and it is too complex.

In this paper, seven (7) advanced ML MPPT algorithms are proposed and compared with each other, using real-time data under PSC. The main objective is to harvest optimal power from a PV module under PSC. Currently most of the MPPT methods have failed to track global maximal power. This causes a PV system to operate at a lower efficiency level, which can damage the batteries and the load connected to it. This is caused by the methods that have large errors after training, testing and validation.

ML techniques depend on the input data for training. Obviously, inaccurate data will result in large errors. This paper proposed the cascading of the MPPT and the PID controllers. It is used to correct and further fine-tune the large errors into smaller errors. It also improves the accuracy of the MPPT algorithm, and ultimately improves the efficiency of the PV panel. The main contribution of this paper is as follows: (i) Highlighting the importance of the MPPT controllers under PSC. (ii) Using different MPPT techniques to harness MPP using Johannesburg real-time data. (iii) Employing a PID controller to reduce error and consequently enhance the performance of MPPT controllers. (iv) Using ML techniques to accelerate tracking speed and eliminate misleading power or minimize mismatching power under PSC.

The synopsis of this paper is arranged as follows: Section II presents the modelling of a PV cell. Section III explains MPPT and the necessity for MPPT algorithms under PSC. Section IV presents the seven machine learning techniques used in this research. Section V explains the system setup. Section VI demonstrates the experimental results and discussions. Section VI concludes this paper.

2 Basic Characteristic of a PV Cell

2.1 Modelling of PV Cell

The Modelling of a PV Cell depends on the environmental factors such as solar insolation and temperature. Other minor factors can be the PV cell material used, and the degree of the PV cell. The Modelling of a PV cell comprises a single diode (practical diode) as shown in Fig. 1 below [23].

Figure 1 above illustrates the PV solar cell equivalent circuit diagram that comprises a shunt resistance (R_{sh}) and a series resistance (R_s); where R_{sh} represents the leakage current losses and R_s represents internal resistance to the

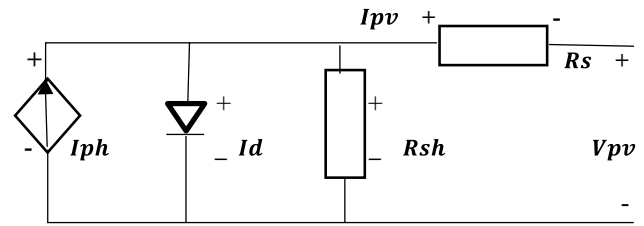


Fig. 1 PV cell equivalent circuit diagram

current flow [24–26]. Table 1 below exemplifies the parameters used to model the PV cell. The mathematical model of this PV cell is given in the following equations:

$$V_t = \frac{KT_c}{Q}, \quad (1)$$

$$a = \frac{NsAKT_c}{Q} = NsAV_t, \quad (2)$$

$$I_{pv} = I_{ph} - I_d - I_{sh}, \quad (3)$$

$$I_d = I_o \left[e^{\frac{V_{pv} + I_{sh} R_s}{a}} - 1 \right], \quad (4)$$

$$I_{sh} = \frac{V + I R_s}{R_p}, \quad (5)$$

Table 1 Photovoltaic parameters

Parameters	Description
k	Boltzmann's constant = $1.38e^{-23}$
V_t	Terminal voltage
q	Charge of an electron = $1.6e^{-19}$
T_{ref}	Reference temperature = 298K
T_c	Actual cell temperature in Kelvin
A	Ideality factor
a	Modified ideality factor
I_{ph}	Photon current
N_s	Number of cells connected in series
N_p	Number of cells connected in parallel
I_d	Diode current
I_{pv}	PV current
I_o	Leakage current
V_{pv}	PV voltage
G_{ref}	Reference insolation at STC = $1000W/m^2$
K_i	Coefficient of short circuit temperature
R_s	Series resistance
R_{sh}	Shunt resistance

$$I_{pv} = I_{ph} - I_o \left[e^{\frac{V_{pv} + I_{pv} R_s}{a}} - 1 \right] - \frac{V + I_{pv} R_s}{R_{sh}}, \quad (6)$$

$$I_{sh} = \frac{G}{G_{ref}} [I_{ph}(ref) + k_i(T_c - T_{ref})], \quad (7)$$

$$I_o = I_o(ref) \left(\frac{T_c}{T_{ref}} \right)^3 \times e^{\left[\left(\frac{qE_g}{AK} \right) \left(\frac{1}{T_{ref}} - \frac{1}{T_c} \right) \right]}, \quad (8)$$

$$V_{pv} = -I_{pv} R_s + K \log \left[\frac{I_{ph} - I_{pv} + I_o}{I_o} \right]. \quad (9)$$

2.2 Partial shading conditions

Generally, one PV module is shaded by partly cloudy weather, flying birds, nearby trees and buildings. This causes one or more PV modules to obtain insufficient solar insolation levels, which in turn produces a low current. The fluctuation of solar irradiance causes multiple power peaks in the P–V characteristic, as shown in Fig. 2 below. The solar irradiance is directly proportional to the PV current and PV power. The other shading can be caused by environmental factors such as heat (temperature), which will reduce or affect the PV voltage. The PV Voltage is inversely proportional to the ambient temperature, which is regarded as power loss. This further causes the PV module to behave as a load and consume power generated by other PV modules [27]. Mismatched power loss takes place, and this poses a threat of damaging the connected load and battery [28].

A bypass diode is then introduced and connected in parallel to each PV module, to avoid this situation, as shown in Fig. 3 below. It allows the string current to flow in the direction even when a PV module is shaded [29, 30]. This will cause multiple power peaks, known as local and global maxima on a P–V curve, as shown in Fig. 2 above.

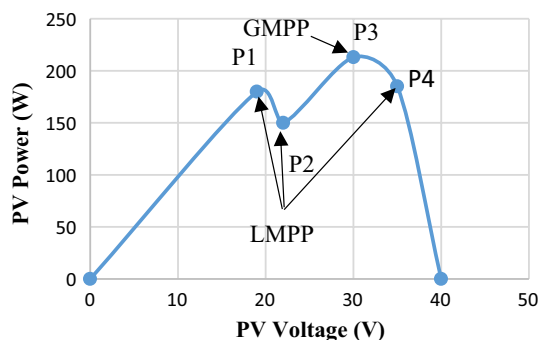


Fig. 2 P–V Curve of a partial shaded PV system

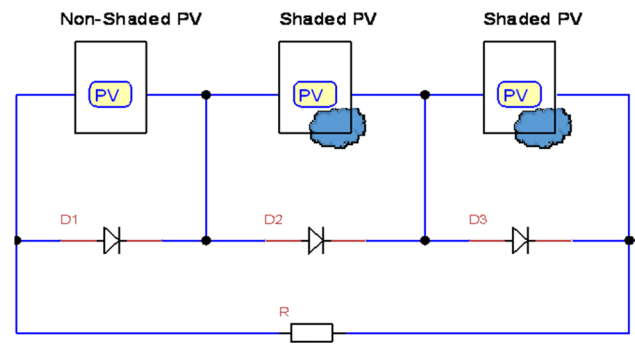


Fig. 3 PV module configuration under different PSCs

3 MPPT Controllers

3.1 MPPT

MPPT is a technique that is employed to ensure that the PV system operates at MPP all the time. MPP is referred to as a point on a P–V curve where $\frac{dp}{dv} = 0$. MPPT algorithms decide the corresponding MPP, as to the duty cycle. The duty ratio is either, enlarged or reduced in the direction of the MPP. The main aim is to match the output PV impedance with load impedance, which is called impedance matching. The impedance matching forces the MPPT to harvest maximum power out of a PV panel.

3.2 Necessity of MPPT Algorithms Under PSC

The PV system is highly affected under sudden changes of solar insolation levels and temperature. This causes the PV system to operate at a very poor efficiency level. This might damage the battery and the load connected to the PV panel. The importance of the MPPT technique is to protect the battery against deep discharging and over-charging, despite harvesting maximum power out of the PV panel. It also protects the load from load variation, which eventually damages the load connected. The main importance of MPPT is to improve the PV Panel operating efficiency. It also accelerates the tracking speed of the MPP, and it reduces the power loss under the PSC. It also minimizes misleading mismatched and power losses under rapidly changing insolation levels.

4 MPPT Techniques

4.1 Decision Tree

A DT classification algorithm is a tree-shaped method used to distinguish dissimilar ways to divide datasets into bough-like divisions. These divisions configure an inverted single tree that begins with a root, testing assigned data per node and categorizing them into a suitable branch, until it

obtains a classification leaf node. DT constructs the explorable model that depicts the applicable set of rules. It is an admired method used for classification, which is comparatively rapid for training and prediction [31–33]. Figure 4 below exemplifies the DT model.

