

Article

A Method to Improve the Accuracy of Simulation Models: A Case Study on Photovoltaic System Modelling

Aekkawat Bupi ^{1,*}, Songkiate Kittisontirak ², Perawut Chinnavornrungrsee ², Sasiwimon Songtraï ³, Phassapon Manosukritkul ⁴, Kobsak Sriprapha ³, Wisut Titiroongruang ¹ and Surasak Niemcharoen ^{1,*}

¹ Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang, Chalongkrung Rd, Ladkrabang, Bangkok 10520, Thailand; wisut.ti@kmitl.ac.th

² National Energy Technology Center (ENTEC), National Science and Technology Development Agency (NSTDA), 112 Thailand Science Park, Thanon Phahonyotin, Tambon Klong Nueng, Amphoe Klong Luang, Pathum Thani 12120, Thailand; songkiate.kittisontirak@nectec.or.th (S.K.); perawut.chinnavornrungrsee@nectec.or.th (P.C.)

³ National Electronics and Computer Technology Center (NECTEC), National Science and Technology Development Agency (NSTDA), 112 Thailand Science Park, Thanon Phahonyotin, Tambon Klong Nueng, Amphoe Klong Luang, Pathum Thani 12120, Thailand; sasiwimon.songtraï@nectec.or.th (S.S.); kobsak.sriprapha@nectec.or.th (K.S.)

⁴ King Mongkut's Institute of Technology Ladkrabang Prince of Chumphon Campus, Chum Kho, Pathio District, Chumphon 86160, Thailand; phasapon.ma@kmitl.ac.th

* Correspondence: aekkawat.bupi@outlook.com (A.B.); surasak.ni@kmitl.ac.th (S.N.)

Abstract: This research presents a method to improve data accuracy for the more efficient data management of the studied applications. The data accuracy was improved using the preciseness function learning model (PFL model). It contains a database in which the amount of data is more or less dependent on all of the possible behavior of the studied application. The proposed model improves data with functions obtained by optimizing curves to represent the data at each point, which estimate the database's diffusion behavior, and functions can be built around all of the various forms of databases. The proposed model always updates its database after processing. It has been learning to optimize the processing precision. In order to verify the precision of the proposed model through its application to a PV system simulation model, the process's database should contain at least one year. This is because the overall behavior of the PV power output in Thailand depends on the seasonal weather; Thailand has three seasons in a period of one year. The testing was performed by comparing the PV power output. The simulation results with the actual measurement data (12 MW PV system) can be divided into two conditions: the daily comparison and the seasonal PV power output. As a result, the proposed model can accurately simulate the PV power output despite the sudden daily climate change. The average nRMSE (normalized RMSE) of the proposed model is very low (1.23%), and ranges from 0.30% to 2.26%. Therefore, it has been proven that this model is very accurate.

Keywords: preciseness function learning model (PFL model); learning; photovoltaic; solar irradiance; module temperature



Citation: Bupi, A.; Kittisontirak, S.; Chinnavornrungrsee, P.; Songtraï, S.; Manosukritkul, P.; Sriprapha, K.; Titiroongruang, W.; Niemcharoen, S. A Method to Improve the Accuracy of Simulation Models: A Case Study on Photovoltaic System Modelling. *Energies* **2021**, *14*, 372. <https://doi.org/10.3390/en14020372>

Received: 7 December 2020

Accepted: 4 January 2021

Published: 11 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The application of mathematical models makes process improvement more convenient because sometimes it is not able to perform experiments and modify the processes. The model helps us to obtain in-depth information to be used for the benefits of education or business. The adequacy of the model which is used relies on its accuracy in processing the results. Suppose the model has low accuracy; it will cause more disadvantages than benefits. The primary source of this research will be a recognition of the importance of the model's accuracy. At present, photovoltaic generators have been installed as a substitute for various sources of electrical energy, for applications such as home appliances, utilities,

external battery packs, and generators, etc. The most critical steps after the installation of the PV system are maintenance and checking. If we estimate the PV system's energy production, we will be available to check for system faults by comparing the measured data with the simulation data for energy management, and checking whether it exceeds its maximum benefit. Data accuracy is a critical factor for effective energy data management. As such, this research will present a PV system estimation model with a preciseness function learning model.

Many researchers have conducted studies on the creation of PV models, from one-panel simulations to systems simulations of solar cells. PV models are mainly affected by the module temperature and solar irradiance. Most of the modeling research is the analysis of single-diode equivalent circuits (five parameters model) [1–18]. These models have five parameters: photocurrent, the ideal factor according to the photovoltaic technology, parallel resistance, leakage current, and series resistance. In many studies, the data accuracy has been improved after simulating the data from the model by different methods. For example, the improvement of the accuracy of PV models by the use of weight functions obtained by 1-year measured data has been studied [1]. This method makes the model more accurate by evaluating the database using a polynomial equation. This equation is a data evaluation in one form of a curve-fitting process. Other research has shown the linear weighting method for PV power forecasting models [2]. This method updates the model data using a linear equation in order to obtain the final result. Based on several studies on the improvement of the accuracy of the database, each method [3–6] has different strengths depending on the distribution behavior of the database. Therefore, this research combines each method's strengths and develops it as a new data precision improvement method in order to make the model more efficient for the best data management.

This research aims to improve the processing accuracy of models in order to make the models more useful. The method used to improve the data accuracy is the incorporation of the strengths of the creation of a function from a database to solve problems if a high distribution of information is created. This model can be made more effective by adding output data into the database to enable the model to be self-learning, and to calculate more precisely in the next calculations. The case studies applied to this process were the power output simulation model of photovoltaic systems, and comparison of the efficiency with the other methods or processes of accuracy improvement researched earlier. The test used data from photovoltaic systems in Thailand ($14^{\circ}10'78.1''$ north latitude and $100^{\circ}16'94.9''$ east longitude). There were two inputs in this model, which are solar irradiance and module temperature, and the output of the model was the power output value of the photovoltaic system. The test was to bring the PV power output value from the accuracy improved model in the different forms, including this research process compared to the actual PV power output in three seasons: the summer, the rainy season, and the winter. The climate changes of the target area have a profound effect on the performance of photovoltaic systems, and also the timing of the installation of the system that causes energy loss in various fields. For this reason, the research devised a method to optimize the data accuracy of the PV system simulation model in order to be as accurate as possible in any climate change. The proposed model detects an abnormality in the power generation of the system and uses the obtained data to check the system for the most efficient system.

2. Materials and Methods of the PV System Simulation Model

2.1. Proposed Model for Accuracy Improvement (Interactive Curve Fitting)

The model was created from the analysis of the most suitable curve to represent the data in each point of interest in order to estimate the value of the data, which is called a 'curve-fitting process' [19,20], and to select the target function in many forms for analysis, which is called the 'optimization process'. The functions used in this model are shown as follows.

Exponential models provide a 1-term and a 2-term exponential model given by Equations (1) and (2).

$$\psi_{exp,1} = ae^{bx} \quad (1)$$

$$\psi_{exp,2} = ae^{bx} + ce^{dx} \quad (2)$$

Exponential models are often used when the quantity's rate of change is proportional to the quantity's initial amount. If the coefficient associated with b and/or d is negative, y represents exponential decay. If the coefficient is positive, y represents exponential growth.

The Fourier series models are the summation of the sine and cosine functions that describe a periodic signal. It represents them in either the trigonometric form or the exponential form. This fitting curve, shown in Equation (3), provides this trigonometric Fourier series form:

$$\psi_{ft} = a_0 + \sum_{i=1}^{n_f} a_i \cos(iwx) + b_i \sin(iwx) \quad (3)$$

where a_0 is the model's a constant (intercept) term in the data, and is associated with the $i = 0$ cosine term; w is the signal's fundamental frequency; n_f is the number of terms (harmonics) in the series $1 \leq n_f \leq 5$.

