Scientific Research Journal,
Vol 17, No 1 (2020) 43-57
https://doi.org/10.24191/srj.v17i1.6321

Grid-Connected Photovoltaic System Performance Prediction Using Long-Term Weather Data

Nor Zaini Zakaria[1,3,4], Hedzlin Zainuddin[1,3,4*], Sulaiman Shaari[1,3,4], Ahmad Maliki Omar[2,4] & Shahril Irwan Sulaiman[2,4]

[1]Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[2]Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[3]SPECTRA Research Group, Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[4]GERC Research Group, Faculty of Electrical Engineering, 40450 Shah Alam, Selangor, Malaysia

Corresponding author's email: zainuddinhedzlin@gmail.com

Received: 27 July 2019
Accepted: 25 September 2019
Online First: 28 February 2020

ABSTRACT

This aim of this paper is to evaluate the accuracy of long-term weather data models for performance prediction of grid-connected photovoltaic (GCPV) systems. The analyses were done for a 6-year old metal deck roof retrofitted GCPV system located in Shah Alam, Malaysia. The monthly and annual energy yield of the actual field data for three consecutive years were compared with the predicted yield using the long-term weather data models. These models were the Typical Meteorological Year (TMY), Model Year Climate (MYC), Microclimate data, and Long-Term statistical Mean for ground station data at Subang. The findings can be a reference for photovoltaic (PV) system designers on the range of accuracy when using the weather data models for performance predictions of GCPV system in Malaysia.

Keywords: *GCPV system, performance prediction, long-term weather data, typical meteorological year, model year climate*

INTRODUCTION

Renewable energy (REN) sources are one of the ways forward for providing energy to mitigate climate change. The global final energy consumption and the growth in power capacity of REN continued to increase. From all the REN sources, solar PV has shown significant leading momentum in the REN generation. By the end of 2018, the energy production by solar PV has reached up to 640 TWh electricity production per year, or approximately 2.4 % global electricity generation per year [1].

Government of Malaysia has provided various schemes as incentives on renewable energy such as Feed-in-Tariff (FiT), Net Energy Metering (NEM) and Large Scale Solar (LSS). To date, under the FiT scheme with seven eligible REN resources, solar PV alone contributes 67% of the total installed capacity and 45% of the total energy generation [2]. For the NEM scheme, 500MW is allocated over five years of 2016 to 2020 [2] while 1000MWac is targeted by 2020 for the LSS scheme [3] These PV systems under these schemes are connected to the utility.

Grid-connected photovoltaic (GCPV) system is a type of configuration that is connected to a utility grid. The system comprises at least a PV array and Balance of System (BOS) components such as grid inverter, protection devices and cables. The systems are normally designed based on some limiting design constraints, such as the available area for installation, energy yield and budget. Among the pertinent tasks of a system, designer isis system sizing and system performance predictions. In the system performance predictions, the expected losses must be calculated as precise as possible as they will affect the accuracy of the energy generation prediction.

The loss factors in GCPV systems can be categorised as technical and environmental. The technical loss factors (LF_Tec) are as specified by the manufacturers, which include the efficiency of the grid inverter, voltage drop in cables, power tolerance and ageing. The environmental loss factors (LF_Env) are site dependence which includes weather data parameters of solar irradiation (H) and ambient temperature (Ta), dirt and as well as module temperature (Tm). Hence, the accuracy and reliability of the environmental loss factors are crucial in predicting the performance of GCPV systems. The energy yield can be calculated as [4]:

$$\text{Energy Yield} = P_{pmp\_array\_stc} \text{ x } LF_{\_Env} \text{ x } LF_{\_Tec} \qquad (1)$$

Whereby the Ppmp_array_stc is the peak array power rated at Standard Test Condition (STC), the values of the environmental loss can be obtained from several sources of weather data, in-situ measurements of Tm and estimation of dirt factor. The main sources of weather data are local meteorological station and global meteorological database. The meteorological stations measured real-time ground data at the site while the global meteorological database obtained data via satellite. Data for microclimate regions or other neighbouring regions can be derived or interpolated from the site data. The interval of data capture is normally in minutes that can be processed into hourly, monthly and annual data.

