Estimating Hourly Energy Generation of Distributed Photovoltaic Arrays: a Comparison of Two Methods

J. Tan[1]
N. A. Engerer[1,2,*]
F. P. Mills[1,3,4]
[1]Fenner School of Environment and Society, The Australian National University, Forestry Building 48, Linnaeus Way, Canberra, ACT 0200, Australia;
[2]National ICT Australia, Canberra Research Laboratory, Canberra, Australia;
[3]Research School of Physics and Engineering, ANU, Canberra, Australia;
[4]Space Science Institute, Boulder, CO, USA.
[*]**nicholas.engerer@anu.edu.au**
http://nickengerer.org
Keywords: solar energy, photovoltaics, radiation, clear-sky index, resource assessment

Abstract

Successful integration of high penetrations of unmonitored grid-connected distributed photovoltaic (PV) generation sites in Australia will require accurate estimates of the real-time energy generated by these systems. However, presently there are not accurate methods available for estimating the real-time total contributions of such systems. This paper explores two methods which may be capable of estimating real-time contributions using PV energy generation from 29 rooftop systems and observations from four pyranometers in Canberra, Australia. The first method uses pyranometer data and PV simulation models to produce estimates of energy generation at a given site. The second produces the estimate using neighboring PV systems by application of the K_{PV} methodology (the clear-sky index for photovoltaics). The first method is shown to have a slightly better performance with RMSE values in the range of 15-20% versus 15-25% for the second method. However we find that the accuracy of the two methods is nearly the same when pairwise distance is 5km or less. We conclude that both methods are viable for producing hourly energy generation estimates.

1. Introduction

In Australia, grid-connected distributed photovoltaic (PV) systems represent the majority of installed PV capacity (2.2 GW of 2.4 GW total in 2012) [24]. Grid-connected distributed PV systems are named for their geographical dispersion and the feeding of their output directly into the distribution network. These systems, typically ~ 1 – 100 kW in size, are expected to maintain vigorous growth in the coming years, even after removal of government incentives, due to falling PV costs and increases in grid power costs [24]. This is notable, as the optimal integration of these distributed PV systems into the Australian electrical grid is non-trivial. In some rural and exceptionally high-penetration areas, problems associated with solar intermittency have resulted in the installation of additional distributed PV generation being stopped [6].

Fortunately, one key finding from a recent study [6] shows that accurate forecasting paired with energy storage and/or fast response alternative stationary energy generation can allow this intermittency to be managed cost-effectively, lifting restrictions on future installations. However, accurate methods for forecasting the energy generation of distributed PV systems are not currently available and are the subject of ongoing research.

A key enabler for accurate forecasts is accurate estimation of the real-time contributions of distributed PV systems. In Australia this is not straightforward as the vast majority of microgenerators are not actively monitored, and their generation is only recorded during quarterly meter readings. This leaves distributors and energy market operators with insufficient information to make operationally useful estimates. It naturally follows that if the real-time contributions of PV systems are not known, then they cannot be forecast accurately and their intermittency cannot be managed optimally.

One proposed solution is to install a network of pyranometers (global horizontal irradiance sensors) across a region in order to sample the radiation resource in real-time, and then generate estimates of PV system performance from these measurements [18, 25]. Another uses satellite imagery of clouds [15], but the lack of regular interval, high-resolution satellite data in Australia limits this option. Another method uses sky cameras and cloud tracking, but cloud deformation and the three-dimensional, multi-directional nature of their movement makes tracking them and estimating the location of their shadows quite difficult [5]. A subset of the PV systems themselves may also work as a sensor network [2].

This paper explores two methods of producing hourly PV energy generation estimates. The first utilizes pyranometer measurements to estimate PV energy generation Engerer [8] (similar to methods discussed elsewhere [4, 3]). The second, uses the energy generation of nearby PV systems as the primary input data [11] (also similar to methods discussed elsewhere [19, 13]).

2. Data

The PV energy generation data came from two primary sources. The first was hourly PV system energy output (Wh) data from 26 Australian Capital Territory (ACT) school sites (ACT Government). The second from half-hourly-averaged PV system power output (W) collected at three CSIRO Weather and Energy Research Unit (WERU) monitoring sites. Half-hourly pyranometer data was collected from three collocated pyranometers at the CSIRO WERU sites and minute-resolution data from a fourth pyranometer at the ANU College of Engineering and Computer Science (CECS) Solar Roof project [1]. Both of these data sources were aggregated to hourly intervals. Finally, the weather data required for PV system simulation (wind speed and temperature) was obtained from the Australian Government Bureau of Meteorology Canberra Airport monitoring station (station number 70351). The locations and site information are presented in Figure 1 and in Table 1.