The Gaussian model fits peaks; and is given by

$$\psi_{gs} = \sum_{i=1}^{n_g} a_i e^{-\left(\frac{x-b_i}{c_i}\right)^2} \quad (4)$$

where a is the amplitude, b is the centroid, c is related to the peak width, and n_g is the number of peaks to fit; this provides a 1-term and a 2-term Gaussian model.

Polynomial models for curves are given by

$$\psi_p = \sum_{i=1}^{n_p+1} p_i x^{n_p+1-i} \quad (5)$$

where $n_p + 1$ is the polynomial order, n_p is the polynomial degree, and $1 \leq n_p \leq 5$. The order gives the number of coefficients to be fitted, and the degree gives the highest power of the predictor variable.

The sum of sines model fits periodic functions, and is given by

$$\psi_{si} = \sum_{i=1}^{n_{si}} a_i \sin(b_i x + c_i) \quad (6)$$

where a is the amplitude, b is the frequency, and c is the phase constant for each sine wave term. n_{si} is the number of terms in the series, and it provides 1-term and a 2-term Sum of Sines models.

Each of the methods above is estimated to be accurately calculated as a function based on the studied database's distributed behavior.

2.2. One Diode Equivalent Circuit

In order to understand solar cells' electrical properties, the equivalent circuit of solar cells [7,8]—as shown in Figure 1—was considered. The solar cell consists of a constant current source, which is the current of the solar cells caused by the solar irradiance (I_{ph}), the p - n junction of a diode, the series resistance in the solar cell (R_s), and the parallel resistance within the solar cell (R_{sh}).

The photocurrent (I_{ph}) from solar irradiance will be divided through the diode and R_s , and then load (R_L). The current flowing through R_s and R_L is I .

$$I_{ph} = I_d + I \quad (7)$$

$$I = I_{ph} - I_d \quad (8)$$

where I_{ph} is photocurrent (A), I_d is the current generated by the diode (A), and I is current that the solar cells produce (A).

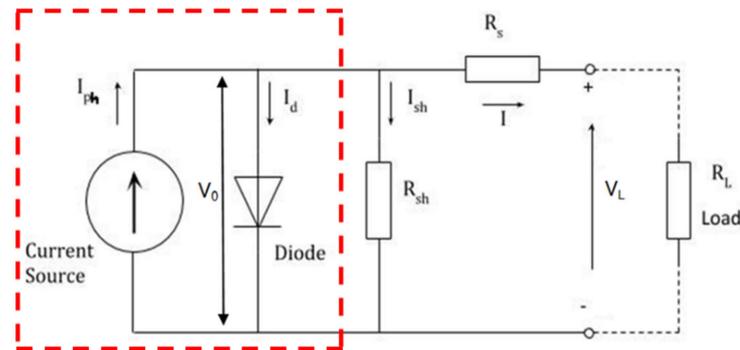


Figure 1. Ideal single diode equivalent circuit of a photovoltaic cell [7,8].

When $R_s = 0$, and I_d is given by

$$I_d = I_0 \left[\exp\left(\frac{qV_0}{nkT_m}\right) - 1 \right] \quad (9)$$

where I_d is the current generated by the diode (A), I_0 is the reverse saturation current of the diode (A), q is the electron charge value (1.602×10^{-19} C), k is the Boltzmann constant (1.381×10^{-23} J/K), T_m is the actual temperature at the PV module (K), R_s is the series resistance (Ω), R_{sh} is the shunt or parallel resistance (Ω), and n is the ideal factor of a diode according to the PV technology involved (chosen from Table 1).

Table 1. The ideal factor depends on the PV technology [18,21].

PV Technology	Values
Monocrystalline	1.2
Polycrystalline	1.3
CdTe	1.5
CIS	1.5
GaAs	1.3
Amorphous silicon single junction	1.8
Amorphous silicon double junction	3.3
Amorphous silicon triple junction	5.0

By substituting Equation (9) into Equation (8), we obtain Equation (10):

$$I = I_{ph} - I_0 \left[\exp\left(\frac{qV_0}{nkT_m}\right) - 1 \right] \quad (10)$$

In the case of a short-circuit, the load to obtain the maximum current value of the solar cell (I_{sc}), $V_o = 0$, is given by

$$I = I_{ph} = I_{sc} \quad (11)$$

In this case, the photocurrent (I_{ph}) is I_{sc} , and is equal to the current generated by the solar irradiance. The circuit is opened at the load point for the maximum solar cell voltage (V_{oc}), $I = 0$, and $V_o = V_{oc}$.

$$I_{ph} = I_0 \left[\exp\left(\frac{qV_{oc}}{nkT_m}\right) - 1 \right] \quad (12)$$

$$\ln \frac{I_{ph}}{I_0} = \frac{qV_{oc}}{nkT_m} - 1 \quad (13)$$

$$\text{So : } V_{oc} = \frac{nkT_m}{q} \left(\ln \frac{I_{ph}}{I_0} + 1 \right) \quad (14)$$

where I_{ph} is the photocurrent (A), I_o is the reverse saturation current of the diode (A), q is the electron charge value (1.602×10^{-19} C), k is the Boltzmann constant (1.381×10^{-23} J/K), T_m is the actual temperature at the PV module (K), V_{oc} is the open-circuit voltage (V), and n is the ideal factor of a diode according to the PV technology involved (chosen from Table 1).

In the case of the load having $0 < R_L < \text{infinitely}$, the solar cell supplies current (I) and voltage (V) to the load between $0 < I < I_{sc}$ and $0 < V < V_{oc}$. The values of the current and voltage that cause both volumes to have the highest value are called the maximum current and the maximum voltage, respectively. This gives the maximum electric power, as shown in Figure 2.

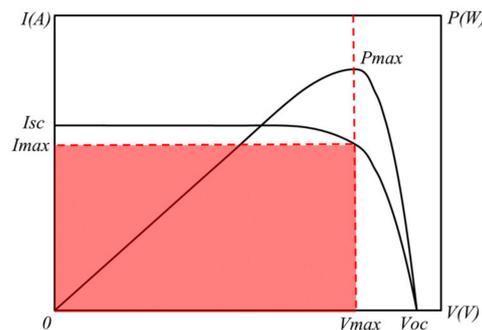


Figure 2. Electrical properties of the solar cells when exposed to solar irradiance.

2.3. Effects of the Series Resistance and Parallel Resistance of Solar Cells

From the equivalent circuit of a solar cell, as shown in Figure 3, while operating, it can be seen that the current from solar irradiance is represented by a constant current source (I_{ph}). It emits an electrical current in the opposite direction to the current that flows through the $p-n$ junction in ideal (I_d).

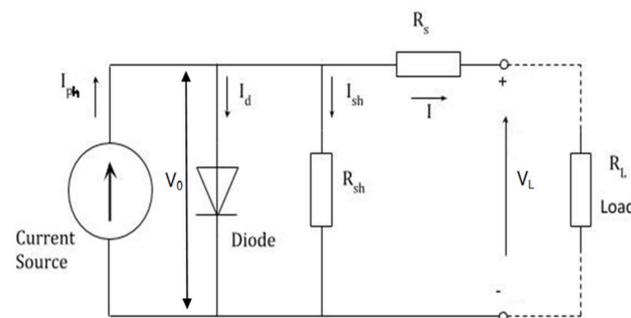


Figure 3. Equivalent circuit (Single diode) of a photovoltaic cell [7,8].

The series resistance occurred because of the semiconductor's resistance, the resistance of the ohmic contact area between the metal and the p and n parts, the resistance of the connecting wires, and the sum of all of the series resistance, abbreviated as R_s . Parallel resistance, R_{sh} , is a hypothetical resistance parallel to the boundary in the complete $p-n$ junction.

Ideally, the value of R_{sh} is infinite and, in the ideal solar cell, the value of R_s is zero. However, in practice, the semiconductor crystal has a breakdown. Joints, especially those with large areas, have defective parts, resulting in an incomplete $p-n$ junction. Therefore, the R_{sh} value is not infinite, and the result of the R_s value is not equal to zero. The effects of the R_s and R_{sh} values will change the properties of the solar cells. In general, the R_{sh} is not high enough to ignore the effect on the solar cell properties, but the R_s resistance will significantly affect the solar cell properties. In designing the structure's characteristics, the builders must then consider the effect of the R_s value.