The main aim is to find the optimum DT by reproducing the general error. DT has several merits; it can solve the missing data by crossing both sub-branches for prediction, using a weighted vote. It is also capable of managing nominal attributes. The amounts of divisions are equal to the amount of nominal values. Nominal values are bunched together to form subsets [34]. The only demerit is that it needs a large set of data to construct enough splits. It is a widely used method by researchers for analysis and decision-making [10].

4.2 Weighted K-Nearest Neighbors

WK-NN is a supervised ML algorithm that is employed for regression and classification [35]. It is dissimilar to other Machine Learning algorithms, as it uses estimated values of neighbors. It focuses less on training and more on prediction. It classifies items based on learning sets of data nearest to the item. ‘k’ represents the natural number, which varies within the scale of 3 to 10 [36, 37]. Predominantly, an excellent prediction is when the k value is an odd number. The option of KNN is inspired by being a very easy and indolent method. It utilizes the Euclidean distance equation to determine the gap between training and testing items.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, \quad (10)$$

where n , represents input data set of the sample.

x_i = i th input to the j th sample for estimation.

The average output matching the KNN is determined using the Eq. (11). This is the estimated sample in a Z sized data set.

$$b^j(x^j) = \frac{1}{k} \sum_{m=1}^k t^m, j = 1, \dots, Z, \quad (11)$$

where k index is determined by enlarging sort distances and storing the matching index connected to the KNN.

$$\text{Sort}[d(x^j, y^l)], l = 1, \dots, N, \quad (12)$$

where N = magnitude of the training dataset.

4.3 Gaussian Process Regression

The world is full of unreliability caused by a deficit of awareness. There is therefore a need for machine learning. A Gaussian Process (GP) algorithm is defined by its mean

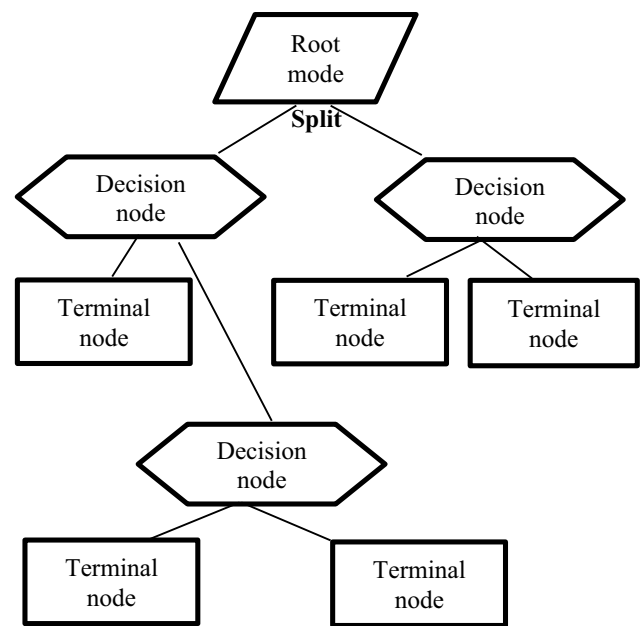


Fig. 4 DT representation [22]

and covariance functions [38]. The GP uses the input training dataset (D) of a process (f) as:

$$D = \{(x_j; y_j) | j = 1, 2, \dots, n\}$$

when n represents observations, x as covariates and y as a response. The trained GP model predicts the global maximum power (Y_{new}) for a new input (X_{new}).

The mean function $m(x)$ is determined by:

$$m(x) = E[f(x)]$$

And the covariance function is given as: $k(x, x')$.

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x')))].$$

And the process of Gaussian is given as:

$$f(x) \sim GP(m(x))(f(x') - m(x')).$$

The GP model is specified as:

$$y = x^T \beta + \epsilon, \dots \epsilon \sim N(0, \sigma^2), \quad (13)$$

where σ and β represents error variance and coefficients, respectively, GP is explicit arbitrary variables count possessing multiple Gaussian distribution. GP model transform (13) by integrating latent variables $f(x_j)$, $j = 1, \dots, n$, from Gaussian process and basis function (h). The $f(x_j)$ guarantees an effortless response [38]. The integration of (1.13) is exemplified in (14).

$$y = h(x)^T \beta + f(x), f(x) \sim GP(0, k(x, x')), \quad (14)$$

where y is given by (15) at any time.

$$P(y_j | f(x_j), x_j) \sim N(y_j | h(x_j)^T \beta + f(x_j), \sigma^2) P(y | f(x_j) \sim N(y | H\beta + f, \sigma^2 I)). \quad (15)$$

4.4 Multivariate Linear Regression

The Multivariable Linear Regression (MLR) algorithm has one or more input (prediction) variables, which in turn, influence the output response. It is commonly used in uncertainty datasets or complicated large dataset learning problems [39]. If the response (Y) has a linear relationship with the predictors, ($x_1, x_2 \dots \dots, x_m$), the Multivariable LR is given by (16):

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m + \epsilon, \quad (16)$$

where β represents the regression coefficient and ϵ represents the training error $E(\epsilon) = 0$.

The main aim is to determine the expected coefficients by substituting the observation values into (17).

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 x_{11} + \dots + \beta_m x_{1m} + \epsilon_1 \\ Y_2 &= \beta_0 + \beta_1 x_{21} + \dots + \beta_m x_{2m} + \epsilon_2 \\ Y_n &= \beta_0 + \beta_1 x_{n1} + \dots + \beta_m x_{nm} + \epsilon_n \end{aligned} \quad (17)$$

The multivariable LR formula should reduce the sum of square of the LR model.

$$Q(\beta) = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \dots - \beta_m x_{im})^2.$$

The regression coefficient matrix is then given by:

$$\beta = (X^T X)^{-1} X^T Y, \quad (18)$$

where:

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_m \end{bmatrix}; X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1m} \\ 1 & x_{21} & x_{22} & \dots & x_{2m} \\ 1 & x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}; Y = \begin{bmatrix} y_1 \\ y_2 \\ y_n \end{bmatrix}.$$

4.5 Linear Discriminant Analysis

LDA is a classification algorithm that perceives the best linear incorporation of distinctive attributes that differentiate two or more classes [40]. It initiates coordinates that provide maximum interspace between a classification center point and maximum overall variation in all classes. These two measures are defined by Eqs. (24) and (25), respectively.

Between classification scatter matrixes, the technique initiates by extrapolating input dataset into the topology, which provides the best classification between classes.

$$Y = L^T \times X, \quad (19)$$

where L represents the linear transformation matrix, X and Y as input and transformed dataset, respectively. The

two measurements mean that vectors between categorized classes are determined using (20) and (21).

$$V_1 = W^T \times V_1. \quad (20)$$

$$V_2 = W^T \times V_1, \quad (21)$$

where V_1 and V_2 represents mean class 1 and 2, respectively.

The class covariance is given by Eqs. (22) and (23).

$$\sum_1 = W^T \sum_1 W, \quad (22)$$

$$\sum_2 = W^T \sum_2 W, \quad (23)$$

between class scatter matrix (SB) is determined by Eq. (24).

$$SB = W^T (V_0 - V_1)(V_0 - V_1)^T W. \quad (24)$$

And scatter matrix within class (SW) is given by Eq. (25).

$$SW = \sum_0 + \sum_2 \dots \quad (25)$$

Finally, this algorithm pursues to increase the ratio of determinant of SB and SW.

$$\text{Ratio of Det} = \frac{\det|SB|}{\det|SW|}. \quad (26)$$

4.6 Bagged Tree (BT)

Bagged Tree is one of the ensemble family techniques. It is the cascaded of the bagging technique and the DT [41]. This algorithm is proposed to enhance the firmness and the precision in regression and classification problems. It uses a bootstrap sampling algorithm for prediction. It eliminates overfitting, minimizes variance and ultimately revamps prediction procedure. It takes the average of bagging samples (m) and determines the best referendum to obtain the gross prediction using Eq. (27) [42].

$$Y(x) = C \arg j \max \sum_{i=1}^N f_i^j(x), \quad (27)$$

where f_i^j represents the output, N repeating iterations for the classifier.

Where the inputs training is represented by Eq. (28).

$$\text{Training dataset } S\{(X1, V1), (X2, V2) \dots (Xm, Vm)\} \quad (28)$$

Bagged tree algorithm $\emptyset$;

Iteration:

For $t = 1, 2, 3 \dots \dots, N$, $ft = \emptyset(D, Dbs)$

where X and V denotes the input (predictors) and $Y(x)$ as the response for the bootstrap algorithm.