3. Methods

PV energy generation was estimated based on two different types of input data as described below.

3.1. PV Energy Generation Estimates from a Pyranometer

In order to generate estimates of the PV system energy generation from a pyranometer measurement, a five-step process following the methodology of Engerer [8] was used. An overview of this modeling methodology is presented in Figure 2. This consists of, first, separating the global horizontal pyranometer measurements into their beam (direct) and diffuse components using the Direct Insolation Simulation Code (DISC) model [20]. Next, the tilted beam and diffuse components are generated in the second and third steps. For the beam component, it is a simple geometric transformation to the plane of array (POA) surface [7]. The tilted diffuse component is more complicated and requires a model. There are many of these models to choose from for hourly data [9]. Here the approach of Reindl et al. [21] was used. This transposition model was selected over others based on the validation undertaken by Gueymard [14], which demonstrated the Reindl model having superior performance in situations where the primary inputs to the model were modeled themselves (e.g. no available observations of the beam and diffuse radiation). Additionally, it was chosen for consistency, as it is also used in the K_{PV} methodology (Section 3.2), a choice based on the model's good performance under clear skies [14, 11].

Next, the performance of the PV system must be simulated. This is accomplished in the fourth and fifth steps using the Sandia Performance Model [16] and the Inverter Performance Model [17]. The closest available panel and inverter types from the System Advisor Model databases [12] are used. In most cases, exact matches were available.

3.2. PV Energy Generation Estimates from a nearby PV System

The second method for estimating PV energy generation follows the K_{PV} methodology [11]. An overview of this methodology is presented in Figure 2. K_{PV} is the clear-sky index for photovoltaics, whose formulation is defined as:

$$K_{PV} = \frac{PV_{MEAS}}{PV_{CLR}} \tag{1}$$

where PV_{MEAS} is the measured energy generation from a system and PV_{CLR} is the simulated clear sky energy generation. Estimates for PV_{CLR} were created using the ESRA clear-sky model [22] and the Sandia Performance and Inverter Models. The choice of the ESRA model is based on its superior performance over other clear-sky models of similar complexity, as shown by Engerer and Mills [11] and Engerer [10]. The Sandia models are chosen due to the versatility of the models' underlying formulae, the open availability of these formulae and the large database of photovoltaic modules and inverter models available for simulation (500+ and 400+ respectively). This clear-sky index formulation removes variations due to diurnal and seasonal cycles, and those due to size, tilt and orientation of the PV systems. A value of 1 indicates perfectly clear skies. Values near 0 indicate thick cloud cover.

This formulation makes it possible to use the energy generated at one PV system (1) to estimate the energy generated at a nearby PV system (2), as long as the clear sky curve can be generated for both systems. This estimate is calculated by:

$$PV_{EST_2} = \frac{PV_{MEAS_1}}{PV_{CLR_1}} \cdot PV_{CLR_2} = K_{PV_1} \cdot PV_{CLR_2} \tag{2}$$

where PV_{CLR_2} represents the clear sky energy generation estimate at the nearby station at the given time. The accuracy of this calculation is primarily dependent on distance, cloud cover, and system specific issues such as the presence of shading, soiling or system level inefficiencies.

4. Results

The overall accuracy of the two PV generation estimation methods was assessed using all available hourly data for 2012-2013. More detailed comparisons examined the accuracy as a function of solar zenith angle and the separation between sites.

4.1. Using Collocated Measurement Sites

The three CSIRO sites included in this analysis provide an excellent point of comparison for the two methods, as they have collocated pyranometer and PV system measurements. Using the observations at these sites, we produced both pyranometer and PV system based estimates at all available sites (i.e. including the ACT Schools). In each instance, we use observations (either radiation or energy generation) from one of these three collocated sites to estimate the energy generation at one of the PV sites. In the course of analyzing the results of these experiments, two important observations were noted.

First, it became clear that there was a strong bias in the estimates from site 642 (see Table 2), which actually went on to reveal what appears to be a calibration error in the pyranometer at that location (visible in Figure 4). This is a very interesting observation because the notion of using PV systems as sensors is often derided as being of too low quality for scientific analysis, but here a possible calibration error in a pyranometer has been detected using PV systems. This suggests that groups of PV systems should perhaps be considered more favorably as tools for tasks such as local solar resource assessments, or as inputs to solar forecasting/nowcasting methods.