$R_s = 0$ and I are given by

$$I = I_{ph} - I_0 \left[\exp\left(\frac{qV_0}{nkT_m}\right) - 1 \right] \quad (15)$$

If $R_s > 0$ and $R_{sh} < \infty$, $V_0 = V_L + IR_s$ is given by

$$I = I_{ph} - I_0 \left[\exp\left(\frac{q(V_L + IR_s)}{nkT_m}\right) - 1 \right] - \frac{(V_L + IR_s)}{R_{sh}} \quad (16)$$

where I_{ph} is the photocurrent (A), I_0 is the reverse saturation current of the diode (A), q is the electron charge value (1.602×10^{-19} C), k is the Boltzmann constant (1.381×10^{-23} J/K), T_m is the actual temperature at the PV module (K), R_s is the series resistance (Ω), R_{sh} is the shunt or parallel resistance (Ω), and n is the ideal factor of a diode according to the PV technology involved (chosen from Table 1).

From Equation (16) and Figure 4 showing R_s ' effect on solar cells, it can be seen that solar cells with a high R_s value will suffer fewer short-circuit currents. The slope of the graph was also decreased, resulting in much less energy from the solar cells. Therefore, when creating solar cells, there should be a way to reduce R_s to the minimum possible.

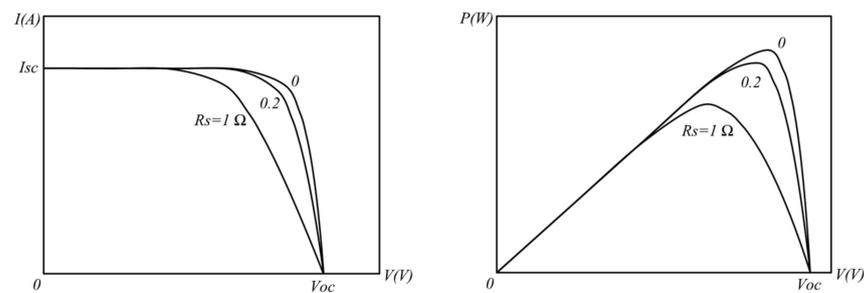


Figure 4. Influence of the series resistance on the solar cell output properties.

2.4. Effect of Solar Irradiance on the Solar Cells

The solar irradiance has a great effect on how the solar cell's electric current and power change, as shown in Figure 5.

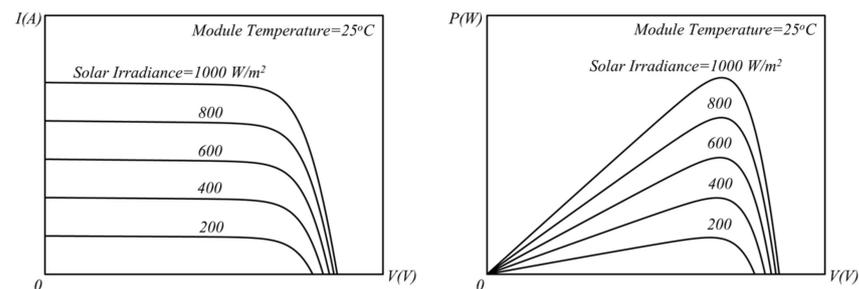


Figure 5. Effect of solar irradiance on the solar cells [7,8].

When $V = 0$, I is given by

$$I = I_{ph} = I_{sc} \quad (17)$$

2.5. Effect of Temperature on the Solar Cells

The temperature of the solar panel has a direct effect on the voltage. When the temperature is high, the voltage will drop and this will result in decreased power, as shown in Figure 6.

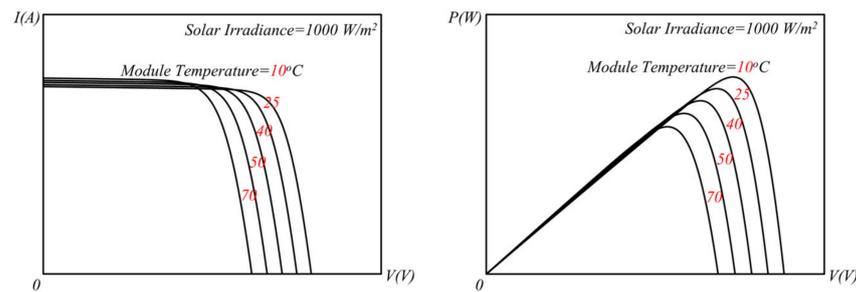


Figure 6. Effect of temperature on the solar cells [7,8].

When $I = 0$, $V_0 = V_{oc}$ is given by

$$V_{oc} = \frac{nkT_m}{q} \left(\ln \frac{I_{ph}}{I_0} + 1 \right) \quad (18)$$

2.6. Equation Using for PV Simulation Model

The basic structure of a solar cell consists of a $p-n$ junction of a silicon semiconductor represented by a diode and current source in an equivalent circuit, as shown in Figure 3. When the solar irradiance falls onto the solar cell, it creates charges at the $p-n$ junction, and moves by the electric field that occurs at the $p-n$ junction. For this reason, it causes the voltage at both poles. When connected to the load, the current flowing into the circuit is directly proportional to the solar irradiance. Equation (19) shows that the solar cells' current and voltage properties are in the form of an exponential equation [7,8]:

$$I = I_{ph} - I_0 \left[\exp \left(\frac{q(V_L + IR_s)}{nkT_m} \right) - 1 \right] - \frac{(V_L + IR_s)}{R_{sh}} \quad (19)$$

where I_{ph} is the photocurrent (A), I_0 is the reverse saturation current of the diode (A), T_m is the actual temperature at the PV module (K), R_s is the series resistance (Ω), R_{sh} is the shunt or the parallel resistance (Ω), and n is the ideal factor of a diode according to the PV technology involved (chosen from Table 1).

The PV system simulation model (1D5P) was created from Equation (19). It consists of a single diode (1D) and five main parameters (5P) [1–3]. The 5P consist of a current photo, I_{ph} ; the ideal factor according to the photovoltaic technology involved, n ; a parallel resistance, R_{sh} , expressing a leakage current, I_0 ; and a series resistance, R_s [7–9]. The PV system simulation model was analyzed using Kirchhoff's law, and was used to estimate the photovoltaic power output.

The photocurrent, I_{ph} (A) depends on the solar irradiance and module temperature. However, changes a lot with the solar irradiance. Consequently, this photocurrent can be expressed as Equation (20):

$$I_{ph} = \frac{G}{G_{ref}} [I_{sc,ref} + \mu_{sc}((T_m - T_{ref}))] \quad (20)$$

where I_{sc} is the short circuit current (A), T_m is the module temperature (K), T_{ref} is the module temperature at the standard test conditions (STC) (298 K), μ_{sc} is the temperature coefficients of I_{sc} ($A/^\circ C$), G is the solar irradiance (W/m^2), and G_{ref} is the solar irradiance at STC ($1000 W/m^2$).

The module temperature directly affects the reverse saturation current of the diode I_0 , as shown in Equations (21)–(23):

$$I_0 = I_{0,ref} \left(\frac{T_m}{T_{ref}} \right)^3 \exp \left[\left(\frac{qE_g}{nk} \left(\frac{1}{T_{ref}} - \frac{1}{T_m} \right) \right) \right] \quad (21)$$

$$I_{0,ref} = I_{sc} \exp\left(\frac{q(-V_{oc})}{nkT_m}\right) \tag{22}$$

$$\text{So : } I_0 = \left[I_{sc} \exp\left(\frac{q(-V_{oc})}{nkT_m}\right) \right] \left(\frac{T_m}{T_{ref}}\right)^3 \exp\left[\left(\frac{qE_g}{nk}\left(\frac{1}{T_{ref}} - \frac{1}{T_m}\right)\right)\right] \tag{23}$$

where I_{sc} is the short circuit current (A), T_m is the module temperature (K), T_{ref} is the module temperature at STC (298 K), V_{oc} is the open-circuit voltage (V), E_g is the band energy gap of the solar cell, and n is the ideal factor according to the PV technology involved.