Historical weather data can be processed via several types of analysis. The most basic analyses are statistical maximum, minimum and mean values. Data is also commonly statistically processed into a one representative year models such as Typical Meteorological Year (TMY) [5],[6], Test Reference Year (TRY) [7], and Model Year Climate (MYC) [8]. The one-year models are widely used in building and energy simulation tools for performance predictions.

The main weather data parameters for solar PV applications are solar irradiation, ambient temperature, relative humidity and wind speed. Various mathematical models have been formulated to predict the most accurate and reliable energy yield for GCPV systems. These mainly focused on the weather parameters and algorithms [9],[10],[11],[12] forecasting for the best or optimum predictions [13],[14] and type of mounting [15],[16],[17],[18]. The concern was on the mathematical models rather than the reliability of the source and type of weather data. Studies comparing the accuracy and reliability of GCPV system performance using one-year weather data models have concluded that these models might not be suitable for PV performance evaluations [19],[20],[21].

Besides, solar irradiation map is another method commonly used as a source of solar irradiation data. For the distribution continuity of the solar irradiation values on the map, the data were usually microclimate data derived or interpolated from several stations data. The latest Solar

Irradiation Map for Malaysia has been produced based on the modified TMY model for 40 sites with historical measured ground data of 2-65 years from 1951-2015 [22]. This map incorporated a matrix for latitude and longitude of sites, thus enabling users to select sites and get the respective H.

In summary, there are many studies on the performance predictions of the GCPV system regarding the mathematical models and algorithms. However, these studies have not evaluated the accuracy and reliability of weather data models. To date, there are limited documented studies on the suitability of the weather data models.

With the rapid growth of solar PV installations in Malaysia, studies for accurate and reliable weather data models would be significant to address the related issues. Thus, this study aims to evaluate the accuracy of long-term weather data models for performance prediction of grid-connected photovoltaic (GCPV) systems. Accordingly, the objectives are to select the appropriate model of weather data, to compare the field energy yield with the predicted yield using the selected weather data and lastly to determine the most accurate model for the energy yield prediction.

METHOD

This section describes the GCPV system, source of weather data, weather data models and the mathematical models for the analyses. The GCPV system chosen provided the field energy yield data used in this study.

GCPV system

The GCPV system was installed at Green Energy Research Center in the Shah Alam campus of University Teknologi MARA. This is a 6-year old 5.4 kWp polycrystalline retrofitted on a metal roof, as shown in Figure 1. The system is connected to an automated data monitoring system where the real-time field environmental and electrical data were logged at a 5-min interval. The environmental data were solar irradiance (G), Ta, and Tm. The electrical data were current, voltage, power and energy yield. Figure 2 shows the schematic of the monitoring system.Figure 1: GCPV

Figure 1: GCPV System Retrofitted on a Metal Roof (Source: Nor Zaini Zakaria,2018

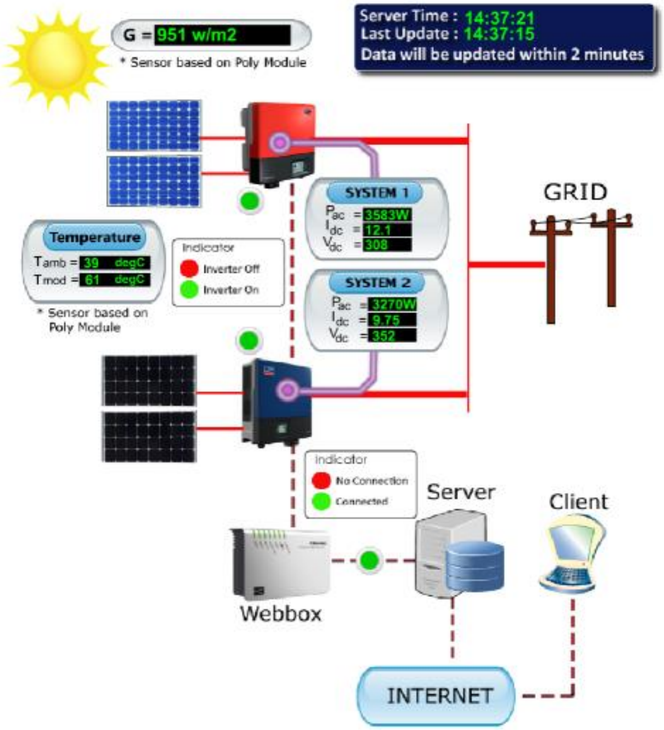


Figure 2: Schematic of the Monitoring System (Source: GERC, 2018)

Source of weather data

The weather data were taken from two main sources, which are the actual field data at the site in Shah Alam and ground data for Subang obtained from the meteorological station. The SEDA Solar Map was also used to provide the solar irradiation data at the site.

