Available online at www.sciencedirect.com

ScienceDirect

Solar Energy 105 (2014) 679–693

SOLAR ENERGY

www.elsevier.com/locate/solener

K_{PV}: A clear-sky index for photovoltaics

N.A. Engerer [a,b,*], F.P. Mills [a,c,d]

[a] Fenner School of Environment and Society, The Australian National University, Forestry Building 48, Linnaeus Way, Canberra, ACT 0200, Australia
[b] National ICT Australia, Canberra Research Laboratory, Canberra, Australia
[c] Research School of Physics and Engineering, ANU, Canberra, Australia
[d] Space Science Institute, Boulder, CO, USA

Received 12 August 2013; received in revised form 8 April 2014; accepted 20 April 2014
Available online 21 May 2014

Communicated by: Associate Editor Jan Kleissl

Abstract

The rapidly growing installed base of distributed solar photovoltaic (PV) systems is causing increased interest in forecasting their power output. A key step towards this is accurately estimating the output from a PV system based on the known output from a nearby PV system. However, each PV system is unique with its own hardware configuration, orientation, shading, etc. Thus, the process of using the power output from one system to estimate the power output of another nearby system is not necessarily straightforward. In order to address these challenges, a modified clear-sky index for photovoltaics is proposed. This index is the ratio of the instantaneous PV power output to the instantaneous theoretical clear-sky power output derived from a clear-sky radiation model and PV system simulation routine. This definition performs better than previous clear-sky indices when both PV systems' characteristics are known and the two PV systems have similar orientations. Through this index, the performance of a nearby PV system can be predicted quite accurately. This is demonstrated through the analysis of power output data from five residential PV systems in Canberra, Australia.

Keywords: Solar energy; Photovoltaics; Radiation; Clear-sky index

1. Introduction

The number of photovoltaic (PV) solar energy systems installed worldwide has grown rapidly over the past decade. Current projections forecast global PV system capacity will grow from 102.1 GW in 2012 to 422 GW by 2017 (Gaetan et al., 2013). However, the integration of these systems into their regional electricity grids poses many challenges, as their power output can be highly variable. In Australia, PV systems are primarily rooftop-scale arrays. There are more than 1 million individual systems installed, with an average system size of approximately 2.4 kW (SunWiz, 2013). These distributed generator (DG) sites pose additional integration challenges beyond those experienced by large solar energy generation plants, as most are not actively monitored. This is potentially troublesome for grid operators, as these systems can collectively generate large localized variations in power output and voltage in response to meteorological variability. The degree of variability experienced by collective groups of PV systems has only very recently begun to be investigated (Jamaly et al., 2013; Lave et al., 2013) and is not yet well understood. Thus, in order to identify potential issues and develop solutions for the successful integration of DGs, accurate, low cost, scalable strategies must be developed for generating estimates of current power

* Corresponding author at: Fenner School of Environment and Society, The Australian National University, Forestry Building 48, Linnaeus Way, Canberra, ACT 0200, Australia. Tel.: +61 2 6125 1658.
E-mail address: nicholas.engerer@anu.edu.au (N.A. Engerer).
URL: http://nickengerer.org (N.A. Engerer).

http://dx.doi.org/10.1016/j.solener.2014.04.019
0038-092X/

contributions from these distributed PV systems and forecasting their future energy production.

A preliminary step is developing estimates of collective power output. This is a challenging problem for a number of reasons. First, the small-scale variations in the radiation resource cannot currently be sampled with traditional approaches. The solar radiation monitoring network maintained by the Australia Bureau of Meteorology (BoM) is far too sparse, with several hundred kilometers of separation, on average, between sites. Satellite coverage from the Japanese Meteorological Agency's Multi-functional Transport Satellites (MTSAT) gives better spatial resolution (0.5 deg), but these are still too coarse to measure differences across suburb-level regions and thus insufficient for this purpose. Although the next generation Himawari satellites will deliver 10 min updates of 1 km resolution, that level of coverage is only guaranteed for regions near Japan. In some cases, fine-scale radiation sensor networks have been deployed, such as in Ota City (Lave et al., 2011b), Alice Springs (CSIRO, 2012), UC San Diego (Lave et al., 2011a) and Oahu (Mills and Wiser, 2010). However, these types of networks are too expensive to establish in all locations with large concentrations of PV systems. Second, even with an adequate assessment of the local radiation resource, transposing this field to the surface of each PV system (each having its own unique azimuth and tilt) is difficult and current modeling approaches are not sufficiently accurate (Gueymard, 2009). Third, once the radiation available at a given array is estimated/known, the amount of power generated depends on the system's characteristics, environment and degradation. Finally, the degree of coherence between locations depends on cloud conditions.