It is impossible to ignore the series resistance (R_s) and the shunt resistance (R_{sh}). The form of the circuit that takes the series resistance (R_s) and the shunt resistance (R_{sh}) is shown in Figure 3. The solar cell connects in series to increase the voltage, and connects in parallel to increase the electric current. The voltage and current will vary according to R_s and R_{sh} .

The shunt or parallel resistance, R_{sh} , is defined in Equation (24):

$$R_{sh} = \frac{V_m + (I_m R_s)}{I_{sc,ref} - I_{sc,ref} \left\{ \exp\left[\frac{q(V_m + I_m R_s - V_{oc,ref})}{nkT_m}\right] \right\} + I_{sc,ref} \left\{ \exp\left[\frac{q(-V_{oc,ref})}{nkT_m}\right] \right\} - \frac{P_m}{V_m}} \tag{24}$$

where R_{sh} is the shunt or parallel resistance (Ω), R_s is the series resistance (Ω), P_m is the rated power of the PV panel (W), V_m is the maximum power voltage of the PV panel (V), I_m is the maximum power current of the PV panel (A), $V_{oc,ref}$ is the open circuit voltage of the PV panel (V), $I_{sc,ref}$ is the short circuit current of the PV panel (A).

3. Method of Accuracy Improvement

3.1. Preciseness Function Learning Model (PFL Model)

The proposed model is the data precision improvement process method based on a database that shows the application’s behavior to be studied. The amount of database data is either large or small, depending on all of the application’s behavior. The proposed model uses a process called the ‘Preciseness function learning model’ (PFL model). The process relies on various curve-fitting principles to bring out each format’s strengths to improve the data accuracy. The structure and method of operation of the process are shown in Figure 7.

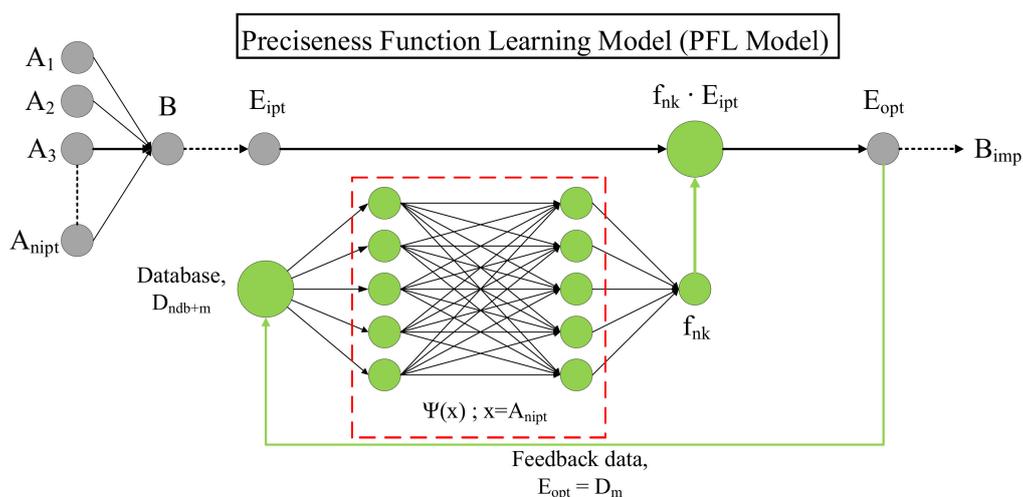


Figure 7. Structure of the Preciseness Function Learning Model (PFL Model).

The details and methods of each step can be explained as follows.

Step 1: The application’s functional format contains an input to obtain the output, and takes the output as the input of the model in order to improve the data’s accuracy. This can be written as Equation (25):

$$f(A_1, A_2, A_3, \dots, A_{n_{ipt}}) = B = E_{ipt} \quad (25)$$

where $A_1, A_2, A_3, \dots, A_{n_{ipt}}$ are the input of the studied application's computational process, B is the output of the studied application's computational process, and E_{ipt} is the input of the proposed model and the data that is used to improve the accuracy.

Step 2: The finding of the factors to improve the data's accuracy can be accomplished as follows:

$$f_{n_k} \times E_{ipt} = E_{opt} = B_{imp} \quad (26)$$

$$\text{So : } B_{imp} = f_{n_k} \times E_{ipt} \quad (27)$$

$$f_{n_k} = \left[\frac{\psi(x)}{E_{ipt}} \right]_{n_{db}} ; x = A_{n_{ipt}} \quad (28)$$

where f_{n_k} is the factor used to adjust the accuracy of the data, E_{opt} is the output of the proposed model and the data that has improved the accuracy, B_{imp} is the data that has been improved for accuracy, $\psi(x)$ is the relative function of the input and output of the studied application with n_{db} databases, and n_{db} is the number of the database.

In order to find the relative function of the input and output of the studied application, we analyzed the database's value into equations by a curve-fitting process in Table 2 (Exponential models, Fourier series models, Gaussian models, Polynomial models, and sum of sines models).

Table 2. Curve-fitting functions of the database.

Types	Functions
Exponential	$\psi_{exp}(x) = a_s e^{b_s x}$
Fourier	$\psi_{ft}(x) = a_{s_0} + \sum_{i=1}^{n_{sf}} a_{s_i} \cos(iw_s x) + b_{s_i} \sin(iw_s x)$
Gaussian	$\psi_g(x) = \sum_{i=1}^{n_{sg}} a_{s_i} e^{-\left(\frac{x-b_{s_i}}{c_{s_i}}\right)^2}$
Linear	$\psi_{li} = a_s x + b_s$
Polynomial	$\psi_p(x) = \sum_{i=1}^{n_{sp}+1} p_{s_i} x^{n_{sp}+1-i}$
Sum of sine	$\psi_{si}(x) = \sum_{i=1}^{n_{ssi}} a_{s_i} \sin(b_{s_i} x + c_{s_i})$

The $A_{n_{ipt}} \in E_{ipt}$ is shown in Figure 8.

Figure 8 shows the trendline from the database's estimation in the form of various equations with a curve-fitting process. The curve-fitting of the database in various forms has different curves, depending on the amount of data or the data distribution. A method of improving the data accuracy is the selection of the equation model using the curve-fitting process with the most accurate curve adjustment, which is chosen from the equation model with the R-square closest to 1, as summarized in Table 3.

The proposed model selects a curve-fitting function of the database with an R-square value closest to 1 from Table 3. The accuracy-improved data is given by

$$B_{imp} = \left[\frac{\psi_x(A_{n_{ipt}})}{E_{ipt}} \right]_{n_{db}} \times E_{ipt} \quad (29)$$

where B_{imp} is the accuracy improved data; $\psi_x(A_{n_{ipt}})$ is the relative function of the input and output of the studied application with n_{db} databases, from which the x function form is selected; E_{ipt} is the input of the proposed model and the data used for the accuracy

improvement, and n_i where I_d is currently generated by the diode (A), I_o is the reverse saturation current of the db is the number of the database.

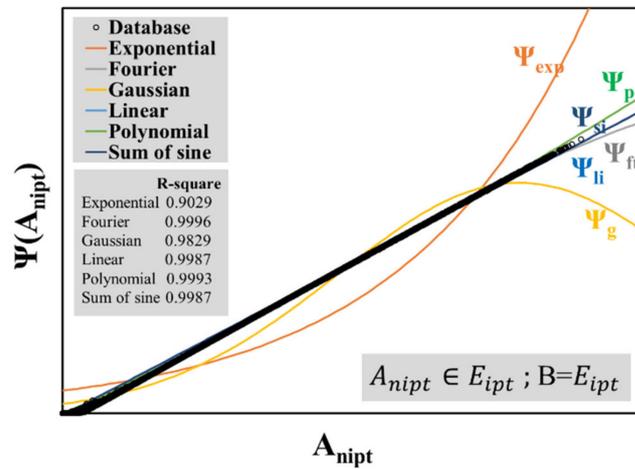


Figure 8. Example of a database of the model with curve-fitting in various forms.