The second observation was that the proximity of site 643 was an issue. It was nearly 11km from the nearest PV site and over 42km from the furthest. For this reason, we chose the centrally located, well calibrated 641 site to produce an overall comparison, which is shown in Figure 3. It reveals that strong linear correlations are present in both of the methods, but with a visibly tighter correlation in the pyranometer approach. Spread is the greatest in both models at mid-range kWh/kWp values, likely owing to less uniform cloud conditions (as discussed in Engerer and Mills [11]). Additionally, the K_{PV} approach shows a bias towards underpredicting the energy generated under clearer conditions. This is likely due to poorer system performance at this site under intense radiation conditions, a hypothesis supported by the disappearance of this bias when we use other PV systems to generate intra-system estimates (e.g. Figure 6). Overall the performance of the two methods is comparable at this site, with better estimates from the pyranometer method at this site.

4.2. Using Sites Grouped by Distance

One of the primary advantages of using a PV system as the primary input towards estimating the energy generation at another site is that they are more widely available. This often means that there are PV systems nearby which are much closer than the nearest radiation monitoring equipment, which suggests that they may provide better estimates (e.g as in Golnas et al. [13]). In order to explore this possible relationship between accuracy and distance, the dataset was broken down into three groupings based on the proximity of stations. The first contains stations grouped according to their four nearest neighbors, the second by the farthest and the third by a random grouping of stations. We have denoted these the "tight", "loose" and "random" clusters. In each instance, each of the five sites are then used to estimate each other.

Results from the Engerer [8] and Engerer and Mills [11] K_{PV} methods are presented in Figures 5 and 6, respectively, with an overview of the errors presented in Table 3. For both methods, the lowest errors are achieved by using groupings of the closest sites, which is indicative of the proximity of sites being important for hourly level data (the same is not necessarily true of other averaging intervals). In the tight clusters, the mean distance and ranges between sites were 3.0 km with a range of 1.1 - 4.2 km for the Engerer [8] method and 0.6 km with a range of 0.3 to 1.1 km for the Engerer and Mills [11] K_{PV} method. For the loose clusters, these mean separations and ranges change to 18.0 km (12.3 - 29.0 km) and 35.2 km (32.8 - 41.1 km), respectively.

When comparing the two methods, the RMSE error for the K_{PV} approach is actually less than the pyranometer method for the tight clusters (13.6% versus 15.2%), with a notable tightening of the linear relationship over that in Figure 3. Overall mean error is still greater and the mean bias is not distinguishably better in either. This is a significant finding, as it demonstrates that PV to PV system hourly energy estimates may be produced with accuracy and mean error comparable to those produced via a pyranometer if the systems are "close enough". In order to explore this further, the RMSE error as plotted by distance is presented in Figure 7. Analysis of this figure indicates that for PV systems or pyranometers within 5km of one another, the hourly energy generation estimates produced by either method are equally suitable, where for distances greater than 5km, pyranometer estimates produce more accurate results. However, where pyranometer estimates are not present, PV system energy generation may be considered suitable (RMSE of 15-25% versus 15-20%). It should be stated that this relationship is only approximate at this stage and requires further, more detailed investigation.

4.3. Errors by Solar Geometry

It is also important to understand how the accuracy and mean error of these two methods changes throughout the day. For this purpose, the RMSE and Mean Absolute Error (MAE) are presented in Figure 8. These plots were generated by grouping the energy generation estimates into 5 degree bins of zenith angle (based on the mean zenith value over the given hour period) and producing the overall error measures in each bin. The pyranometer based method displays a clear advantage throughout the day using both error measures. Both methods show a flat response in RMSE with low zenith angles (15-50°), which then climbs quickly until falling slightly after 80-85°. MAE, unsurprisingly falls with increased zenith angle, as the PV systems are producing less energy at these low sun angles.

5. Conclusion

By comparing two methods for producing estimates of hourly energy generation from PV systems in Canberra, Australia, we have been able to explore the relationships that determine the accuracy of each. The first method uses a pyranometer to produce energy generation estimates as shown in Engerer [8] and the second uses the energy generation of another PV system via the K_{PV} methodology of Engerer and Mills [11]. In general, the estimates using a pyranometer have superior performance, with reduced absolute and relative errors. However, the difference between the two methods is not large (e.g. RMSE values of approximately 15-25% versus approximately 15-20%). The accuracy of the two methods is nearly the same for pairwise distances of approximately 5km or less. Furthermore, we have shown that in some cases a group of PV systems is even capable of revealing calibration errors in pyranometers. Future work in this area will focus on assessing the potential applications of these methods for PV project resource assessments, as well as further exploring the viability of these two methods in more depth. This is to include exploring the relationships between these two methods and the intermittency of solar radiation, overall cloud cover and using multiple sites for a single site estimate. Overall, the use of PV system energy generation for future scientific analysis and as cost-effective monitoring equipment is promising.