One sensor network able to sample PV systems at sufficient resolution with low additional cost is the PV systems themselves. As the number of installed PV systems has risen, the number of PV power output monitoring systems and/or smart meters capable of recording interval level data has also grown. This has led to unprecedented growth in the amount of historical and live PV power output data available for analysis globally. Currently, large amounts of these data are being made publicly available through various websites or can be provided by a central authority (e.g. an electric utility) upon request. For example, pvoutput.org has recorded over 2.2 million measurements from approximately 12,000 sites (http://pvoutput.org, March 2013). Another example is SMA's SunnyPortal, which hosts data from over 109,000 individual sites (http://www.sunnyportal.com, March 2013). These data provide excellent information about the real world performance of solar energy systems and are often used for assessing measured versus expected performance. However, few studies have gone beyond these elementary applications.

One useful extension of this data is to use one PV system's power output to estimate the power output of another, thereby turning monitored PV sites into "sensors" of PV performance (Golnas et al., 2011; Lonij and Jayadevan, 2012; Engerer, 2012). This type of estimation, however, is not a straightforward process. As previously discussed, each PV system has its own unique characteristics, but most of these can be modeled with sufficient accuracy to remove their effects via normalization. In this paper, we introduce a modified clear-sky index for photovoltaics: K_{PV}. Like its predecessors, K_{PV} normalizes a measured value by that PV system's clear-sky output. Data used to test this procedure are described in Section 2. The details of the K_{PV} calculation are presented in Section 3. Results are discussed in Section 4.

1.1. Overview of clear-sky, clearness, and cloudiness indices

The concept of normalizing radiation measurements to their clear-sky potential was first introduced by Black et al. (1954), wherein the ratio of daily radiation measurements to their theoretical values under a perfectly transparent atmosphere was used to develop regression equations for estimating daily radiation from sunshine hour observations. However, the basis for most modern interpretations of this concept are the formulations first suggested in Liu and Jordan (1960). These are K_T, K_d and K_D which respectively represent the measured global, diffuse and beam (direct) radiation received on a horizontal surface divided by the horizontal extraterrestrial radiation. These were first used as normalized coordinates for investigating the relationships among the three measurements in order to provide estimates of the daily diffuse and direct radiation for a given location based on measured global radiation. It was in the course of this investigation that K_T was first termed the "cloudiness index". The usefulness of this index for estimating the performance of tilted flat-plate solar collectors was illustrated shortly thereafter in Liu and Jordan (1963), and since that time the index, in various forms, has remained of fundamental importance in estimating solar energy system performance. It is important to note that the term "clearness index" is often used interchangeably with cloudiness index, as discussed in Duffie and Beckman (2006), and will henceforth be referred to as such herein.

A further evolution of the clearness index concept has become possible with the development of proficient clear-sky radiation modeling (e.g. Bird and Hulstrom, 1981). By modeling the clear-sky radiation arriving at the surface of the Earth, the denominator of the clearness index (previously the extraterrestrial component of irradiance) can be replaced with this clear-sky estimate, thereby changing it to the well-known "clear-sky index" which most modern methods now utilize. The primary usefulness of this clear-sky index is the removal of diurnal and seasonal signals from a given set of radiation data to apply advanced analysis techniques such as wavelets (Woyte et al., 2007), or to compute fluctuation power content (Lave et al., 2011a). This method has been used in more modern assessments of solar variability for solar energy purposes (Lave and Kleissl, 2010; Lave et al., 2011a; Hoff and Perez, 2010)

and as an input to and output from machine-learning-based forecasts of solar radiation (Sfetsos and Coonick, 2000; Yang et al., 2012; Benghanem and Mellit, 2010). It has also become a key variable within solar radiation modeling (e.g. in transposition models (Reindl et al., 1990) or for variability (Skartveit and Olseth, 1992)). It has even been used to classify cloud types (Calbó et al., 2001; Pages et al., 2003), in numerical weather modeling based predictions (Mathiesen and Kleissl, 2011) and in the computation of satellite derived irradiance estimates (Zarzalejo et al., 2009). We hope the new formulation of clear-sky index proposed herein will lead to the replication and extension of many of these methods using the power output from photovoltaic solar energy systems.