Table 3. Example of the R-square values of the curve-fitting types.

Curve Fitting Types	R-Square
Exponential	0.9029
Fourier	0.9996
Gaussian	0.9829
Linear	0.9987
Polynomial	0.9993
Sum of Sine	0.9987

Step 3: The proposed model always updates its database after processing. That is, the proposed model learns to further optimize the processing precision. The Learning Method Structure process of the proposed model is shown in Figure 9.

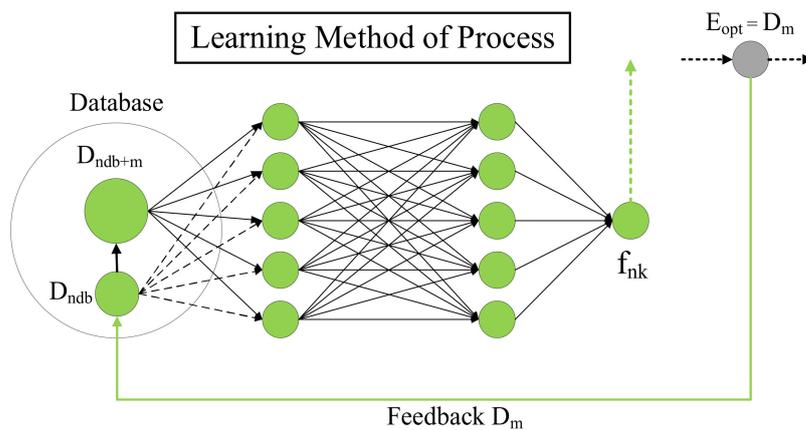


Figure 9. Learning method structure of the preciseness function learning model (PFL Model).

The details can be explained in equations, as follows:

$$Database = D_{n_{db}} \tag{30}$$

$$n_{db} = \frac{\sum_{i=0}^{n_{db}} x_i}{\bar{x}_{n_{db}}} \tag{31}$$

where D_{ndb} is the database at the first precision improvement, n_{db} is the number of the database at the first precision improvement, x_i is the number of databases, and $\bar{x}_{n_{db}}$ is the average amount of databases.

$$Data_{new} = B_{imp} = E_{opt} = D_m \quad (32)$$

$$m = \frac{\sum_{j=0}^m x_j}{\bar{x}_m} \quad (33)$$

where D_m is the accuracy improved data, m is the number of the data, x_j is the amount of data from the precision improvement, and \bar{x}_m is the average amount of data from the precision improvement.

When Equation (31) + Equation (33), we obtain Equation (34), as follows:

$$n_{db} + m = \frac{\sum_{i=0}^{n_{db}} x_i + \sum_{j=0}^m x_j}{\bar{x}_{n_{db}+m}} \quad (34)$$

where n_{db+m} is the number of the new database, and $\bar{x}_{n_{db}+m}$ is the average amount new databases.

$$\text{So : } Database_{new} = D_{n_{db}+m} \quad (35)$$

where $D_{n_{db}+m}$ is an updated database that contains the number of data n_{db+m} .

Through precision improvements from self-learning models, the data adds enhanced data to the database, in order to be more efficient in the next processing. This is given by

$$B_{imp} = \left[\frac{\psi_x(A_{n_{ipt}})}{E_{ipt}} \right]_{n_{db}+m} \times E_{ipt} \quad (36)$$

where B_{imp} is the accuracy improved data; $\psi_x(A_{n_{ipt}})$ is the relative function of the input and output of the studied application with n_{db+m} databases, and its selected x function form; E_{ipt} is the input of the proposed model, with the data that is used to improve the accuracy; and n_{db+m} is the number of the database.

3.2. Photovoltaic System Simulation Model with Preciseness Function Learning Model (PFL Model)

Figure 10 shows the structure of the photovoltaic simulation model with the preciseness reciprocation process. The structure describes the system of the model, which is divided into 2 parts:

Part 1: Photovoltaic simulation model, which has two inputs (solar irradiance and module temperature) and one output (PV power).

Part 2: Improving the accuracy of the photovoltaic simulation model's output by the proposed model (Preciseness Function Learning Model (PFL Model)).

The reason to improve the data's accuracy is that the results from the PV power output of the PV simulation model are the ideal result due to the loss of energy in various fields. Photovoltaic systems have many components that can affect electricity production, which is the source of this research. The structure's database is the PV systems data in the central region, Thailand. The database should contain at least one year because the overall behavior of photovoltaic power generation in Thailand depends on the seasonal weather. Thailand has three seasons in a period of one year.

The function format of the application is the 1D5P PV systems simulation model, according to Equation (19), which is given by

$$f(A_1, A_2) = B = E_{ipt} \quad (37)$$

where A_1 is the solar irradiance ($G, W/m^2$), A_2 is the module temperature ($T_m, ^\circ C$), and $B = E_{ipt}$ is the PV power output from the 1D5P PV systems simulation model (P_{max}, MW). $G \in P_{max}$ is shown in Figure 11.

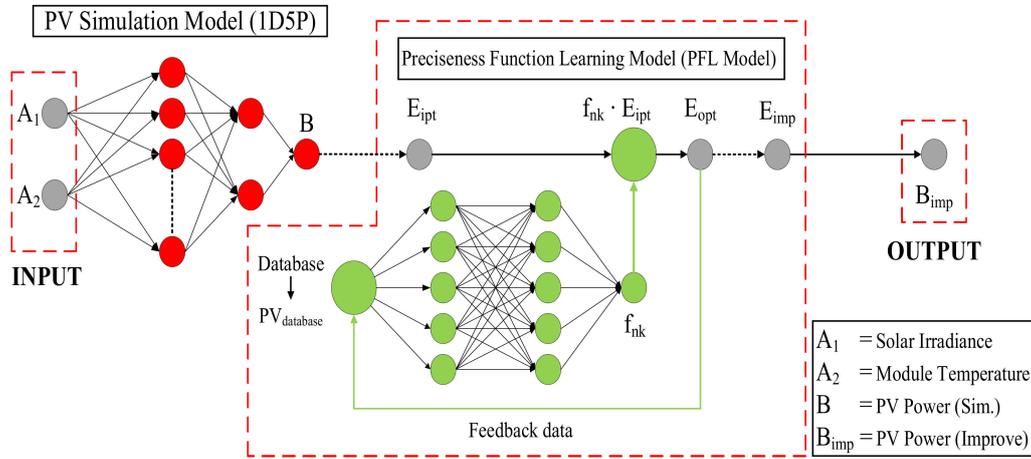


Figure 10. Structure of the photovoltaic simulation model with preciseness function learning model (PFL model).

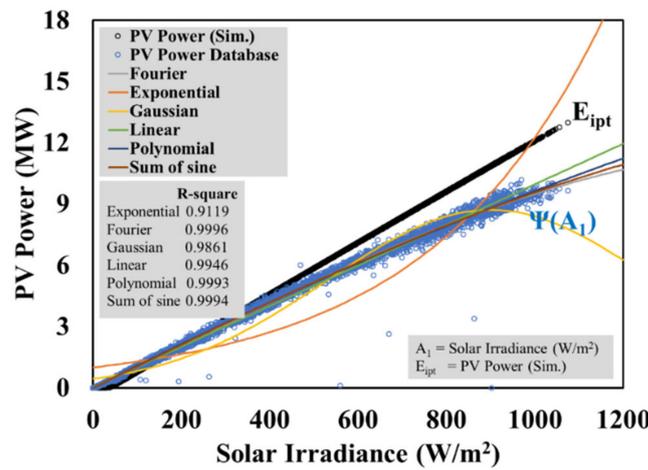


Figure 11. Accuracy improvement of the proposed model.

The proposed model selects a Fourier function with an R-square value closest to 1 (0.9996) from Table 4. The accuracy improved data is given by Equation (38). The simulation result is shown in Figure 12.