Acknowledgements

The authors would like to thank CSIRO's Weather and Energy Research Unit, headed by Alberto Troccolli, and the ACT Solar Schools program, managed by Melanie Pill and Kendra Wasiluk, for the provision of their PV and solar radiation data. Thanks also to the Australian Renewable Energy Agency who provided partial support for this project [2]. Finally, NAE would like to thank the U.S. National Science Foundation Graduate Research Fellowship Program and NICTA for providing partial support for this project.

References

[1] ANU College of Engineering and Computer Science, 2014. ANU CECS Solar Roof.
URL http://weather.cecs.anu.edu.au

[2] ANU-NICTA, 2012. Machine-learning-based Forecasting of Distributed Solar Energy Production.
URL http://arena.gov.au/project/forecasting-distributed-solar-energy-production/

[3] Cameron, C., Stein, J., Hansen, C., 2011. Evaluation of PV Performance Models and Their Impact on Project Risk. PV Rollout Conference.

[4] Cameron, C. P., Boyson, W. E., Riley, D. M., May 2008. Comparision of PV System Performance-Model Predictions with Measured PV System Performance. 2008 33rd IEEE Photovolatic Specialists Conference, 1–6.

[5] Chow, C., Urquhart, B., Lave, M., Dominguez, A., Kleissl, J., Shields, J., 2011. Intra-hour Forecasting with a Total Sky Imager at the UC San Diego Solar Energy Testbed. Solar Energy 85 (11), 2881–2893.

[6] CSIRO, 2012. Solar Intermittency: Australia's Clean Energy Challenge. Tech. Rep. June.

[7] Duffie, J. A., Beckman, W., 2006. Solar Engineering of Thermal Processes. John Wiley & Sons.

[8] Engerer, N. A., 2011. Simulating Photovoltaic Array Performance Using Radiation Observations from the Oklahoma Mesonet. Masters, University of Oklahoma.
URL http://nickengerer.org

[9] Engerer, N. A., 2014. Minute Resolution Estimates of the Diffuse Fraction of Global Irradiance for Southeastern Australia. Submitted to Solar Energy.
URL http://nickengerer.org

[10] Engerer, N. A., 2014. Validating Nine Clear-Sky Radiation Models in Australia. Manuscript in Preparation.
URL http://nickengerer.org

[11] Engerer, N. A., Mills, F. P., 2013. Kpv: A Clear-Sky Index for Photovoltaics. Solar Energy (Accepted) August.
URL http://nickengerer.org

[12] Gilman, P., Blair, N., Mehos, M., Christensen, C., Janzou, S., Cameron, C., 2008. Solar Advisor Model User Guide for Version 2.0. Tech. rep., NREL.

[13] Golnas, A., Bryan, J., Wimbrow, R., Hansen, C., Voss, S., Clemente, S., 2011. Performance Assessment Without Pyranometers: Predicting Energy Output Based on Historical Correlation. In: 37th IEEE Photovoltaic Specialists Conference (PVSC). Sandia National Laboratories.

[14] Gueymard, C., Mar. 2009. Direct and Indirect Uncertainties in the Prediction of Tilted Irradiance for Solar Engineering Applications. Solar Energy 83 (3), 432–444.

[15] Hammer, G. L., Nicholls, N., Mitchell, C. D., 2000. Applications of seasonal climate forecasting in agricultural and natural ecosystems. Springer.

[16] King, D., Boyson, W., Kratochvill, J., 2004. Photovoltaic Array Performance Model. Tech. Rep. December, Sandia National Laboratories.

[17] King, D., Gonzalez, S., Galbraith, G., Boyson, W., 2007. Performance Model for Grid-Connected Photovoltaic Inverters. Tech. Rep. September.

[18] Kuszamaul, S., Ellis, A., Stein, J., Johnson, L., 2010. Lanai High-Density Irradiance Sensor Network for Characterizing Solar Resource Variability of MW-Scale PV System. In: Paper presented at the 35th IEEE Photovoltaic Specialists Conference. Honolulu , HI, pp. 4–9.

[19] Lonij, V., Brooks, A., Koch, K., Cronin, A., 2012. Analysis of 80 Rooftop PV Systems in the Tucson, AZ Area. 38th IEEE Photovoltaic Specialists Conference (PVSC).

[20] Maxwell, E., 1987. A Quasi-Physical Model for Converting Hourly Global Horizontal to Direct Normal Insolation. Tech. rep., Solar Energy Research Institute.

[21] Reindl, D., Beckman, W., Duffie, J., Jan. 1990. Evaluation of Hourly Tilted Surface Radiation Models. Solar Energy 45 (1), 9–17.