1.2. Review of previously utilized methods for PV based power output estimation

The concept of using one PV system or set of PV systems as predictors for the power output of another system or set of systems is not new. Golnas et al. (2011) first introduced the concept of completing performance assessments in the absence of pyranometer data by using PV system power output. They proposed that the long term performance of neighboring PV systems is correlated well enough that one group of systems can be used to predict the power output of another group of systems. They facilitated this via a "Bird Performance Index" (BPI), which is defined as the measured energy generation divided by the plane-of-array irradiance as calculated by the Bird clear-sky model (Bird and Hulstrom, 1981) (it should be noted that the authors do not specify how the transposition from horizontal to tilted surfaces is completed). They apply this method using 15-min interval data from 55 systems (30–500 kW_p) in New Jersey to calculate the daily value of BPI and then use the mean output value from a subset of stations to estimate the BPI value at a given site. By completing this for n x n systems over a training period, they were able to produce very good results for monthly and weekly data with >75% of estimates being within 5% of measured output. For time scales of a day, 2 h and 15 min, the number of estimates within 5% falls from approximately 58% to 24% to 16%, respectively.

Another successful approach can be found in (Lonij et al., 2012; Lonij and Jayadevan, 2012; Lonij et al., 2013) in which the authors demonstrate a method of analyzing system performance for fleets of PV systems (80 PV systems in Tuscon, Arizona). They defined a system's clear-sky performance as the 80th percentile of the system's performance at the same time of day on the previous 15 days. This allowed them to identify outages, partial shading and cloud occurrence within the fleet. They also demonstrate a way to estimate system tilt and orientation with the aid of a simple model for module power output and measurements of global horizontal irradiance. A clearness index (K) was proposed in which the measured power output is divided by the clear-sky power output, as defined above. The clearness index was used for power output forecasts and interpolated spatially to intermediate locations.

The present study differs from these previous studies primarily because it explores a fully modeled approach for the clear-sky reference, as first presented in Engerer (2012). This means historical data are not required to generate clear-sky estimates (as in Lonij and Jayadevan, 2012) and the entire process from solar radiation to PV system power output is simulated (in contrast to Golnas et al. (2011)). Our approach requires detailed information on each PV system but generates a more robust "clear-sky reference," and, thus, more accurate estimates of the power output at nearby PV systems. Further advantages of our new approach are discussed in more detail in Section 3.3.

2. Data

Publicly available PV power output time series were collected for systems in Canberra, Australian Capital Territory, from http://pvoutput.org, a public webpage with user submitted power output data. These data were submitted via commercially available monitoring products (e.g. the SMA 'SunnyWebbox') connected to the inverter system by bluetooth, serial or ethernet connections. Power production from the PV system is averaged over five minute intervals before being time-stamped and reported to the PVOutput.org servers. Users are also give the chance to upload information about their system, including the rating, make and number of modules and inverters, as well as the approximate orientation, tilt and shade conditions of their installation. In order to undertake this study, five of the highest quality stations were hand selected from the available systems in Canberra. These PV systems reported data from early 2011 through 2012, providing an excellent source of real-world power output data. These systems and their respective characteristics are presented in Table 1. In total, the dataset analyzed herein contained over 500,000 measurements.

Additionally, temperature and wind speed data were obtained from the Canberra airport (BoM station 70351). These measurements were incorporated into the performance modeling detailed in Section 3.2 which includes corrections for changes in module performance at different temperatures and wind speeds (i.e. increased temperatures reduce power output).

3. Methods

In order to calculate a clear-sky index for a PV system, its theoretical clear-sky performance must be estimated. This has two primary steps. First, the calculation of the transposed clear-sky radiation available to the PV system, and second, the simulation of the PV system given this available radiation resource. In order to obtain these estimates, a series of models were used. The ESRA clear-sky model (Rigollier et al., 2000) was used to estimate beam,

Table 1
PV system sites used in this study. Information on the system was provided by the owner. Also listed are the module and inverter models from the Sandia performance model database used to simulate their performance.