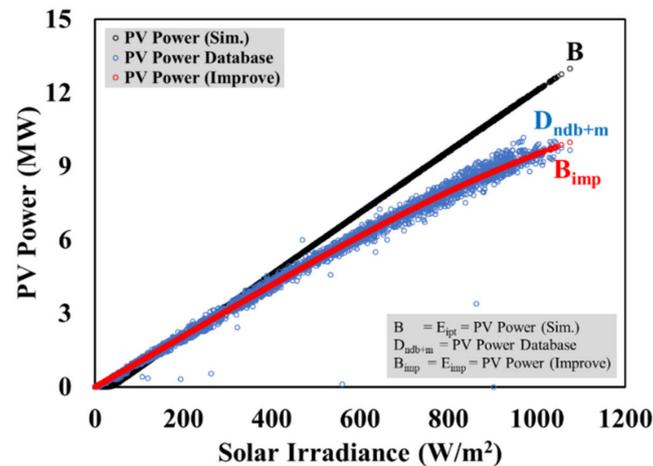
$$\text{So : } B_{imp} = \left[\frac{\psi_{ft}(A_1)}{E_{ipt}} \right]_{n_{db}+m} \times E_{ipt} \tag{38}$$

$$\psi_{ft}(A_1) = a_{s_0} + \sum_{i=1}^{n_{sf}} a_{s_i} \cos(iw_s A_1) + b_{s_i} \sin(iw_s A_1) \tag{39}$$

where B_{imp} is the accuracy improved PV power output (P_{max}, MW); a_{s_0} is the model and a_{s_i} is the intercept constant term in the data, which is associated with the $i = 0$ cosine term; w_s is the fundamental frequency of the signal; and n_{sf} is the number of harmonics terms in the series, in which $1 \leq n_{sf} \leq 5$.

Table 4. R-square values of the curve-fitting types.

Curve Fitting Types	R-Square
Exponential	0.9119
Fourier	0.9996
Gaussian	0.9861
Linear	0.9946
Polynomial	0.9993
Sum of Sine	0.9994

**Figure 12.** Simulation result of the accuracy improvement.

4. Evaluate Results and Discussions

4.1. Information and Simulation of the PV System

The case study applies the proposed model for the improvement of the PV system simulation model's accuracy using 2-year measured data (2018–2019) from the PV system in the central region of Thailand ($14^{\circ}10'78.1''$ north latitude and $100^{\circ}16'94.9''$ east longitude). The PV system consists of 48,980 PV panels and 12 inverters; one string consists of 24 PV panels connected in series, and two series are connected in parallel; seven arrays consisting of 10 strings are connected in parallel; two arrays consisting of eight strings are connected in parallel, and nine arrays are connected in one inverter. The experiment of the PV system is as shown in Table 5. The PV system has monitoring systems for all of the parameters, which are recorded every 1 min.

Table 5. Information about the PV systems for the experiment.

PV Systems	Rate Capacity	Panels	Number of PV Panel
Thailand (Central region)	12 MW	245 Wp	48,980

The pyranometer uses the KIPP&ZONEN band (CMP series) by installing it on the same plane as the PV panel, and the thermometer is installed under the PV panel. The PV panels uses the REC peak energy series band (REC245PE), as shown in Table 6.

In order to verify the accuracy of the proposed model with the application to the PV system simulation model, the testing and comparison of the simulation results with the actual measurement data can be divided into two conditions. These are the daily PV power output in two weather conditions (cloudy day and sunny day) and the effect of climate change for the electricity production of the photovoltaic systems. In this study, the model was designed to calculate the PV power output in different weather conditions, which are divided into seasons. These are in the summer (February–May), the rainy season (June–September), and the winter (October–January). The division of the seasons in Thailand

was taken from the Thai Meteorological Department. The test uses data from photovoltaic systems in Thailand (2018–2019). In order to verify the model that was used to improve the accuracy, the most accurate results to compare the % RMSE [22] values with the measured data is given by

$$\%RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\text{simulate}_i - \text{measured}_i}{\overline{\text{measured}}} \right)^2} \times 100 \quad (40)$$

where simulate_i is the simulation data, measured_i is the measured data and $\overline{\text{measured}}$ is the average value of the measured data.

Table 6. Information of the PV panel for the experiment.

Parameters of PV Panel	Value and Units
Rated Power of PV Panel, P_m	245 Wp
Open Circuit Voltage of PV Panel, V_{oc}	37.1 V
Short Circuit Current of PV Panel, I_{sc}	8.8 A
Maximum Power Voltage of PV Panel, V_m	30.1 V
Maximum Power Current of PV Panel, I_m	8.14 A
Ideality factor of PV Panel (n)—Poly Crystalline [18]	1.3

4.2. Daily PV System Simulation

Figure 13 shows a comparison of the PV power output between the proposed model and the actual measurement data of two weather conditions. The test results are as follows.

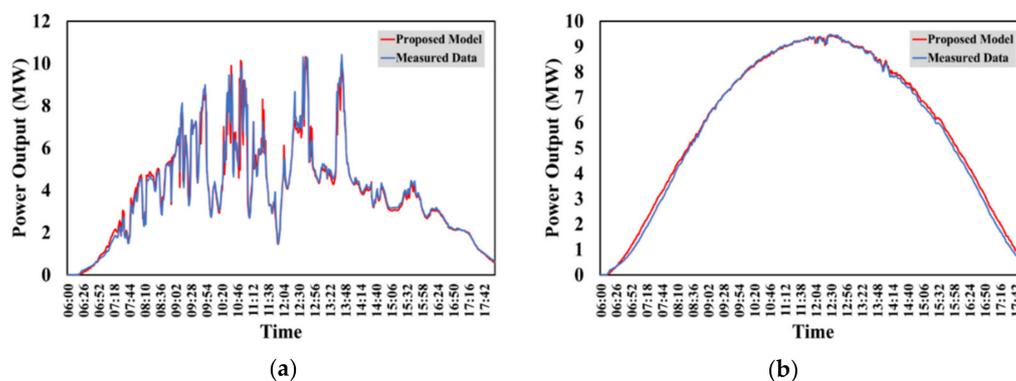


Figure 13. The proposed model and the measured data of the PV power output on a cloudy day (a) and a sunny day (b).

The PV power output of the proposed model tends to be in the same direction as the data from the actual measurement data on a cloudy day. This shows that this model can accurately simulate the PV power output in spite of sudden climate change. The nRMSE (normalized RMSE) is deficient (3.19%), as shown in Figure 13a.

The proposed model can accurately simulate the PV power output on a sunny day, and is very useful on days with the right weather conditions. The nRMSE (normalized RMSE) is lower than the cloudy day (1.79%), as shown in Figure 13b.

The PV power output of the proposed model and the actual measurement data are summarized in Table 7. The nRMSE (normalized RMSE) shows the accuracy of the proposed model.

Table 7. Comparison of the results on a sunny day and a cloudy day.

Case Day	Measured Data (MW)	Proposed Model (MW)	nRMSE (Normalized RMSE)
Sunny Day	71.43	72.70	1.79
Cloudy Day	64.57	66.63	3.19

4.3. Seasonal PV System Simulation

In order to compare the electricity generation data of the proposed model and the actual measurement, the conditions are divided into three seasons: the summer, the rainy season, and the winter. The climate in each season is different. The climate is an essential variable for the electricity generation of the PV system, including solar irradiance, ambient temperature, and module temperature, etc. In this research, the proposed model’s accuracy is checked with the PV system simulation model application that calculates the PV system’s electricity generation in all climatic conditions. In the test, this research randomly generated electricity for one week of the three seasons in 2019.

Figure 14 compares the PV power output between the proposed model and the actual measurement data for the random week of the different seasons. The simulation results of the sampling of the photovoltaic system’s electricity production, the simulation results, and the data from the actual measurements tend to show a similar trend. The daily electricity generation data for all three seasons have different weather conditions. However, the proposed model shows that this model can accurately simulate the electricity production in the different weather conditions.