Weather data models

i. TMY Subang
This is the modified TMY model from 15 years Subang data for the years 1999-2013.

ii. MYC Subang
This is the MYC model from 21 years Subang data for the years 1975-1995.

iii. SEDA Solar Map
This is microclimate interpolated model for the actual site in Shah Alam from the Solar Irradiation Map of Malaysia. The respective H is obtained via GPS coordinate of Latitude 3.00° and Longitude 101.5°[22]

iv. AVGR
This is the long-term mean model of 15 years Subang data for the years 1999-2013.

Data analysis

The monthly and annual actual Field Energy Yield (FY) were compiled for the years 2015, 2016 and 2017. The Predicted Energy Yield (PY) were calculated by substituting H of the four selected weather data models as expressed by [4]

$$PY = P_{pmp\_array\_stc} \times PSH \times DF_{\_total} \tag{2}$$

Where,

PSH is the ratio of H to G at sea level (hour)

$DF_{\_total}$ is the total derating factors (technical & environmental)

The PY from the TMY, MYC, SEDA Solar Map and AVGR models were compared with the FY. The accuracy was determined by the percentage difference between the FY and the PY of each weather data model as expressed by:

$$\%Diff = 2 \times (FY\text{-}PY) / (FY + PY) \tag{3}$$

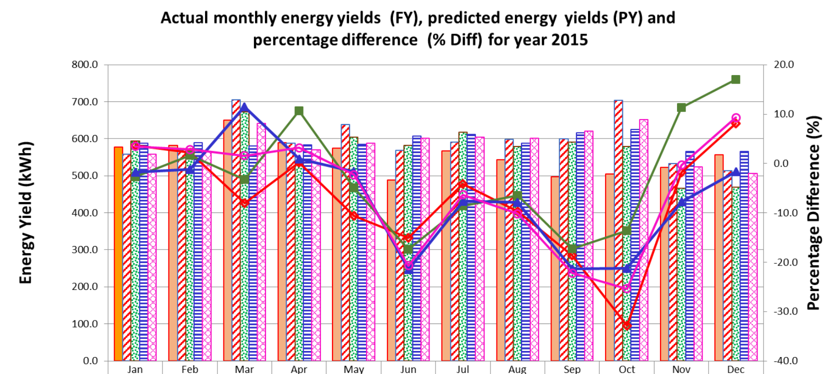
RESULTS AND DISCUSSION

In this study, the GCPV system performance indices are limited to only FY. In applying the mathematical model as in Equation 2, the Ppmp_array_ stc was a constant value of 5.405 kWp. The total derating factor DF_total for each month were calculated from the actual FY and was taken as a constant for the monthly analysis of PY for each TMY, MYC, SEDA Solar Map and AVRG weather data models. The PSH taken from the respective models were the only parameter varied in the calculation of PY in Equation 2. The % Diff of each model was calculated and analysed for three consecutive years of 2015, 2016 and 2017. This can be used for the performance predictions provided that other loss factors can be determined or assumed with an acceptable degree of accuracy. The results are discussed for each year and summarized for comparisons.

The year 2015

For the year 2015, as shown in Figure 3, the highest energy generated (FY) by the GCPV system was 651.0 kWh in March. This is consistent with the typical projection of PV system generation in Malaysia due to high solar irradiation between February, March and April [23],[24]. The lowest FY was 489.1 kWh in June that needs to be further investigated because the lowest solar irradiation recorded in a similar study was between November, December and January [24],[23].