[22] Rigollier, C., Bauer, O., Wald, L., 2000. On the Clear Sky Model of the ESRA - European Solar Radiation Atlas - With Respect to the Heliostat Method. Solar Energy 68 (1), 33–48.

[23] Watergroup Pty. Ltd & ACT Government Education and Training Directorate, 2014. ACT Public Schools Pulse Metering. URL `http://www.watergroup.com.au/actschools`

[24] Watt, M., Passey, R., 2013. PV in Australia 2012 Executive Summary. Tech. rep., Australian PV Association.

[25] Yang, D., Jirutitijaroen, P., Walsh, W. M., Dec. 2012. Hourly Solar Irradiance Time Series Forecasting Using Cloud Cover Index. Solar Energy 86 (12), 3531–3543.

Name	Rating (W)	Orientation (°)	Tilt (°)	Module type and number	Inverter type and number	Latitude	Longitude	Data Start
Campbell Primary School	10640	12	0	56 x TSM-190-DC01A	1 x STP10000TL	-35.289	149.156	12/10/2012
Lake Ginninderra College	10260	0	30	56 x TSM-190-DC01A	1 x STP10000TL	-35.238	149.074	18/08/2012
Black Mountain School	10260	-3	30	54 x TSM-190-DC01A	1 x STP10000TL	-35.264	149.112	18/08/2012
Chapman Primary School	10360	-22	30	56 x SF160-24-1M185	1 x Aurora PVI10.0-OUTD-AU	-35.356	149.042	18/08/2012
Duffy Primary School	10360	-5	15	56 x SF160-24-1M185	1 xAurora PVI10.0-OUTD-AU	-35.335	149.033	21/11/2012
Wanniassa Primary School	15540	-18	25	84 x SF160-24-1M185	1 x Sunny Tripower 15000TL	-35.398	149.085	14/09/2012
Majura Primary School	10260	2	30	54 x TSM-190-DC01A	1 x STP10000TL	-35.238	149.155	19/09/2012
Yarralumla Primary School	10260	-16	23	54 x TSM-190-DC01A	1 x STP10000TL	-35.306	149.102	18/08/2012
Charles Conder Primary School	15540	45	25	84 x SF160-24-1M185	1 x Sunny Tripower 15000TL	-35.464	149.098	18/08/2012
Lyneham High School	10260	-7	30	54 x TSM-190-DC01A	1 x STP10000TL	-35.252	149.130	18/08/2012
Canberra College	10260	10	23	54 x TSM-190-DC01A	1 x STP10000TL	-35.339	149.088	1/09/2012
Weetangera Primary School	10350	0	25	45 x SR-P660230	1 x SMC10000TL	-35.251	149.047	13/06/2013
Dickson College	20010	-5	15	87 x SR-P660230	2 x SMC10000TL	-35.249	149.153	25/05/2013
Lyneham Primary School	10350	-53	30	45 x SR-P660230	1 x SMC10000TL	-35.251	149.125	25/05/2013
Amaroo School	20010	0	10	87 x SR-P660230	2 x SMC10000TL	-35.164	149.128	16/06/2013
Southern Cross ECS	10350	15	20	45 x SR-P660230	1 x SMC10000TL	-35.232	149.038	2/06/2013
Namadgi School	29640	33	25	156 x STP190S-24/Ad+	3 x SMC10000TL	-35.391	149.067	16/06/2013
Stromlo High School	29670	-43	18	129 x SR-P660230	3 x SMC10000TL	-35.354	149.054	13/06/2013
Arawang Primary School	10350	0	20	45 x SR-P660230	1 x SMC10000TL	-35.355	149.059	13/06/2013
Lyons ECS	10350	0	38	45 x SR-P660230	1 x SMC10000TL	-35.341	149.076	16/06/2013
Melrose High School	20010	48	17	87 x SR-P660230	2 x SMC10000TL	-35.359	149.088	2/06/2013
Mawson Primary School	10350	-5	22	45 x SR-P660230	1 x SMC10000TL	-35.357	149.097	13/06/2013
Alfred Deakin High School	20010	0	18	87 x SR-P660230	2 x SMC10000TL	-35.324	149.094	13/06/2013
The Woden School	10350	-51	20	45 x SR-P660230	1 x SMC10000TL	-35.326095	149.092	13/06/2013
Red Hill Primary School	10350	28	25	45 x SR-P660230	1 x SMC10000TL	-35.338	149.132	16/06/2013
Narrabundah ECS	10350	-21	30	45 x SR-P660230	1 x SMC10000TL	-35.328	149.150	16/06/2013
Black Mountain (641)*	1560	-38	36	8 x SF160-24-1M195	1 x Aurora PVI-6000-OUTD-xx	-35.275	149.113	1/08/2012
Namadgi (642)*	4940	33	25	26 x STP190S-24/Ad+	1 x SMC10000TL	-35.392	149.067	2/02/2013
Wombat Hill (643)*	4000	-31	21	20 x STP200-18/ud	1 x SB3800	-35.534	149.155	1/08/2012
ANU CECS**	-	-	-	-	-	-35.276	149.1190	1/08/2012