4.7 Naïve Bayes Classifier (NBC)

The Naïve Bayes Classifier provides effortless perspective, with intelligible denotation, and utilizes probabilistic understanding. It is based on Bayes' rule with powerful individualistic supposition [43]. The NBC method is a well-known algorithm because of its simplicity accuracy and fastness. It was meant to make the computations easier and less complicated; hence the name "naïve". It is based on Eq. (29).

$$P(M|N) = \frac{P(N|M) * P(M)}{P(N)}, \quad (29)$$

where M and N are the events and $P(N) \neq 0$. Where $P(M)$ and $P(N)$ are the probabilistic of perceiving M and N without considering each other. $P(M|N)$ denotes the probabilistic condition, which perceives the probabilistic event M provided N is correct. $P(N|M)$ is the probabilistic of perceiving event N provided M is accurate enough.

4.8 Recurrent Neural Network (RNN)

The Recurrent Neural Network (RNN) considers the prior output state as input. It studies the relationship between the input and output dataset [44]. RNN can reserve the prior data in its central memory. Figure 5 shows the RNN typical architecture, where hidden layers get feedback from the preceding neurons multiplied by the weight [45].

4.9 Support Vector Machine (SVM)

The Support Vector Machine (SVM) is a supervised ML algorithm. It is a ML technique that scrutinizes data for classification and regression [46]. It uses a historical input data set (called predictors) and the prior output data set (called responses) to predict the responses of the new input data. Equation (30) defines the classification mathematically.

$$Y_{fit} = \text{trainedClassifier.PredictFcn}(X), \quad (30)$$

where X and Y_{fit} represents the prediction and response respectively [47].

SVM is utilized to work out two distinct classes by locating the best Hyperplane. A Hyperplane differentiates the positive from the negative data set with maximum distance.

The maximum distance is called Margin, where positive data is given by $w.X^+ + b \geq 1$ whilst the negative data is given by $w.X^- - b \leq -1$, where w represents the hyperplane normal vector and x is the input vector. The separating Hyperplane (u) is equal to zero. The points next to the separating Hyperplane are $u = \mp 1$. Therefore, Margin (M) is represented by Eq. (31) [48, 49].

$$M = \frac{(X^+ - X^-).w}{|w|}. \quad (31)$$

Maximizing Margin is given by Eq. (32).

$$\text{Maximize } m = \frac{2}{|w|}. \quad (31)$$

Subject to $Y_i(w.x + b) \geq 1, \forall i, \dots \dots$

5 System setup

The Soltech 1STH-215-P PV panel is used in this research. The MPPT controllers use existing Johannesburg South African weather data obtained from the PSIM software. This compartment presents the three case studies conducted, as well as the experimental setup. The simulated experiments were conducted to ascertain the performance and the preciseness of the MPPT algorithms under PSC and unvarying conditions. The nine proposed ML-based MPPT algorithm were modelled and simulated by using MATLAB SIMULINK. Table 2 below illustrate the specifications of the Soltech 1STH-215-P PV panel and Boost DC-DC converter used in the experiments.

The Boost Converter is employed to balance the source and load impedance. The duty ratio is utilized to manage the flow of power by altering the duty cycle ON/OFF. The duty ratio is the proportion of time during which an electronic device is operated and can be expressed as a ratio or percentage. It is enlarged by a minor factor to reach the desired MPP. The GMPP is reached when the duty ratio approaches close to one. The PV panel output voltage is fed to the DC/DC

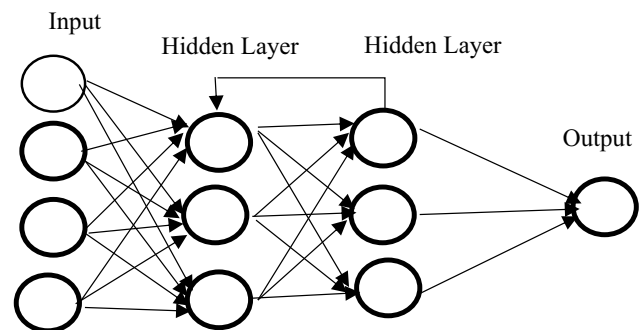


Fig. 5 Recurrent Neural Network architecture

converter. The converter produces continuous output voltage to ensure effortless energy-saving in the battery on the load compartment. As a result, the acquired accumulated energy in an inductor is discharged to the load at a high-rise voltage. In this design, the ripple current is selected to be twenty percent (20%) of the output current and the voltage ripple to be one-point five percent (1.5%), respectively. The duty cycle D is obtained by using Eq. (33), where the average and input voltages of the converter is represented by V_{out} and V_i [50].

$$D = 1 - \frac{V_i}{V_{out}}. \quad (33)$$

Equations (34) and (35) are used to determine the appropriate inductors and capacitor value, where f represents the switching frequency [51].

$$C = \frac{D}{R \left(\frac{\Delta V_o}{V_o} \right) f}, \quad (34)$$

$$L = \frac{D(1-D)^2 R}{2f}. \quad (35)$$

Figure 6 above demonstrates the complete block diagram of the proposed network, which comprises the PV Panel, MPPT techniques, the PID controller, the PWM generator, the Boost DC-DC converter and the 20 Ohms load resistor. The PV efficiency and efficiency at the load is then calculated by using Eqs. (36) and (38). Finally, the Boost converter losses are then calculated by using Eq. (38).

$$PV \text{ efficiency at MPPT} = \frac{\int_0^t P_{PV(max)t} dt}{\int_0^t P_{PV(mppt)t} dt}, \quad (36)$$

Table 2 Specification of 1STH-215-P PV panel and Boost DC/DC converter

PV Panel specification		Boost DC/DC converter specification	
PV model	1STH-215-P	L1	4 mH
Short circuit current (I_{sc})	7.84 A	C1	100 μ F
Open circuit voltage (V_{oc})	36.3 V	C2	100 μ F
Maximum Voltage (V_{mpp})	29 V	Ro	20 Ω
Maximum current (I_{mpp})	7.35 A		
Maximum power (P_{mpp})	213.15 W		
Number of cells in series (Ns)	60		
Temperature coefficient of I_{sc}	$-0.36099\%/^{\circ}\text{C}$		
Temperature coefficient V_{oc}	$0.102\%/^{\circ}\text{C}$		
Diode ideality factor (A)	0.98117		
Series resistance (R_s)	0.39383 Ω		
Shunt resistance (R_{sh})	313.3991 Ω		

$$\text{Efficiency at MPPT} = \frac{\int_0^t P_{out(mppt)t} dt}{\int_0^t P_{PV(mppt)t} dt}, \quad (37)$$

$$\text{Boost losses} = \text{Input power} - \text{Output power}, \quad (38)$$

where $P_{pv(max)}$ is defined as PV power harvested from the PV module and where $P_{pv(mppt)}$ defines the 1STH-215-P rated power at standard test condition (STC), which is equal to 213.15 W. $P_{out(mppt)}$ that defines the maximum power on the load compartment.

If the PV module is under STC with uniform solar irradiance, the values of solar irradiance $G_1, G_2 \dots G_n = 1000 \text{ W/m}^2$. Under uniform solar irradiance, the mismatch ($G_1 - G_2$) value is approximately equal to zero based on the slope of I-V curve. If the mismatch is greater than 40 W/m^2 , PSC prevails. This is represented by the equations below:

$$|G_1 - G_2| < \frac{40 \text{ W}}{\text{m}^2} \text{ Unchanging Solar irradiance}, \quad (39)$$

$$|G_1 - G_2| > \frac{40 \text{ W}}{\text{m}^2} \text{ PSC}. \quad (40)$$

Table 3 below demonstrates the three (3) case studies conducted in this research. Case study A is a standard test condition, which explains that the PV system is obtaining maximum irradiance ($G_1 = G_2 = 1000 \text{ W/m}^2$) from the sun, under normal operating temperatures ($T_1 = T_2 = 25^{\circ}\text{C}$). The normal operating temperature of the PV module is 298 K (approximately 25°C). Case study B demonstrates partial shading conditions, where the irradiance is reduced from $G_1 = 1000 \text{ W/m}^2$ to $G_2 = 850 \text{ W/m}^2$ and the ambient temperature is decreased from 25 to 20°C . The mismatch ($G_1 - G_2$) is equal to 150 W/m^2 (PSC). Case study C illustrates a very low solar insolation level, which changed from $G_1 = 1000 \text{ W/m}^2$ to $G_2 = 500 \text{ W/m}^2$ and a very high temperature of

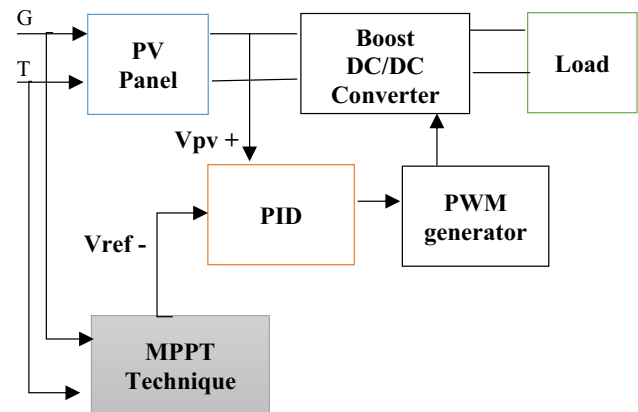


Fig. 6 Block diagram of the proposed system

T2 = 35 °C. Case study C mismatch (G1–G2) is equal to 50 $0W/m^2$ (PSC).