In the comparison evaluation for each model, the % Diff ranges from the lowest of 0.3 % in April up to the highest of 32.9 % in October both by the TMY model. All models show the lowest % Diff for February while the highest were September and October. The highest % Diff was observed in September and October for all of the models which could be due to the anomalies in the actual field weather for those months.

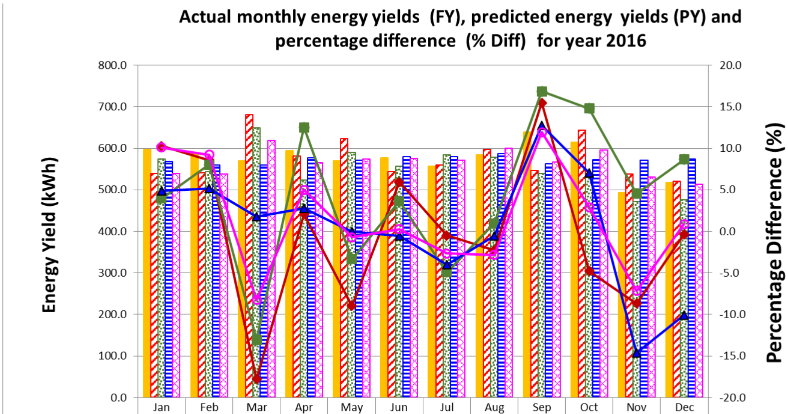


	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
FY_field	578.4	582.5	651.0	589.2	575.2	489.1	567.4	542.8	497.2	505.5	523.0	556.4
PY_TMY	557.7	569.7	705.6	587.7	639.2	569.0	591.5	598.0	599.6	704.5	532.6	513.3
PY_MYC	594.0	572.8	672.8	529.4	604.2	582.3	618.2	578.8	591.3	579.0	466.8	469.2
PY_SEDA	588.5	589.6	580.1	583.7	584.7	607.0	612.5	587.4	616.3	625.8	565.3	565.8
PY_AVRG	558.8	566.1	641.0	570.8	588.5	602.1	604.5	601.1	621.3	652.4	524.7	507.0
% diff_TMY	3.6	2.2	-8.0	0.3	-10.5	-15.1	-4.2	-9.7	-18.7	-32.9	-1.8	8.1
% diff_MYC	-2.7	1.7	-3.3	10.7	-4.9	-17.4	-8.6	-6.4	-17.3	-13.6	11.4	17.0
% diff_SEDA	-1.7	-1.2	11.5	1.0	-1.6	-21.5	-7.6	-7.9	-21.4	-21.3	-7.8	-1.7
% diff_AVRG	3.5	2.9	1.5	3.2	-2.3	-20.7	-6.3	-10.2	-22.2	-25.4	-0.3	9.3

Figure 3: Energy Yield and % Diff for the Year 2015

The year 2016

In this year, the highest actual FY was 638.6 kWh for September, as shown in Figure. 4. The month of November shows the lowest actual FY of 493.4 kWh and comparable to the lowest monthly irradiation data recorded[24],[23].



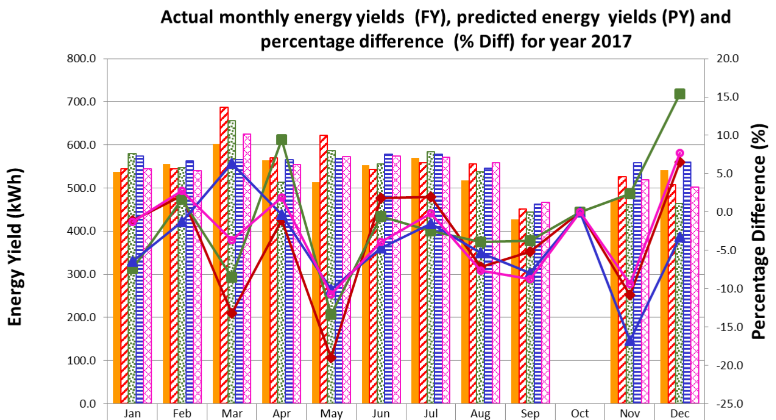
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
FY_Field	597.4	589.6	569.3	593.7	570.3	577.3	556.7	583.7	638.6	614.0	493.4	518.3
PY_TMY	539.0	541.1	680.5	581.5	623.8	544.0	559.4	597.3	547.0	644.0	538.1	520.1
PY_MYC	574.1	544.0	648.8	523.8	589.7	556.8	584.8	578.1	539.3	529.4	471.6	475.4
PY_SEDA	568.8	560.0	559.5	577.5	570.6	580.3	579.3	586.7	562.1	572.1	571.1	573.4
PY_AVRG	540.0	537.7	618.2	564.7	574.4	575.7	571.7	600.3	566.7	596.4	530.1	513.8
% diff_TMY	10.3	8.6	-17.8	2.1	-9.0	5.9	-0.5	-2.3	15.5	-4.8	-8.7	-0.3
% diff_MYC	4.0	8.0	-13.1	12.5	-3.3	3.6	-4.9	1.0	16.9	14.8	4.5	8.6
% diff_SEDA	4.9	5.2	1.7	2.8	0.0	-0.5	-4.0	-0.5	12.7	7.1	-14.6	-10.1
% diff_AVRG	10.1	9.2	-8.2	5.0	-0.7	0.3	-2.7	-2.8	11.9	2.9	-7.2	0.9