Table 1: Basic site information for the photovoltaic and pyranometer measurement stations. The first 26 sites are schools in the Australian Capital Territory that have rooftop PV systems. Three of the sites, denoted *, are CSIRO Weather and Energy Research Unit sites that have collocated PV and pyranometer measurement stations. The last site, denoted **, is a single pyranometer located on the campus of The Australian National University on the roof of the College of Engineering and Computer Science Building. A much more detailed table of information, including the module and inverter simulation information is available at the corresponding author's webpage (http://nickengerer.org).

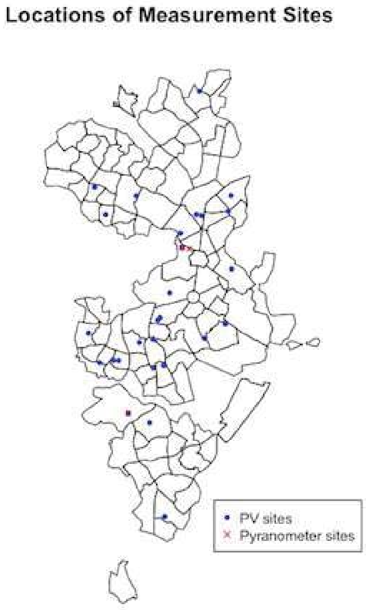

Figure 1: Map of the measurement sites presented in Table 2. Blue dots represent the ACT Schools PV systems, purple squares the CSIRO sites with PV and pyranometers collocated and the red marks standalone pyranometers (e.g. the CECS site). The sites are overlaid by a suburb map of the ACT.

Method	Site	RMSE(%)	MAPE(%)	MBE(%)
	641	15.9	75.0	0.7
Pyranometer [8]	642	19.9	130.7	5.9
	643	19.9	41.5	1.0
	CECS	16.2	64.8	2.1
	641	18.8	64.4	6.5
K_{PV} [11]	642	19.8	53.1	-6.5
	643	20.7	42.8	5.3
	CECS	-	-	-

Table 2: Overall error results for the two methods as reported for the hourly 2012-2013 data for the four pyranometer sites, three of which have collocated PV systems. Note that as the CECS site does not have a PV system, the K_{PV} method [11] cannot be evaluated at that location.

Method	Site	RMSE(%)	MAPE(%)	MBE(%)
	Tight	15.2	34.2	0.0
Pyranometer [8]	Loose	17.9	58.3	-3.7
	Random	14.4	71.6	-2.6
	Tight	13.6	59.3	-3.0
K_{PV} [11]	Loose	28.5	68.3	-0.5
	Random	21.0	63.1	-1.9

Table 3: A summary of the errors produced creating energy generation estimates at PV sites using a pyranometer [8] and using another PV system [11]. Error measures are broken down into tight, loose and random clusters, as based on pairwise distance.