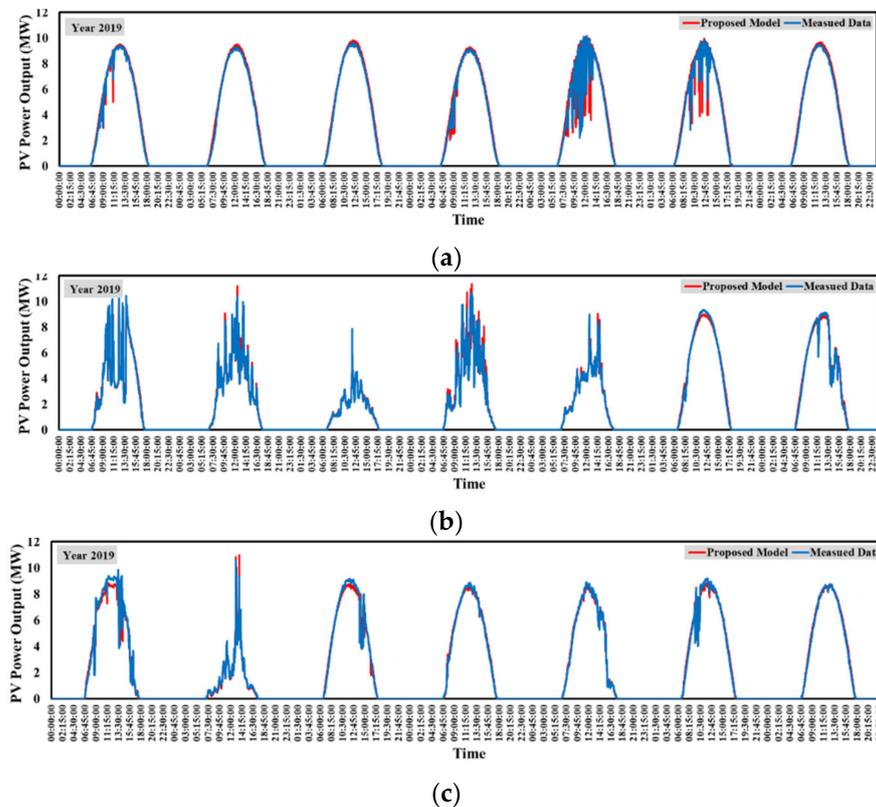


Figure 14. The proposed model and the measured data of the PV power output in the three seasons. (a) In the summer; (b) in the rainy season; (c) in the winter.

The accuracy of the proposed model of the monthly PV energy simulation compared with the actual measurement data in 2018 and 2019. When analyzing the simulation results of each season, such as in the summer, the rainy season, and the winter in Thailand’s

tropical climate, the proposed model can simulate the PV energy precisely in all weather conditions, as shown in Figure 15.

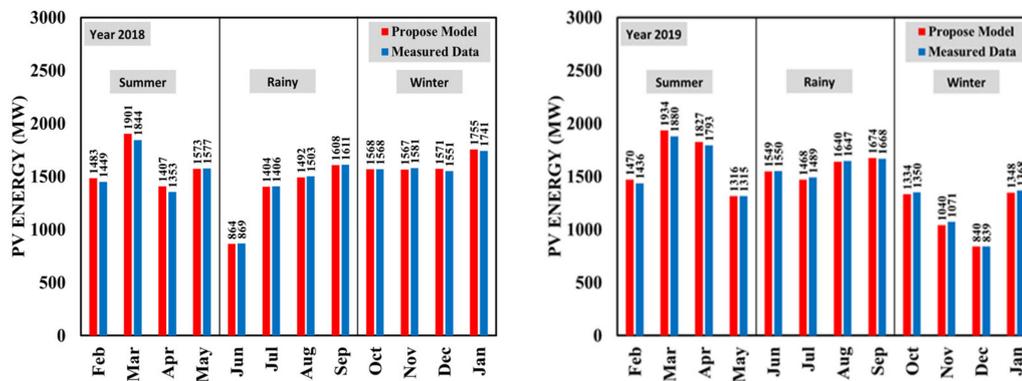


Figure 15. Comparison of the PV energy of the simulated and measured data of the different seasons.

Table 8 shows the proposed model of the monthly PV energy and the actual measurement data (2018). The proposed model results from 12 months of PV energy in Thailand, showing that the average PV energy is 1516 MWh/month, and the PV energy is 18193 MWh/year. The actual measurement data has an average PV energy of 1504 MWh/month, and PV energy of 18053 MWh/year. The nRMSE (normalized RMSE) ranges from 0.06% to 4.05%, which is very low. The average nRMSE (normalized RMSE) is 1.19% for 12 months 2018).

Table 8. Comparison of the results of the proposed model and the measured data in 2018–2019.

Data	2018				2019				2018–2019
	Summer	Rainy	Winter	Average	Summer	Rainy	Winter	Average	Average
Measured Data (MW)	1556	1347	1610	1504	1606	1589	1157	1450	1477
Proposed Model (MW)	1591	1342	1615	1516	1637	1583	1141	1453	1485
nRMSE (Normalized RMSE)	2.43	0.39	0.75	1.19	1.81	0.59	1.40	1.26	1.23

The test results of the PV energy of the proposed model and the actual measurement data on monthly basis (2019) showed that the average PV energy is 1453 MWh/month, and that the PV energy is 17440 MWh/year. The average PV energy is 1450 MWh/month, and the PV energy is 17406 MWh/year in the actual measurement data. The nRMSE (normalized RMSE) ranges from 0.07% to 2.91%, which shows that this value is deficient. The average nRMSE (normalized RMSE) is 1.26% for 12 months (in 2019).

In the comparison of the PV energy of the 24 months data of the proposed model (2018–2019) with the actual measurements for all three seasons, the average nRMSE (normalized RMSE) is 2.12% in the summer, 0.49% in the rainy season, and 1.08% in the winter. These are the results of the proposed model.

In the rainy season, the model is at its highest accuracy. We found that the solar cells in this area performed at their best because the dust was washed off the front of the solar panels by the rain, and because the temperature accumulation of the solar panels is not as high as it is in the summer. Because of this, it causes relatively little energy loss.

In the winter, the model is less accurate than it is in the rainy season, but more accurate than it is in the summer because there is a lot of dust on the front of the solar panel, which is a waste of energy for the solar cells. However, no heat is accumulated on the solar cells; thus, the models presented can be processed better than those in summer.

In the summer, the model is less accurate than both seasons because the summer temperature is high. The heat accumulation of the solar panel causes the temperature to rise.

In the actual case of a solar farm, the panel cleaning is not performed every day, but there is a cleaning cycle. Therefore, the accumulated dust affects the solar cell's electrical power generation efficiency.

The main factors that affect the PV system's electricity production are solar irradiance, which varies according to the solar cell's current, and the panel's temperature, which varies according to the voltage of the solar cell. The proposed model has a method that uses the relationship of solar irradiance and electricity generation to help improve the accuracy process, but the module temperature is not used as a parameter in this model. For this reason, the accuracy of the model in the summer is lower than it is in the rainy season.

4.4. Comparison with the PV Model with the Weight Function

The precision of the proposed model was verified by the comparison with the PV Model with the weight function [1] and the actual data from the one-year measured data, as shown in Figure 16.

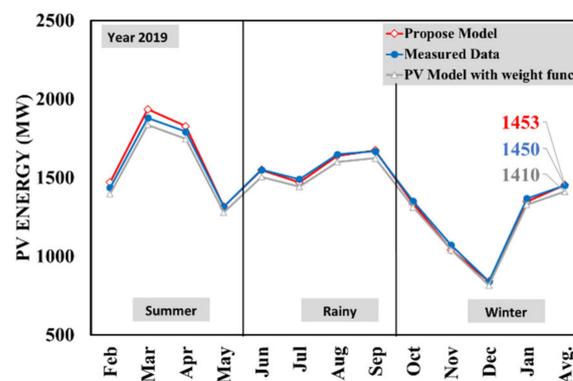


Figure 16. Comparison of the PV energy of the simulated and measured data of the different seasons.