Figure 4: Energy Yield and % Diff for the Year 2016

In the comparison for each model, the percentage difference ranges from 0% in May by SEDA Solar Map model to highest of 17.8 % in March by TMY model. All models show the lowest percentage of less than 5 % difference for July and August. The worst months of highest % Diff was September whereby all models show over prediction ranging from 11.9 % up to 16.9 %. The month with the highest % Diff was also September for the year 2015.

The year 2017

For the year 2017, FY data for October was missing, as shown in Figure 5 due to some technical problems. The highest and lowest FY recorded was 602.4 kWh in March and 427.7 kWh in September respectively. For the % Diff, the lowest was recorded in February and July for all models.



	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
FY_Field	538.0	555.6	602.4	564.4	513.7	552.3	570.1	517.2	427.7	0.0	472.5	541.7
PY_TMY	544.0	543.7	687.8	570.3	621.5	542.6	559.0	555.6	450.5	0.0	526.7	507.8
PY_MYC	579.4	546.6	655.8	513.7	587.5	555.3	584.3	537.7	444.2	0.0	461.6	464.2
PY_SEDA	574.0	562.7	565.5	566.3	568.5	578.8	578.8	545.7	463.0	0.0	559.1	559.8
PY_AVRG	545.0	540.3	624.8	553.8	572.2	574.2	571.2	558.4	466.8	0.0	518.9	501.6
% diff_TMY	-1.1	2.2	-13.2	-1.0	-19.0	1.8	2.0	-7.2	-5.2	0.0	-10.9	6.5
% diff_MYC	-7.4	1.6	-8.5	9.4	-13.4	-0.5	-2.5	-3.9	-3.8	0.0	2.3	15.4
% diff_SEDA	-6.5	-1.3	6.3	-0.3	-10.1	-4.7	-1.5	-5.4	-7.9	0.0	-16.8	-3.3
% diff_AVRG	-1.3	2.8	-3.7	1.9	-10.8	-3.9	-0.2	-7.7	-8.7	0.0	-9.4	7.7

Figure 5: Energy Yield and % Diff for the Year 2017

Annual comparison

GCPV system performance isis also usually analysed as annual data. Thus, the above monthly data were calculated for the annual % Diff for each model, as shown in Table 1. This was used to determine the most accurate model for the energy yield prediction.

Table 1: Comparison of Annual Percentage Difference for All Models Weather Data Models

Wether Data Models	SEDA (%)	TMY (%)	MYC (%)	AVRG (%)	Best Model
Year 2015	8.8	9.6	9.6	9.0	SEDA
Year 2016	5.3	7.1	7.9	5.2	AVRG
Year 2017	5.8	6.4	6.3	5.3	AVRG
Mean	6.6	7.7	7.9	6.5	

For the year of 2015, the best model was the SEDA model with the lowest % Diff of 8.8%. This is followed by AVRG model of 9.0 % and the TMY and MYC models both had 9.6 %. The best model for the year 2016 was AVRG with the lowest % Diff of 5.2 %, followed closely by SEDA with 5.3%, TMY with 7.1% and lastly the MYC model with 7.9 %. The AVRG was also the best model for the year 2017 with 5.3 %. Thus, the AVGR model was the best model with accuracies ranges from 5.2 % to 9.0 %.