Figure 7 above demonstrates the MPPT algorithm training block diagram. The 130 data sample was utilized for training, testing and validation of the nine ML-based MPPT techniques. Below is a summary of steps taken when training the proposed ML-based techniques:

- 130 quality datasets are collected (Data collection).
- Start the proposed ML algorithm application.
- Load the input dataset into an application for training using pre-collected data.
- The data is branched into training and testing procedure, where 70% of the data was used for training and 30% for testing purposes (data preparation).
- The solar insolation level (G) and ambient temperature (T) (predictors) were utilized as the input parameters, and the PV voltage as the output parameter (Yfit).
- Training of an MPPT-based ML algorithm begins.
- The fitness function $Y_{fit} = \text{trainedClassifier.PredictFcn}(X)$ is then generated by using the predictors, where $\text{OutputresponseVref} = Y_{fit}$.
- The comparison of the operating PV voltage and predicted response V_{ref} takes place to attain the error signal ($\text{error} = V_{pv} - V_{ref}$).
- Evaluation of the model takes place by using types of errors.
- Make predictions by utilizing the test set for a better approximation of model performances.
- The optimal duty cycle (D) is obtained by processing the error signal to the PID controller block for tuning.
- The error signal is sent to PWM at 10 k Hz, to activate the Mosfet of the Boost DC/DC converter.

The performance of the Machine Learning algorithms is measured by using the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the Mean Squared Error (MSE) and R-squared (R^2) which are determined by using Eqs. (41)–(44).

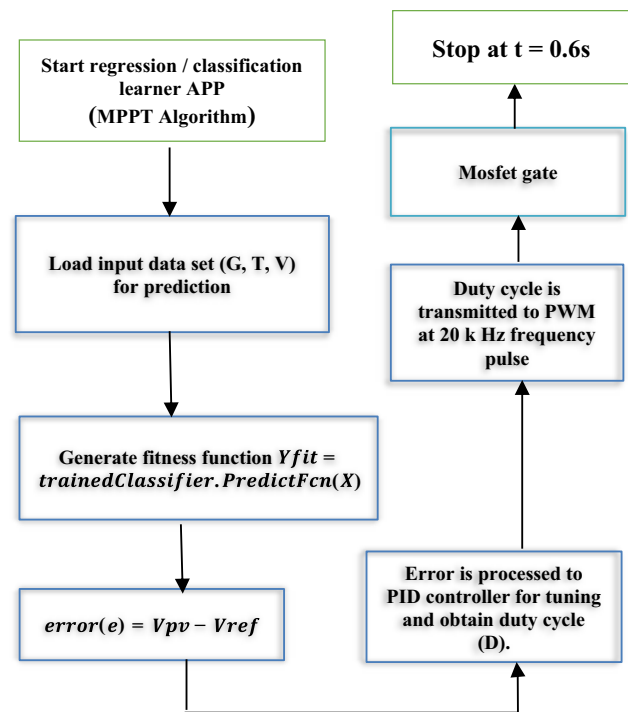


Fig. 7 MPPT algorithm training block diagram

$$MSE = \frac{1}{N} \sum_{i=1}^N (y - t)^2, \quad (41)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y - t)^2} = \sqrt{MSE}, \quad (42)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |(y - t)|, \quad (43)$$

$$R^2 = 1 - \frac{MSE(model)}{MSE(baseline)}. \quad (44)$$

6 Experimental Results

As mentioned on the system setup, the performance of the ML-based techniques is measured by using four types of errors. Table 4 below shows the training error of the nine (9) Machine Learning techniques, using an identical percentage for training, testing and validation in a tabulated form.

Table 4 above the SVM and WK-NN MPPT algorithms performance is the finest, while the Bagged Tree (BT) recorded the worst in terms of training, testing and validation error. SVM recorded RSME of 0.14, R-squared error of 0.99 in 1.1178 s (shorter period than any other algorithm).

Table 3 Case studies conducted

Weather Conditions	Irradiance (W/m^2)	Temperature (°C)
Case Study A – STC	G1 = 1000	T1 = 25
	G2 = 1000	T2 = 25
Case Study B – PSC	G1 = 1000	T1 = 25
	G2 = 850	T2 = 20
Case Study C – Worse PSC	G1 = 1000	T1 = 25
	G2 = 500	T2 = 35

Table 4 MPPT testing error tabulated results

Error type	DT	WK-NN	MLR	LDA	BT	GPR	NBC	SVM	RNN
RMSE	0.42	0.37	0.44	0.48	0.73	0.4	0.51	0.14	0.36
R^2	0.96	0.98	0.96	0.96	0.91	0.97	0.95	0.99	0.98
MSE	0.18	0.14	0.2	0.21	0.53	0.16	0.26	0.06	0.19
MAE	0.20	0.23	0.3	0.29	0.52	0.18	0.26	0.19	0.24
Prediction speed (obs/sec)	5600	8700	830	3500	1900	3800	1600	7200	4400
Training time (s)	0.91	0.78	6.17	2.32	2.35	5.04	8.56	1.1178	8.9

WK-NN recorded an RSME of 0.37, R-squared error of 0.98 in 0.78 s.

RNN, GPR, MLR, DT performed slightly higher with RMSE of 0.36, 0.4, 0.44 and 0.42 independently, and R-squared of 0.98, 0.97, 0.96 and 0.96 respectively. NBC and LDA also performed better than BT recording RMSE of 0.51 and 0.48 and R-squared of 0.95 and 0.96 independently, whilst BT performed worst at RMSE of 0.73 and R-squared of 0.91. However, the Naïve Bayes classifier has more training time than any other algorithms used in this research study.

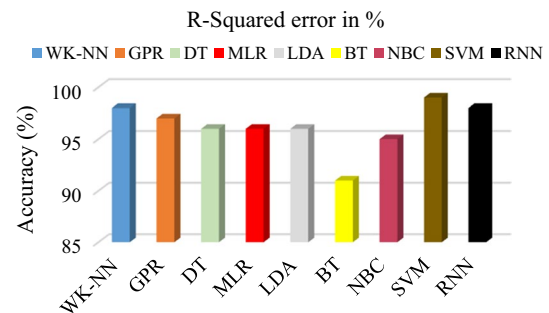
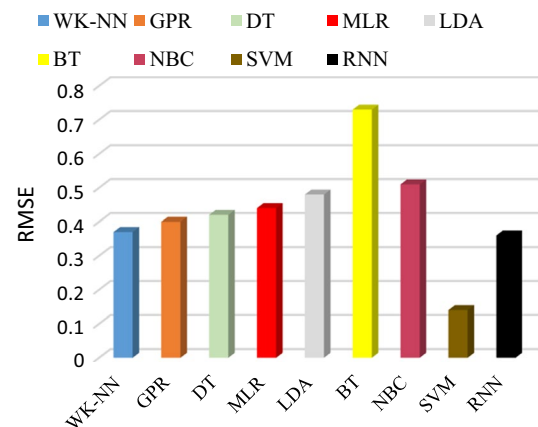
Figure 8 below demonstrates the graphical results of R-squared error obtained from MATLAB. The results are in percentages up to a hundred (100%). Figures 9 and 10 below also show the graphical representation of RMSE and training time of the MPPT techniques gotten from the MATLAB, respectively.

6.1 Case Study A

Case Study A (STC): As mentioned above, it represents full sun. Table 5 below and Figs. 11, 12 and 13 below illustrate the tabulated and graphical results received from this experiment, using nine (9) different Machine-Learning algorithms.

From the Case Study A results, SVM performed better than any other method with an operating efficiency of 99.77% and 96.76% at the PV panel and the load, respectively. RNN became the second highest with efficiency of 98.58% and 96.67% at the output and the load individually. WK-NN, GPR, DT, MLR, LDA and NBC performed slightly lower than SVM and RNN with a PV efficiency of 209.1%, 97.8%, 97.7%, 97.6%, 97.23% and 97.1% distinctly. BT performed worse than any other method with the PV power of 206.2 W and efficiency of 96.8

Figures 14 and 15 show the simulated P–V curve and I–V curve for Case Study A respectively. The nine ML MPPT techniques are compared with Conventional P&O under PSC as shown in Fig. 14. The conventional P&O performed poorly when compared to the ML algorithms under STC. It obtained a PV power of 190 W and PV efficiency of 89% respectively. The I–V curve demonstrates that the conventional MPPT algorithm struggles to optimize the maximum current under Standard test conditions as shown in Fig. 15.