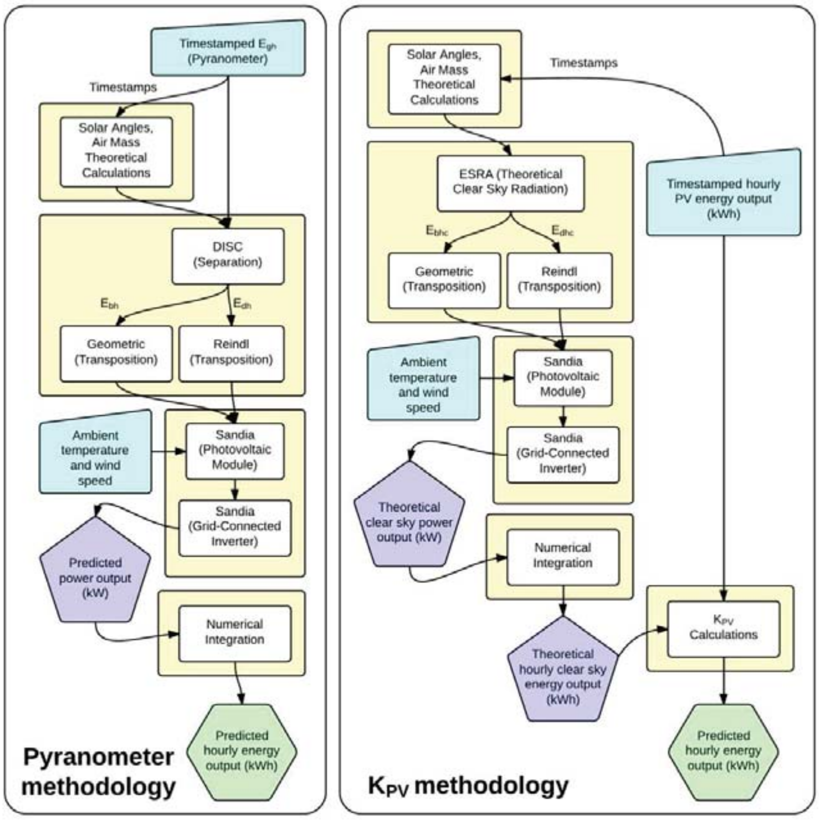

Figure 2: A flow-diagram for each of the two methods tested herein. The pyranometer method [8] is presented at left, showing how a system of models can be used to produce an energy generation estimate for a PV system using a measurement of global horizontal irradiance E_{gh} from a pyranometer. At right, the K_{PV} methodology [11], in which PV energy output is the primary input variable. The diagram is structured such that blue boxes represent inputs, white boxes are processes, purple pentagons are interim outputs and green hexagons are final outputs.

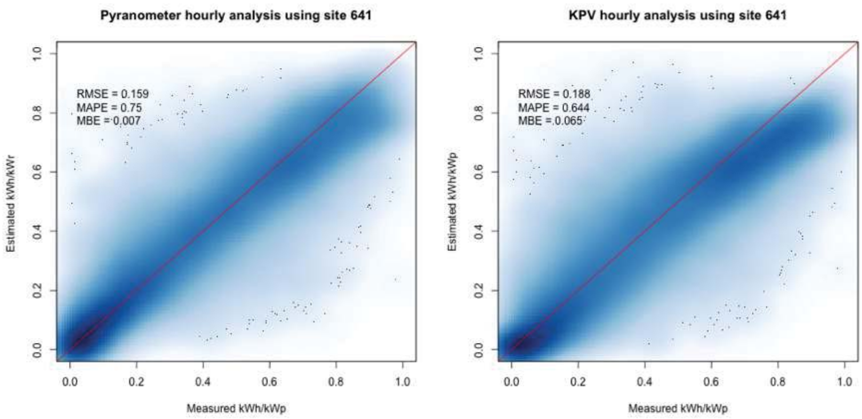

Figure 3: At left, the pyranometer based method [8], at right, the PV system based method [11] as computed for all hourly data for all sites from the 2012-2013 period and using site 641 to generate the estimates. Areas of darker shading indicate areas of a greater density of points. RMSE, MAPE and MBE errors are presented in terms of fractions of the respective mean (e.g. 0.159 = 15.9%).

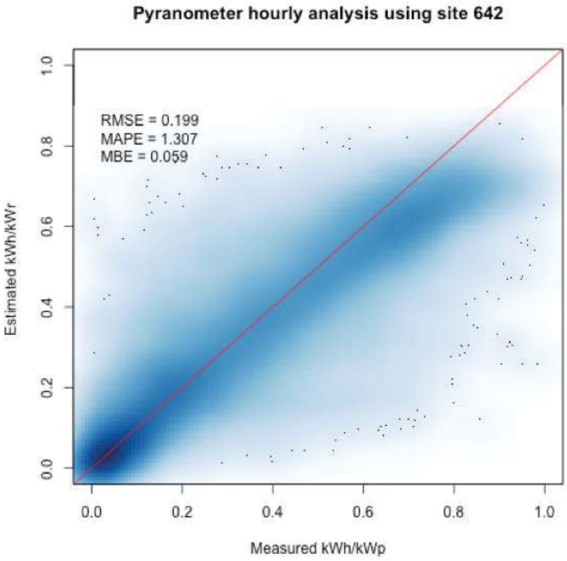

Figure 4: Pyranometer based estimates from 2012-2013 using site 642, which show a consistent under-prediction of PV system energy generation under clear conditions, which is indicative of a calibration error.