In summary, the best model comparing the three years was the AVRG model. This model was generated from the typical statistical mean of the long-term Subang data. The second-best SEDA model provided the site data at Shah Alam obtained from the SEDA Solar Map that was interpolated from the Subang ground data. The one-year models of TMY and MYC ranked third and fourth showing comparable accuracy with the average mean value of 7.7 % and 7.9 % respectively. These findings are agreeable with the previous study that concluded that one-year weather data models might not be suitable for PV performance evaluations [19],[20],[21]. In the selection of the one-year models, data with big variations from the typical values are rejected. Therefore, these models do not account for the extreme weather conditions.

Nevertheless, in this study, it is worth noted that the AVGR model was generated from 15 years of weather data. For further studies, it is recommended to evaluate the accuracy of this long-term statistical mean model if lesser numbers of years were used. Besides, it is also worthwhile to observe the reliability of the one-year TMY (1999-2013) and MYC (1975-1995) models despite the big gap in historical data years.

CONCLUSION

In this study, the AVRG model was found to be the best long-term weather data model of H in predicting GCPV system performance, followed by SEDA Solar Map, TMY and MYC models. Nevertheless, this study also shows that the SEDA Solar Map model is comparably accurate to AVGR model and it is a more practical approach via only GPS coordinates in obtaining the relevant H. However, many factors need to be evaluated indepth, to derive to an accurate long-term weather data models that can be used in many other applications of solar-related technology.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the Sustainable Energy Development Authority (SEDA) Malaysia for appointing the authors as consultants for processing the TMY data and developing it into the Solar Irradiation Map for Malaysia. SEDA Malaysia is also greatly appreciated for granting permission to share the TMY data for knowledge and practical applications. Deep gratitude is also extended to the Ministry of Education, Malaysia, for a research grant FRGS/1/2019/TK07/UITM/02/13 where the funding has given a significant contribution to this study.

REFERENCES

[1] REN 21, Renewables 2019 Global Status Report.15 Feb 2020 2019. [Online]. Available: https://www.ren21.net/wp-content/uploads/2019/05/gsr_2019_full_report_en.pdf

[2] SEDA Annual Report, SEDA Malaysia. Accessed on Nov. 2, 2018 from www.seda.gov.my.

[3] Energy Commission Malaysia, 2018. *Guidelines on Large Scale Solar Photovoltaic Plant for Conenection to Electricity Networks [Electricity Supply Act 1990 (Act447)]*. Accessed from www.st.gov.my

[4] SEDA Malaysia, 2016. SEDA Malaysia Grid-Connected Photovoltaics Systems Design Course. *Sustainable Energy Development Authority Malaysia.*

[5] K. Skeiker, 2007.Comparison of methodologies for TMY generation using 10 years data for Damascus, Syria. *Energy Conversion and Management, 48*(7), pp. 2090-2102. https://doi.org/10.1016/j.enconman.2006.12.014

[6] S. Wilcox and W. Marion, 2008. "Users manual for TMY3 data sets, National Renewable Energy Laboratory.

[7] Kwanho Lee, Hochun Yoo and Geoff J.Levermorec, 2010. Generation of typical weather data using the ISO Test Reference Year (TRY) method for major cities of South Korea. *Building and Environment,* 45(4), pp. 956-963. https://doi.org/10.1016/j.buildenv.2009.10.002

[8] A. Zain-Ahmed, 2000. Daylighting and shading for thermal comfort in Malaysian Building. Unpublished PhD thesis, University of Hertfordshire, United Kingdom.