**Fig. 8** R-squared error results in percentage**Fig. 9** RMSE graphical results

6.2 Case Study B

Case Study B (PSC, as mentioned above represents a sudden change in weather conditions. Table 6 and Figs. 16, 17 and 18 below demonstrates the tabulated and graphical results extracted from this experiment, by using nine (9) different MPPT techniques.

For Case Study B, SVM continued to outperform the six MPPT algorithms with the PV power of 206.1 W under sudden change of environment and recorded the remarkable 96.76% efficiency on the PV output. WK-NN has a faster tracking speed than any other MPPT algorithm. It managed

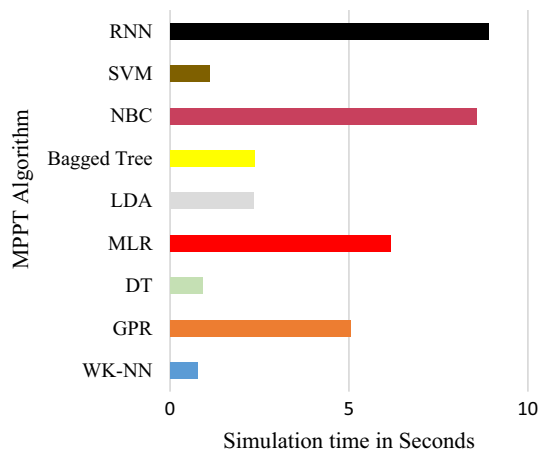


Fig. 10 MPPT Training time results

to reach and settle at MPP in 0.02 s. WKNN performed slightly lower than SVM with PV efficiency of 95.12% and PV power of 202.6 W, respectively. SVM and WK-NN have less oscillation around MPP, as shown in Fig. 15 below.

DT, MLR, LDA and NBC showed competency by producing PV efficiency of 92.96%, 92.56%, 92.45% and 91.46% under PSC independently. BT underperformed once again, with the PV power of 194 W and PV efficiency of 91.01%. With Case Study B, it can be concluded that RNN performs poorly under PSC when compared to other Machine Learning techniques. RNN recorded PV efficiency of 88.59% and 86.24% efficiency at the load end.

Figures 19 and 20 demonstrate the P–V and I–V curve simulated results from PSC B. The conventional P&O is compared to the nine proposed ML MPPT algorithms under light PSC. Under PSC, the PV current of P&O dropped, which ultimately reduces the PV power as shown in Fig. 19. The P&O recorded the PV power of 180 W and PV efficiency of 84.4% respectively. The conventional MPPT fails to distinguish between the GMPP and LMPP under different weather conditions. It ends up trapped in LMPP and

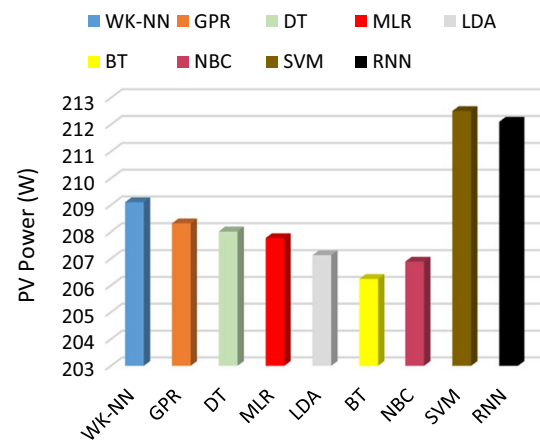


Fig. 11 MPPT techniques PV Power at STC

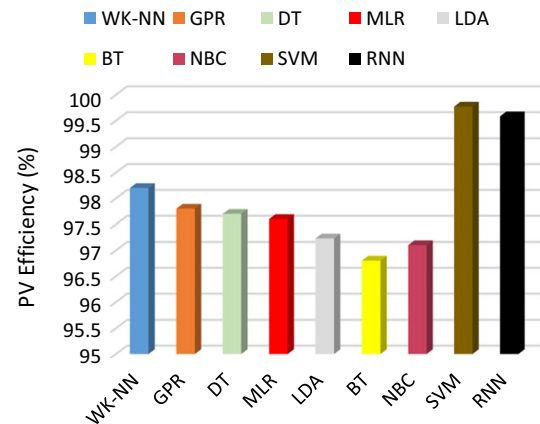


Fig. 12 MPPT techniques PV efficiency at STC

thus loss of power. The ML techniques managed to locate GMPP on a P–V curve and ultimately eliminated unnecessary power losses.

Table 5 Case Study A MPPT experimental results

Measurement	DT	WK-NN	MLR	LDA	BT	GPR	NBC	SVM	RNN
I _{pv} (A)	7.2	7.21	7.18	7.1	7.1	7.2	7.2	7.2	7.185
V _{pv} (V)	29	29.1	28.92	28.8	28	29	28.7	29.43	29.52
P _{pv} (w)	208	209.1	207.8	207.1	206.2	208.3	206.9	212.5	212.1
PV η (%)	97.7	98.2	97.6	97.23	96.8	97.8	97.1	99.77	99.58
I _o (A)	– 3.2	– 3.2	– 3.16	– 3.16	– 3.17	– 3.17	– 3.16	– 3.21	– 3.2
V _o (V)	– 63	– 63.4	– 63.67	– 63.7	– 63	– 63.7	– 63.6	– 64.21	– 64.34
P _{out} (W)	201.6	203	201.5	201.12	200.3	201.9	201.0	206.1	205.9
P _{out} η (%)	94.6	95.1	94.6	94.4	94	94.8	94.4	96.76	96.67
Losses (W)	6.4	6	6.3	5.98	5.9	6.4	5.89	6.4	6.2

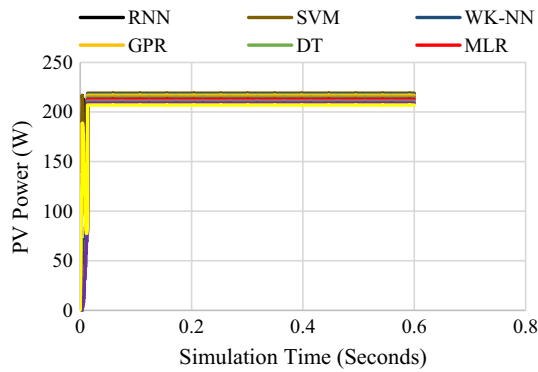


Fig. 13 Case study A graphical results

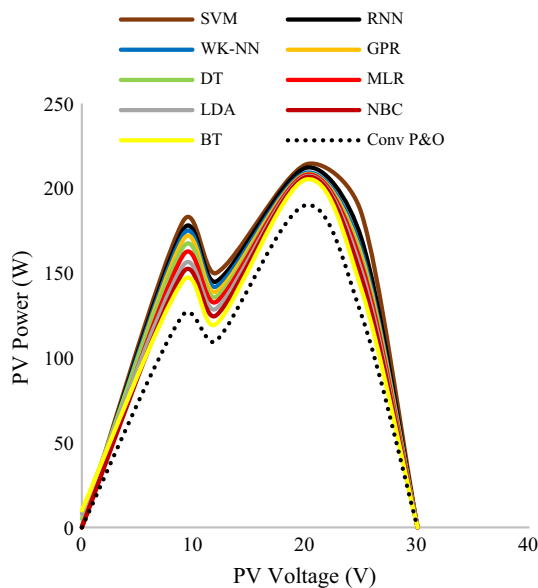


Fig. 14 Case study A P-V Curve

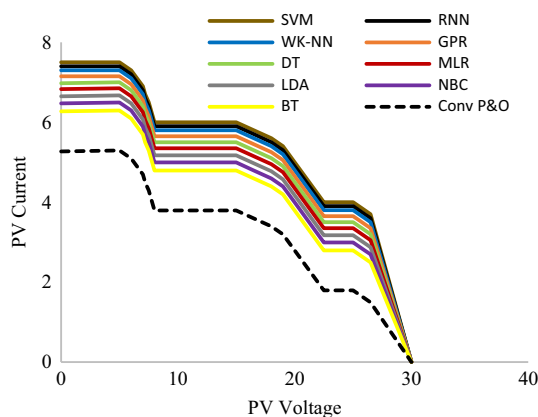


Fig. 15 Case study A I-V Curve

6.3 Case Study C

Case Study C (worst PSC, as mentioned above, represents a drastic change in weather conditions). Table 7 and from Figs. 21, 22 and 23 demonstrate the tabulated and graphical results extracted from this experiment, using nine (9) different MPPT algorithms.