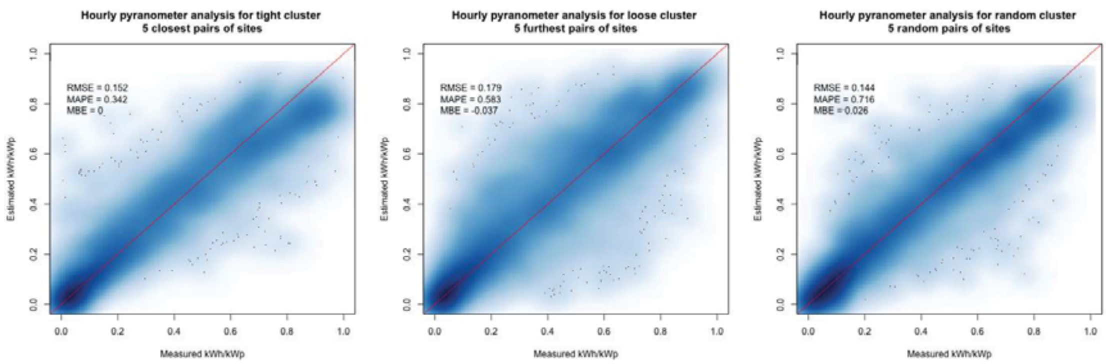

Figure 5: Energy generation estimates at PV sites using a pyranometer to make the estimate [8] for data from 2012-2013, broken down into tight, loose and random clusters, as based on pairwise distance. Shading indicates the density of points, where darker regions are those of greater density.

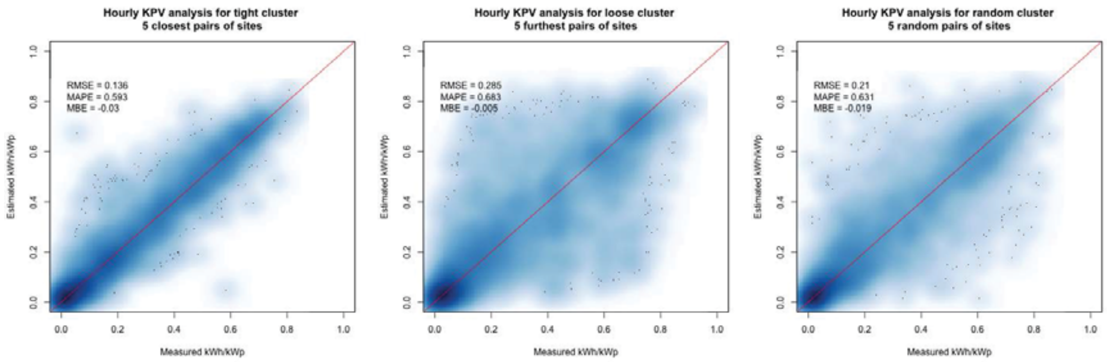

Figure 6: Energy generation estimates at PV sites using another PV system to make the estimate [11] for data from 2012-2013, broken down into tight, loose and random clusters, as based on pairwise distance. Shading indicates the density of points, where darker regions are those of greater density.

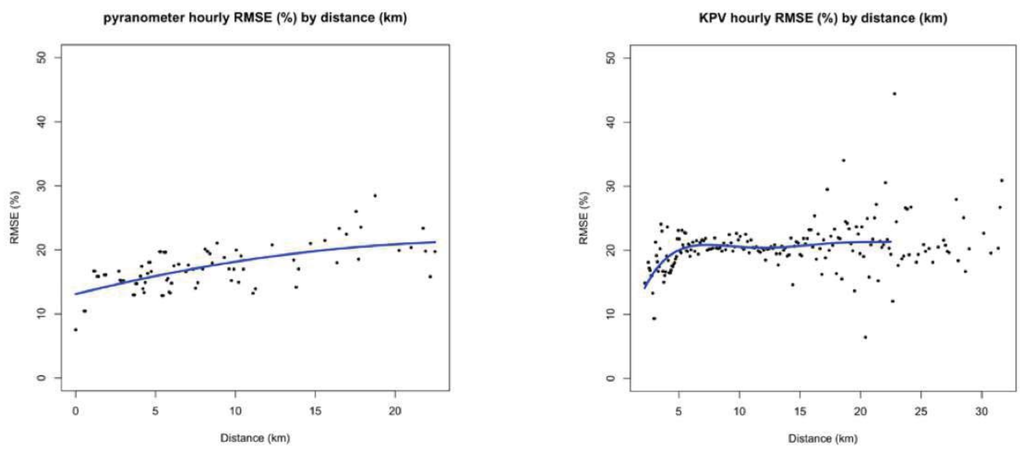

Figure 7: RMSE by distance for the two respective methods. Plots were generated with all of the data from 2012-2013 and breaking the overall error measure down by 0.15 km bins. The blue line indicates a low-order polynomial fit, for the purpose of clarifying the relationship with distance.

Figure 8: Mean Absolute Error and RMSE were generated for zenith angles throughout the day by grouping the energy generation estimates into 5 degree bins and then producing the overall error measures in each bin. The red line indicates the K_{PV} method [11], the blue, the pyranometer base method [8].