[9] M. Z. Hussin, A. M. Omar, Z. M. Zain and S. Shaari, 2013. Performance of grid-connected photovoltaic system in equatorial rainforest fully humid climate of Malaysia. *International Journal of Applied Power Engineering (IJAPE), 2*(3), pp. 105-114. DOI: 10.11591/ijape.v2i3.2090

[10] S. I. Sulaiman, T. K. A. Rahman, I. Musirin, S. Shaari and K. Sopian, 2012. An intelligent method for sizing optimization in grid-connected photovoltaic system. *Solar Energy, 86*(7), pp. 2067-2082. https://doi.org/10.1016/j.solener.2012.04.009

[11] H. Zainuddin, S. Shaari, A. M. Omar, Z. Mahmud and F. M. Darus, 2011. Prediction of module operating temperatures for free-standing (FS) photovoltaic (PV) system in Malaysia. International Review on Modelling and Simulations (I.RE.MO.S.), 4(6), pp. 3388-3394.

[12] H. Zainuddin, 2014. Module Temperature Modelling for Free-Standing Photovoltaic System in Equatorial Climate. In The Doctoral Research Abstract, Vol 5(5) May 2014. Universiti Teknologi MARA, Shah Alam, Malaysia, 2014.

[13] K. A. Baharin, H. A. Rahman, M. Y. Hassan and C. K. Gan, 2016. Short-term forecasting of solar photovoltaic output power for tropical climate using ground-based measurement data. *Journal of Renewable and Sustainable Energy*, 8, p. 053701. https://doi.org/10.1063/1.4962412

[14] A. Mellit and A. M. Pavan, 2010. A 24-h forecast of solar irradiance using artificial neural network: Application for performance prediction of a grid-connected PV plant at Trieste, Italy. *Solar Energy, 84*(5), pp. 807-821. https://doi.org/10.1016/j.solener.2010.02.006

[15] N. Yusoff, N. Z. Zakaria, H. Zainuddin and S. Shaari, 2016. Operating temperature of photovoltaic module for retrofitted grid-connected photovoltaic system on metal roof. *International Journal of Simulation: Systems, Science and Technology*, 17, pp. 541-545. DOI: 10.5013/IJSSST.a.17.41.54

[16] N. Z. Zakaria, H. Zainuddin and Shaari "Thermal impact on cell temperature of retrofitted roof-mounted photovoltaic module," Int. J. Appl. Eng. Res, 9, pp. 7921-7927, 2014.

[17] N. Z. Zakaria, H. Zainuddin, S. Shaari, 2013. Critical factors affecting retrofitted roof-mounted photovoltaic arrays: Malaysian case study," in *Clean Energy and Technology* (CEAT), 2013 IEEE Conference 2013, pp. 384-388.

[18] N. Mat Yusoff, 2017. A revised model for power output of grid-connected photovoltaic system in Malaysian climate. Unplished Master thesis, Universiti Teknologi MARA, Shah Alam, Malaysia.

[19] S. R. Dean and P. ASA, 2010. Quantifying the variability of solar PV production forecasts in 39th *American Solar Energy Society* (ASES) National Solar Conference, Arizona, USA.

[20] P. Storck, J. Mccaa, S. Eichelberger, and A. Etringer, 2010. Integration of short-duration ground based measurements with long-duration satellite based measurement and numerical weather simulations to provide a more accurate assessment of solar energy availability and variability. In SOLAR 2010 *National Conference* 2010, Arizona, USA.

[21] C. Yimprayoon and M. Navvab, 2011. "Impact of different weather data sets on photovoltaic system performance evaluation," in ARCC Conference Repository, 559-568. https://doi.org/10.17831/rep:arcc%y364

[22] SEDA Malaysia, 2016. *Solar Irradiation Data for Malaysia*. Sustainable Energy Development Authority Malaysia, Putrajaya.

[23] M. M. M. Hanifah, H. Zainuddin and M. Z. Hussin, 2018. Establishment of standard reference environment for photovoltaic nominal operating cell temperature testing with dedicated approach for tropical region. *Indonesian Journal of Electrical Engineering and Computer Science, 12*(1), 433-440 . DOI: 10.11591/ijeecs.v12.i1.pp433-440

[24] M. M. M. Hanifah, H. Zainuddin and M. Z. Hussin, 2017. Evaluation of standard reference environment for photovoltaic nominal operating cell temperature testing in Malaysia. *Indonesian Journal of Electrical Engineering and Computer Science*, 8, pp. 245-252. DOI: 10.11591/ijeecs.v8.i1.pp245-252