From Case Study C, it can be noticed that SVM outperformed all the other methods under worse PSC. SVM took the lead with PV power of 178.19 W and an operating efficiency of 83.66. GPR became the second recording the PV power of 176.7 W. GPR tracked the MPP at 0.029 s, which makes it the fastest MPPT method. WK-NN became the third highest performed method this time around; WK-NN works slightly better under PSC and performs much better with a slight change in environment factors. It recorded a PV power of 173.6 W and efficiency of 81.5%, respectively. In this case, WK-NN suffered from larger oscillations than any other method. DT, MLR, LDA and NBC attained a PV efficiency of 79.34%, 78.98%, 78.83% and 78.07% independently. BT underperformed by obtaining the PV power of 165 W and 77.48% PV efficiency. RNN underperformed under mismatch when compared to other ML-based MPPT techniques by extracting PV power of 161.3 W and 75.77% PV efficiency.

Figures 24 and 25 represents the P-V curve and I-V curve for nine ML MPPT algorithm and conventional P&O under heavy PSC. As shown in Fig. 25, P&O shows multiple power peaks, which result from the different weather conditions. The solar irradiance is reduced, which decreases the PV current and PV power as shown on the I-V and P-V curve. The conventional method dismally fails to track the true GMPP under heavy PSC. It recorded a PV power of 154.6 W and PV efficiency of 72.5% separately. The ML algorithms performed significantly better than the conventional method because they use historic data, which has high accuracy, particularly SVM, RNN and WK-NN.

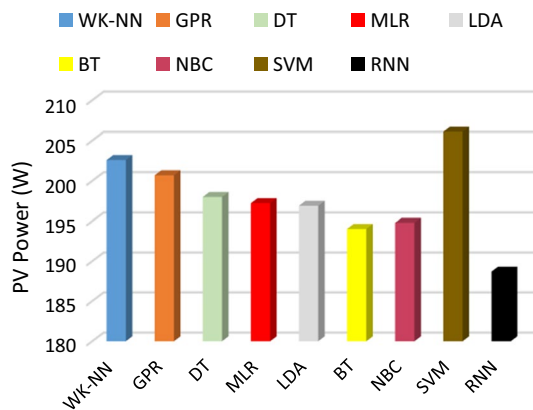
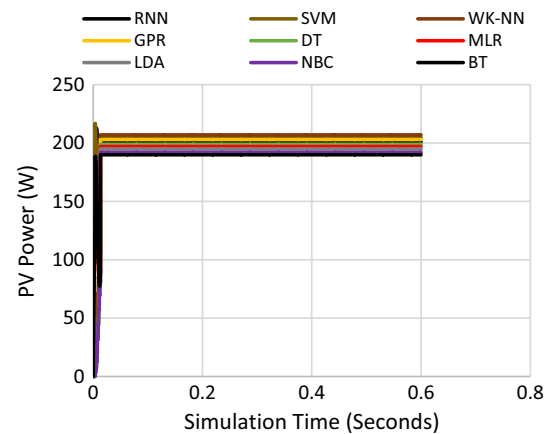
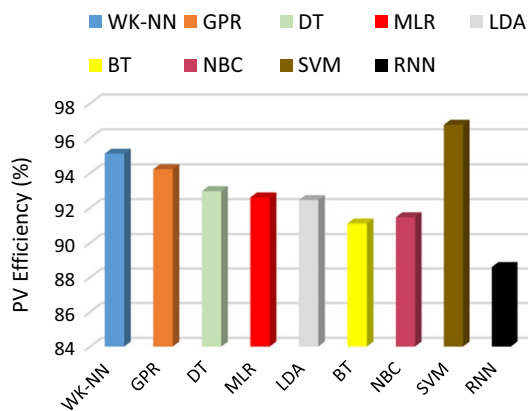
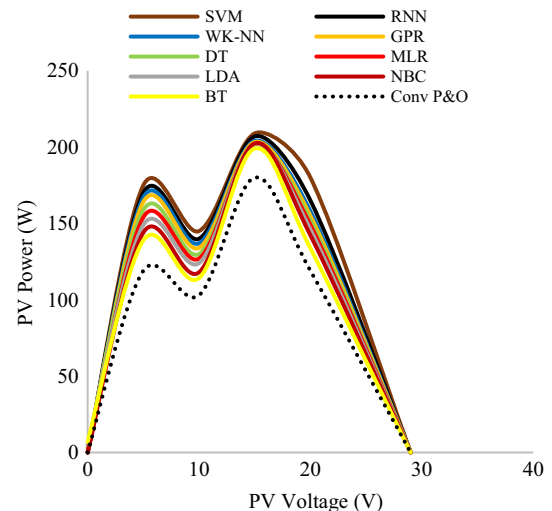
7 Conclusion

This paper formally introduced nine (9) ML techniques to harvest MPP out of PV panel in case of no sensor. The algorithms presented also have a fast-tracking speed and minimized oscillations around MPP, using a PID controller for further fine-tuning of errors.

The nine advanced classification and regression ML algorithms were constructed and tested under distinct weather conditions, using Johannesburg weather conditions. The ML techniques performance was evaluated by using MSE, RMSE, MAE and R-squared error.

Table 6 Case Study B MPPT experimental results

Measurement	DT	WK-NN	MLR	LDA	BT	GPR	NBC	SVM	RNN
I_{pv} (A)	6.97	7	7	6.88	6.87	6.98	6.88	7.12	6.449
V_{pv} (V)	28.41	29	28.14	28.58	28.24	28.8	28.4	28.95	29.27
P_{pv} (w)	198	202.6	197.2	196.9	194	200.7	194.8	206.1	188.7
PV η (%)	92.96	95.12	92.56	92.45	91.01	94.23	91.46	96.76	88.59
I_o (A)	- 3.14	- 3.16	- 3.14	- 3.13	- 3.09	- 3.15	- 3.13	- 3.17	- 3.07
V_o (V)	- 61	- 62.0	- 60.9	- 61	- 61	- 61.8	- 60.5	- 63.08	- 61
P_{out} (W)	191.6	196.03	191.3	190.95	188.5	194.6	189.21	199.97	187.3
P_{out} η (%)	90	92	90	89.6	88.5	91.4	88.8	93.88	86.24
Losses (W)	6.1	5.97	5.9	5.95	5.5	6.1	5.59	6.13	5.0

**Fig. 16** MPPT techniques PV Power at under PSC**Fig. 18** Case study B graphical results**Fig. 17** MPPT techniques PV efficiency under PSC**Fig. 19** Case study B P-V Curve

From the experimental results, SVM and WK-NN method showed more accuracy and a quicker computational time than any other methods. BT and RNN had less precision

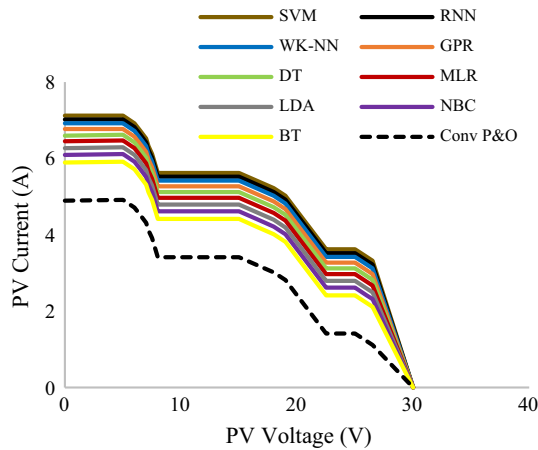


Fig. 20 Case study B I-V Curve

than any other techniques. The Naïve Bayes Classifier has a high computational time but is complicated and expensive to construct.

The experimental results attained suggested that SVM, WK-NN and GPR could be utilized to harvest maximum power from a PV module. The results showed that SVM and WK-NN works better when the solar irradiance is high and under diverse environmental conditions. However, under the worst PSC, GPR performed better than WK-NN. It is clear that with these methods, a PV system can be used as a primary power source, not only because it is a cleaner and cheaper source of energy, but because these methods can optimize maximum power under PSC. The proposed

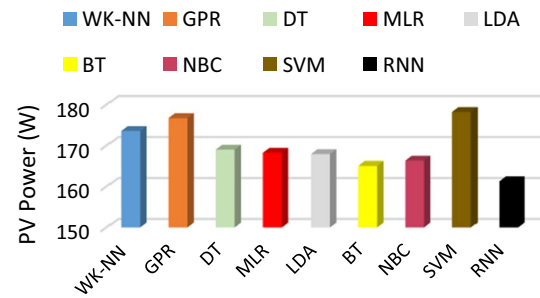


Fig. 21 MPPT algorithms PV power under worst PSC

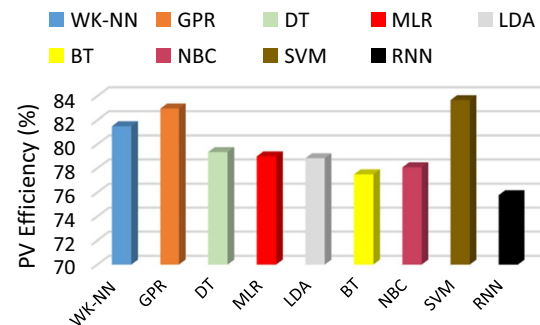


Fig. 22 MPPT algorithms PV efficiency under worst PSC

methods of cascading the MPPT and PID controllers are significantly better because training errors are as minimized as possible.