The PV Model with the weight function [1] uses a method to quantify the difference between the simulated data and the actual measurement data in the polynomial equation form in order to improve the simulated data's accuracy. This method is accurate with a significant increase in the database, but as the database becomes more diffused, the accuracy of this method is lower.

On the other hand, the proposed model was designed to analyze various database behaviors diffused as equations in different forms of 'curve-fitting process'. The proposed model is a learning model that increases in accuracy every time it analyses and calculates data.

The results show that the proposed model has very high accuracy. The proposed model has a lower nRMSE than the PV Model with the weight function in all three seasons, as shown in Table 9.

Table 9. Comparison of the proposed model and the PV Model with the weight function in 2019.

Data	Proposed Model				PV Model with Weight Func.			
	Summer	Rainy	Winter	Average	Summer	Rainy	Winter	Average
Measured Data (MW)	1606	1589	1157	1450	1606	1589	1157	1450
Simulation Data (MW)	1637	1583	1141	1453	1564	1543	1124	1410
nRMSE (Normalized RMSE)	1.81	0.59	1.40	1.26	2.63	2.89	2.88	2.80

5. Conclusions

This research presents a simulation of a method to improve the accuracy of mathematical models (the Preciseness Function Learning Model (PFL Model)) using a PV simulation

model of the PV system, a precision test of the model, and a comparison of the 2-year data of the proposed model with the actual measurements of the PV system. The accuracy depends on the function selected from the curve-fitting process.

The results show that the proposed model has the highest accuracy in the rainy season. The average nRMSE (normalized RMSE) of the proposed model is very low (1.23%), and it ranges from 0.30% to 2.26%. It has been proven that this model is very accurate. The proposed model is a learning model that can optimize the higher precision with which the proposed model performs. It can be said that, after the accuracy improvement, the PV power output will be added to the proposed model database in order to enable the model to learn and improve the accuracy of the data in the next prediction results. In the future, we plan to develop models to improve the data accuracy by correlating more than one input to the outputs, in order to reduce the loss or error from the calculation during the data precision improvement.

Author Contributions: A.B., S.K., P.C. designed this experiment and prepared the manuscript. The experiments were carried out by A.B., S.K., P.C., K.S. and A.B., S.K., A.B., S.K., P.C., S.S., P.M., K.S., W.T., S.N. have analyzed the results and discussed the manuscript during the preparation. All authors discussed the results and implications and commented on the manuscript at all stages. All authors have read and agreed to the published version of the manuscript.

Funding: Details of organizations that funded the research and publication of article:

1. Thailand's National Electronics and Computer Technology Center.
2. Faculty of Engineering, King Mongkut's Institute of Technology Ladkrabang.
3. Thailand Graduate Institute of Science and Technology, Scholarship recipient code TG-44-22-59-016D.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of King Mongkut's Institute of Technology Ladkrabang.

Data Availability Statement: Data available on request due to restrictions e.g., privacy or ethical. The data presented in this study are available on request from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kittisontirak, S.; Bupi, A.; Chinnavornrungssee, P.; Sriprapha, K.; Thajchayapong, P.; Titiroongruang, W. An Improved PV Output Forecasting Model by Using Weight Function: A Case Study in Cambodia. *Int. J. Photoenergy* **2016**, *2016*, 1–10. [[CrossRef](#)]
2. Li, P.; Zhou, K.; Lu, X.; Yang, S. A hybrid deep learning model for short-term PV power forecasting. *Appl. Energy* **2020**, *259*, 114216. [[CrossRef](#)]
3. Guifang, G. Modeling of solar photovoltaic cells and output characteristic simulation based on Simulink. *J. Chem. Pharm. Res.* **2014**, *6*, 1791–1795.
4. De Soto, W.; Klein, S.A.; Beckman, W. Improvement and validation of a model for photovoltaic array performance. *Sol. Energy* **2006**, *80*, 78–88. [[CrossRef](#)]
5. Karatepe, E.; Boztepe, M.; Colak, M. Neural network based solar cell model. *Energy Convers. Manag.* **2006**, *47*, 1159–1178. [[CrossRef](#)]
6. Brano, V.L.; Orioli, A.; Ciulla, G.; Di Gangi, A. An improved five-parameter model for photovoltaic modules. *Sol. Energy Mater. Sol. Cells* **2010**, *94*, 1358–1370. [[CrossRef](#)]
7. Bupi, A.; Kittisontirak, S.; Sriprapha, K.; Siriwongrungsan, W.; Titiroongruang, W.; Niemcharoen, S. A simplified model for the estimation of energy production of PV module. In Proceedings of the 2017 International Electrical Engineering Congress (iEECON), Pattaya, Thailand, 8–10 March 2017.
8. Bupi, A.; Kittisontirak, S.; Sriprapha, K.; Siriwongrungsan, W.; Titiroongruang, W.; Niemcharoen, S. Comparison of PV module power output measurements. In Proceedings of the 2017 International Electrical Engineering Congress (iEECON), Pattaya, Thailand, 8–10 March 2017.
9. Vinod; Kumar, R.; Singh, S.K. Solar photovoltaic modeling and simulation: As a renewable energy solution. *Energy* **2018**, *4*, 701–712. [[CrossRef](#)]
10. Maninder Kaur, E. Simulation Model for PV Array & its Characteristics. *Int. J. Electr. Eng. Educ.* **2014**, *1*, 1694–2426.
11. Tsai, H.L. Insolation-oriented model of photovoltaic module using Matlab/Simulink. *Sol. Energy* **2010**, *84*, 1318–1326. [[CrossRef](#)]

12. Altas, I.H.; Sharaf, A.M. A photovoltaic array simulation model for Matlab-Simulink GUI environment. In Proceedings of the International Conference on Clean Electrical Power, Capri, Italy, 21–23 May 2007; pp. 341–345.
13. Vengatesh, R.P.; Rajan, S.E. Investigation of cloudless solar radiation with PV module employing Matlab-Simulink. In Proceedings of the International Conference on Emerging Trends in Electrical and Computer Technology, Nagercoil, India, 23–24 March 2011; pp. 141–147.
14. Maffezzoni, P.; Codecasa, L.; D'Amore, D. Modeling and Simulation of a Hybrid Photovoltaic Module Equipped with a Heat-Recovery System. *IEEE Trans. Ind. Electron.* **2009**, *56*, 4311–4318. [[CrossRef](#)]
15. Pandiarajan, N.; Muthu, R. Mathematical modeling of a photovoltaic module with Simulink. In Proceedings of the International Conference on Electrical Energy Systems, Chennai, India, 3–5 January 2011; pp. 258–263.
16. Rustemli, S.; Dincer, F. Modeling of Photovoltaic Panel and Examining Effects of Temperature in Matlab/Simulink. *Electron. Electr. Eng.* **2011**, *3*, 35–40. [[CrossRef](#)]
17. Nema, S.; Nema, R.K.; Agnihotri, G. MATLAB/Simulink based study of photovoltaic cells/modules/array and their experimental verification. *Int. J. Energy Environ.* **2010**, *1*, 487–500.
18. Bellia, H.; Youcef, R.; Fatima, M. A detailed modeling of photovoltaic module using MATLAB/SIMULINK. *J. Astron. Geophys.* **2014**, *3*, 53–61. [[CrossRef](#)]
19. Levy, E.C. Complex-curve fitting. *IRE Trans. Automatic Control* **1959**, *AC-4*, 37–43. [[CrossRef](#)]
20. O'Hagan, A. Curve Fitting and Optimal Design for Prediction. *J. R. Stat. Soc. Ser. B Methodol.* **1978**, *40*, 1–24.
21. McIntosh, K.R.; Honsberg, C.B. The Influence of Edge Recombination on a Solar Cell's IV Curve. In Proceedings of the 16th Euro-pean Photovoltaic Solar Energy Conference, Glasgow, UK, 1–5 May 2000.
22. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [[CrossRef](#)]