Table 7 Case Study B MPPT experimental results

Measurement	DT	WK-NN	MLR	LDA	BT	GPR	NBC	SVM	RNN
I_{pv} (A)	6.27	6.3	6.25	6.24	6.17	6.39	6.42	6.44	5.474
V_{pv} (V)	26.97	27.5	26.92	26.89	26.77	27.67	25.89	27.68	29.48
P_{pv} (w)	169	173.6	168.2	167.9	165	176.7	166.3	178.19	161.3
PV η (%)	79.34	81.5	78.98	78.83	77.48	82.96	78.07	83.66	75.77
I_o (A)	- 2.72	- 2.73	- 2.71	- 2.69	- 2.7	- 2.73	- 2.7	- 2.7	- 2.53
V_o (V)	- 62.1	- 63.6	- 62.1	- 62.4	- 61.13	- 64.72	- 61.57	- 64.75	- 61.0
P_{out} (W)	162	166.5	161.34	161.0	158.17	169.7	159.4	171.99	154.48
P_{out} η (%)	76	78.17	75.75	75.61	74.3	79.67	74.84	80.75	72.52
Losses (W)	6.9	7.1	6.89	6.87	6.87	7	6.88	6.2	6.82

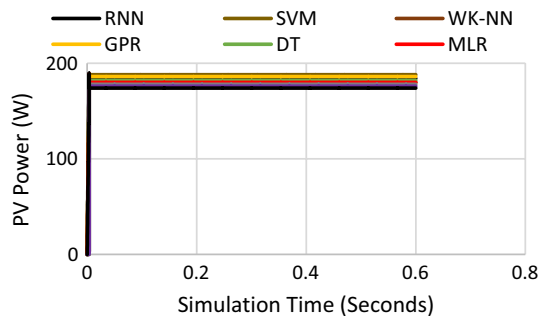


Fig. 23 Case study C graphical results

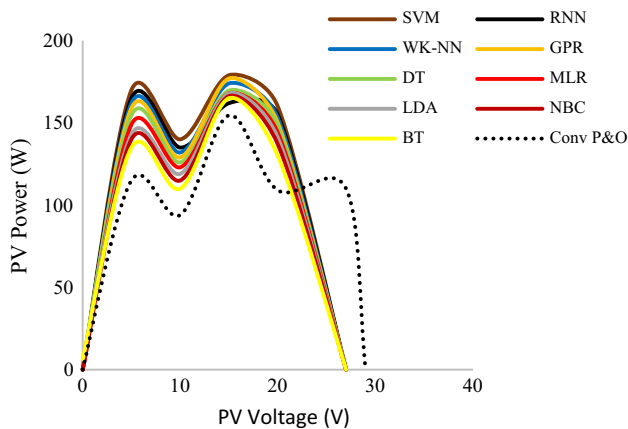


Fig. 24 Case study C P-V curve

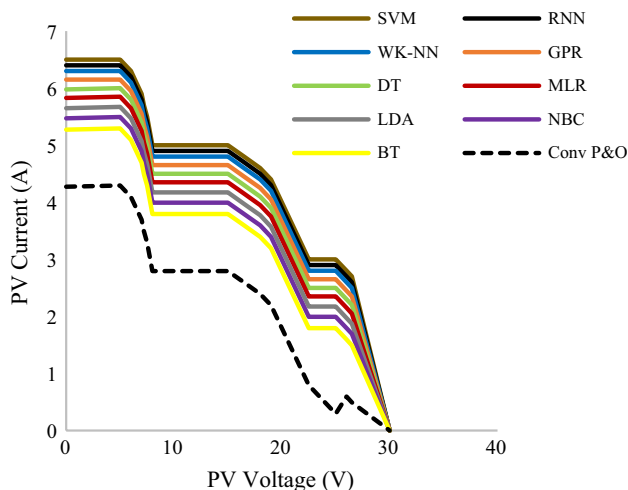


Fig. 25 Case study C I-V curve

Funding This research was supported by the Energy Cloud R&D Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT (2019M3F2A1073164).

References

1. Lyden S, Haque ME (2016) A simulated annealing global maximum power point tracking approach for PV modules under partial shading conditions. *IEEE Trans Power Electron* 31(6)
2. Radjai T, Rahmani L, Mekhilef S, Gaubert JP (2014) Implementation of a modified incremental conductance MPPT algorithm with direct control based on a fuzzy duty cycle change estimator using dSPACE Science Direct. *Elsvier Solar Energy* 110:325–337
3. Mpho SN, Ali NH, AA (2019) MPPT under partial shading conditions based on perturb & observe and incremental conductance. In: 2019 11th ELECO international conference
4. Geem ZW, Yoon Y (2017) Harmony search optimization of renewable energy charging with energy storage system. *Int J Electr Power Energy Syst* 86:120–126
5. Shiqing T, Yize S, Yujie C, Yiman Z, Yunhu Y, Warren So (2017) An enhanced MPPT method combining fractional order and fuzzy logic control. *IEEE J Photovolt* 2156–3381
6. Pakkiraiah B, Sukumar GD (2016) Research survey on various MPPT performance issues to improve the solar PV system efficiency. *J Solar Energy* 2016:1–20
7. Farayola AM, Hasan AN, Ali A, Twala B (2018) Distributive MPPT approach using ANFIS and perturb&observe techniques under uniform and partial shading conditions. *Artif Intell Evol Comput Eng Syst Adv Intell Syst Comput* 668:27–37 (Springer, Singapore 2018)
8. Bounechba H, Bouzid A, Nabti K, Benalla H (2014) Comparison of perturb & observe and fuzzy logic in maximum power point tracker for PV systems Science Direct, Elsevier. *Energy Procedia* 50:677–684
9. Saravanan S, Babu NR (2016) Maximum power point tracking algorithms for photovoltaic system—a review. *Renew Sustain Energy Rev* 57:192–204
10. Fraunhofer Institute for Solar Energy Systems (2016) ISE, “PHOTOVOLTAICS REPORT,” Fraunhofer ISE. Freiburg, Germany
11. Prasanth RJ, Sudhakar BT (2017) A comprehensive review on solar PV maximum power point tracking techniques. *Renew Sustain Energy* 67(3):826–847 (Elsevier)
12. Guerard JB (2013) Introduction to Financial Forecasting in Investment Analysis. Springer Science+Business Media 1, 19–45
13. de Medeiros T, Antunes FLM, dos Reis FS (1998) An artificial neural network-based real time maximum power tracking controller for connecting a PV system to the grid. In: *Proceeding of IEEE the 24th annual conference on industrial electronics society*, 1, 554–558
14. Adedayo MF, Ali NH, Ahmed A (2018) Efficient photovoltaic MPPT system using coarse gaussian support vector machine and artificial neural network techniques. *Int J Innov Comput Inf Control (IJICIC)* 14(1)
15. Farayola AM, Hasan AN, Ali A (2017) Curve fitting polynomial technique compared to ANFIS technique for maximum power point tracking. In: 8th IEEE International Renewable Energy Congress (IREC) 2017, Amman Jordan
16. Mirhassani SM, Golroodbari SZM, Golroodbari SMM, Mekhilef S (2015) An improved particle swarm optimization based maximum power point tracking strategy with variable sampling time. *Int J Electr Power Energy Syst* 64:761–770. <https://www.sciencedirect.com/science/article/pii/S0142061514005183>
17. Kermadi M, Berkouk E (2015) A maximum power point tracker abased on particle swarm optimization for pv-battery energy

- system under partial Shading conditions. In: Control, engineering information technology (CEIT), 2015 3rd international conference on, pp 1–6
18. Yu Y, Tian YF, Yin ZF (2005) Multiuser detector based on adaptive artificial fish school algorithm. In: Proceeding of IEEE international symposium on communications and information technology, Beijing, China, 1480–1484
 19. Murtaza MC, De Giuseppe M, Boero D (2014) Adaptive cycle optimization-based hybrid maximum power point tracking technique for photovoltaic systems. *Int J. Electr Power Energy Syst* 59:141–154
 20. Geem ZW, Kim J-H (2018) Application of computational intelligence techniques to an environmental flow formula. *Int J Fuzzy Logic Intell Syst* 18(4):237–244
 21. Shaiek MB, Smida AS, Mimouni MF (2013) Comparison between conventional methods and {GA} approach for maximum power point tracking of shaded solar{PV}generators. *Solar Energy* 90:107122. <https://www.sciencedirect.com/science/article/pii/S0038092X13000091>
 22. Tajuddin MFN, Ayob SM, Salam Z, Saad MS (2013) Evolutionary based maximum power point tracking technique using differential evolution algorithm. *Energy Build* 67:245252. <https://www.sciencedirect.com/science/article/pii/S0378778813004763>
 23. Geem ZW (2012) Size optimization for a hybrid photovoltaic-wind energy system. *Int J Electr Power Energy Syst* 42(1):448–451
 24. Morrison JW, Zapata SK, Perez MA, Meynard TA, Renaudineau H (2016) Partial power dc-dc converter for photovoltaic two stage string inverters. In: 2016 IEEE energy conversion congress and exposition (ECCE)
 25. Zhang F, Thanapalan K, Procter A, Carr S, Maddy J (2013) Adaptive hybrid maximum power point tracking method for a photovoltaic system. *Energy Convers IEEE Trans* 28(2):353–360
 26. Adedayo F, Ali NH, Ahmed A (2017) Implementation of modified incremental conductance and fuzzy logic MPPT techniques using MCUK converter under various environmental conditions. *Appl Solar Energy Springer* 53(2)
 27. Ahmed J, Salam Z (2015) An improved method to predict the position of maximum power point during partial shading for PV arrays. *Ind Inf IEEE Trans* 11(6):1378–1387
 28. Tey KS, Mekhilef S (2014) Modified incremental conductance algorithm for photovoltaic system under partial shading conditions and load variation. *IEEE Trans Industr Electron* 61(10):5384–5392
 29. Liu Y-H, Chen J-H, Huang J-W (2015) A review of maximum power point tracking techniques for use in partially shaded conditions. *Renew Sustain Energy Rev* 41:436–453
 30. Ahmed J, Salam Z (2017) An accurate method for MPPT to detect the partial shading occurrence in a PV system. *IEEE Trans Industr Inf* 13(5):2151–2161
 31. Adedayo MF, Ali NH, Ahmed A (2018) Optimization of PV systems using data mining and regression learner MPPT techniques TELKOMNIKA Indonesian. *J Electr Eng* 10(3):1080–1089
 32. Villalva MG, Gazoli JR (2009) Comprehensive approach to modeling and simulation of photovoltaic arrays. *Power Electron IEEE Trans* 24(5):1198–1208
 33. Teshome D, Lee CH, Lin YW, Lian KL (2016) A modified firefly algorithm for photovoltaic maximum power point tracking control under partial shading. *IEEE J Emerg Sel Topics Power Electron*
 34. Naghibi SA, Pourghasemi HR, Dixon B GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran," *Environmental monitoring and assessment*, vol. 188,no.1, 44, Naghibi SA, Pourghasemi HR, Dixon B
 35. Agrawal R (2014) K-nearest neighbor for uncertain data (0975–8887). *Int J Comput Appl*. 105(11):13–16
 36. Parvin H, Alizadeh H, Minati MB, Bidgoli MKNN (2008) Modified K-Nearest Neighbor. In: Proceeding soft he world congress on engineering and computer science WCECS. 2008. ISBN: 978-988-98671-0-2, pp 1–4
 37. Mutrofin S, Kurniawandhani A, Izzah A, Masrur M Optimasi Teknik Klasifikasi Modified K-Nearest Neighbor Menggunakan Lappalainen K, Valkealahti S (2017) Effects of irradiance transition characteristics on the mismatch losses of different electrical PV array configurations. *Inst Eng Technol (IET) Renew Power Gen* 11(2):248–254 (**Algoritma Genetika". Jurnal GAMMA. 2016. Vol. 10 No. 1, pp. 1–5**)
 39. Ndamulelo T, Ali NH, Thokozani S (2020) Proposed fuzzy logic system for voltage regulation and power factor improvement in power systems with high infiltration of distributed generation. *Energies* 13(16):4241
 40. El Telbany ME, Youssef A, Zekry AA (2014) Intelligent techniques for MPPT control in photovoltaic systems: a comprehensive review. In: 2014th international conference on artificial intelligence with applications in engineering and technology (ICAiet 2014), Kinabalu, Malaysia
 41. Clemencon S, Marine D, Nicolas V (2009) Bagging ranking trees. In: IEEE international conference on machine learning and applications (ICMLA09), Florida, USA
 42. Sutton CD (2005) Classification and regression trees, bagging, and boosting. *Handb Stat* 24:303–329
 43. Ting SL, Ip WH, Tsang AC (2011) Is Naïve Bayes a Good Classifier for Document Classification? *Int J Softw Eng Appl* 5(3)
 44. Pascanu R, Mikolov T, Bengio Y (2013) On the difficulty of training recurrent neural networks. In: International conference on machine learning, 1310–1318
 45. Nkambule MS, Hasan AN, Ali A (2019) Proportional study of perturb & observe and fuzzy logic control MPPT algorithm for a PV system under different weather conditions. In: Computer Science; 2019 IEEE 10th GCC Conference & Exhibition (GCC)
 46. Rizzo SA, Salerno N, Scelba G, Sciacca A (2018) Enhanced hybrid global MPPT algorithm for PV systems operating under fast-changing partial shading conditions. *Int J Renew Energy Res* 8(1)
 47. Morrison JWZ, Kouro S, Perez MA, Meynard TA, Renaudineau H (2016) Partial power dc-dc converter for photovoltaic two stage string inverters. In: 2016 IEEE energy conversion congress and exposition (ECCE)
 48. Kareim AA, Bin-Mansor M (2013) Efficiency improvement of the maximum power point tracking for PV systems using support vector machine technique. In: The 4th International Conference on Energy and Environment (ICEE2013), Mashhad, Iran, Al Alin, United Arab Emirates
 49. Yoon Y, Geem ZW (2015) Parameter Optimization of Single-Diode Model of Photovoltaic Cell Using Memetic Algorithm. *Int J Photoenergy*, Article ID 963562, 2015
 50. Alsharif MH, Yahya K, Geem ZW (2020) Strategic market growth and policy recommendations for sustainable solar energy deployment in South Korea. *J Electr Eng Technol* 15(2):803–815
 51. Shin H, Geem ZW (2019) Optimal design of a residential photovoltaic renewable system in South Korea. *Appl Sci* 9(6):1138

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Mpho Sam Nkambule received his Bachelor and MPhil degrees in electrical and electronics engineering from University of Johannesburg, Johannesburg, South Africa in 2017 and 2020, respectively. Mpho is currently pursuing the Ph.D. degree with the University of Johannesburg. His research interests include solar photovoltaic systems, power electronics, artificial intelligence and MPPT controller designs for energy harvesting systems.



Ali Hasan received B. Eng. in Electrical and Electronic Engineering from the Hashemite University in Jordan in 2005. Dr. Hasan completed his M. Eng. in Electrical Engineering from North West University in 2010 and his D. Ing (PhD) degree in Electrical and Electronic Engineering in the year 2014 from University of Johannesburg in South Africa respectively. Dr. Hasan is currently working at the department of electrical engineering at the Higher Colleges of Technology in Abu Dhabi, UAE

as Assistant Professor. Dr. Hasan's research work focuses on the applications of Artificial Intelligence in energy, power and communication systems.



Ahmed Ali received his Ph.D. degree in Electrical Engineering from University of Johannesburg in 2017. He is presently working as a Senior Lecturer of Electrical Engineering from University of Johannesburg- South Africa. His current research interest includes Machine learning, Artificial Intelligence, Renewable Energy Resources and Power Electronics.



Junhee Hong received B.S., M.S., and Ph.D. degrees in electrical engineering from Seoul National University, Seoul, South Korea, in 1987, 1989, and 1995, respectively. Since 1995, he has been a Professor of the Department of Energy IT at Gachon University, Seongnam, South Korea. His research interests include smart grid, super-grid, renewable energy, digitalization of electric power, and energy policy.



Zong Woo Geem is an Associate Professor in Department of Energy IT at Gachon University, South Korea. He has obtained B.Eng. in Chung-Ang University, Ph.D. in Korea University, and M.Sc. in Johns Hopkins University, and researched at Virginia Tech, University of Maryland - College Park, and Johns Hopkins University. He invented a music-inspired optimization algorithm, Harmony Search, which has been applied to various scientific and engineering problems. His research interest

includes phenomenon-mimicking algorithms and their applications to energy, environment and water fields. He has served for various journals as an editor or guest editor (Engineering Optimization; Swarm & Evolutionary Computation; Int. Journal of Bio-Inspired Computation; Journal of Applied Mathematics; Applied Sciences; Complexity; and Sustainability).