Site	Module make	Module simulated	Inverter make	Inverter simulated
1	215 W Suntech	215 W SunPower	SMA SB 3000TL	SMA SB 3000TL
2	235 W REC	235 W Suniva Titan	SMA SB 3000TL	SMA SB 3000TL
3	220 W Suntech	220 W Solar Semiconductor	Xantrex GT2.8	Xantrex GT2.8
4	235 W REC	235 W Suniva Titan	SMA SB 2000HF	SMA SWR2100
5	190 W Suntech	190 W Evergreen	SMA SB 3000TL	SMA SB 3000TL
Site	Rating (kW)	# of Modules	Azimuth (0°N)	Tilt (°)
1	2.15	10	-40	22.5
2	3.055	13	-100	20
3	2.42	11	10	20
4	2.115	9	28	20
5	3.04	16	18	28
Site	Latitude	Longitude		
1	−35.37	149.1		
2	−35.39	149.05		
3	−35.45	149.12		
4	−35.42	149.08		
5	−35.44	149.09		

diffuse and global radiation components. The diffuse component was then transposed to a tilted surface via the formulae of Reindl et al. (1990). Finally, the PV and inverter performances were modeled via the Sandia Performance Model, developed by Sandia National Laboratories (King et al., 2004, 2007). It is necessary to note that a wide variety of clear-sky radiation models (Reno et al., 2012), transposition models (Gueymard, 2009) and PV simulation models (Klise and Stein, 2009) are available, and could be used in the K_{PV} calculation to replicate the analyses completed herein.

3.1. Estimated clear-sky radiation input

In order to generate estimates of the clear-sky performance of the selected PV systems, the total (global) amount of clear-sky radiation incident upon a tilted array surface (E_{gtc}) must be determined. E_{gtc} consists of two components: the tilted beam radiation (E_{btc}) and tilted diffuse radiation (E_{dtc}). These are estimated by two radiation models. The ESRA model computes both the clear-sky normal beam component (E_{bnc}) and the horizontal diffuse component (E_{dhc}). E_{btc} is computed based on the system's geometry (3.1.3). E_{dtc} is estimated via the Reindl transposition model.

3.1.1. Clear-sky model selection

We are not aware of any published evaluation of clear-sky models for Australia, so we compared six clear-sky modeling techniques. They were chosen based on their relative simplicity (i.e. not requiring sparse measurements such as the concentrations of particular atmospheric gases, etc.) and good performance in northern hemisphere clear-sky comparisons (Gueymard, 2003; Ineichen, 2006; Reno et al., 2012). Model calculations of E_{bnc} were compared with one minute interval solar radiation data from the Australian Bureau of Meteorology Wagga Wagga station (72150), which is approximately 160 km from Canberra, for 82 mostly clear-sky days in 2011. The selected models are Ineichen (1983); Ineichen and Perez (2002), Bird and Hulstrom (Bird and Hulstrom, 1981), ESRA (Rigollier et al., 2000), Molineaux (Molineaux et al., 1995) and MAC (Davies and McKay, 1982).

Clear-sky periods were identified based on the two conditions provided by Ineichen (2006), that:

$$E_{bn} \geqslant 0.9 \cdot E_{bnc} \tag{1}$$

and that the variability is $\leqslant$10%, where variability is defined as:

$$\triangle E_{bn} = \left(E_{bn}^{n-1} + E_{bn}^{n+1}\right)/E_{bn}^{n} \tag{2}$$

and $n-1$, n, and $n+1$ represent three sequential points in the 1-min data time series.

The results are shown in Fig. 1, which contains each detected clear-sky period. Perfect predictions lie along the identity line, and Mean Bias Error (MBE) and Root Mean Square Error (RMSE) are provided as measures of their accuracy. MBE and RMSE values have been accepted as standards for evaluation of clear-sky solar radiation models (Badescu et al., 2012). Based on our analysis herein, the Bird and ESRA models perform best, with MAPE values of 7.3% and 5.4%, and RMSE values of 71.8 and 60.1 W/m^2, respectively. However, the Bird model requires more input values, including surface pressure, ozone, water vapor, aerosol optical depth and surface albedo. Coincident measurements of these parameters are not available so climatological averages would need to be used. The ESRA model, conversely, requires only an estimate of the Linke turbidity coefficient which can be inferred from beam radiation measurements (Ineichen and Perez, 2002). Given

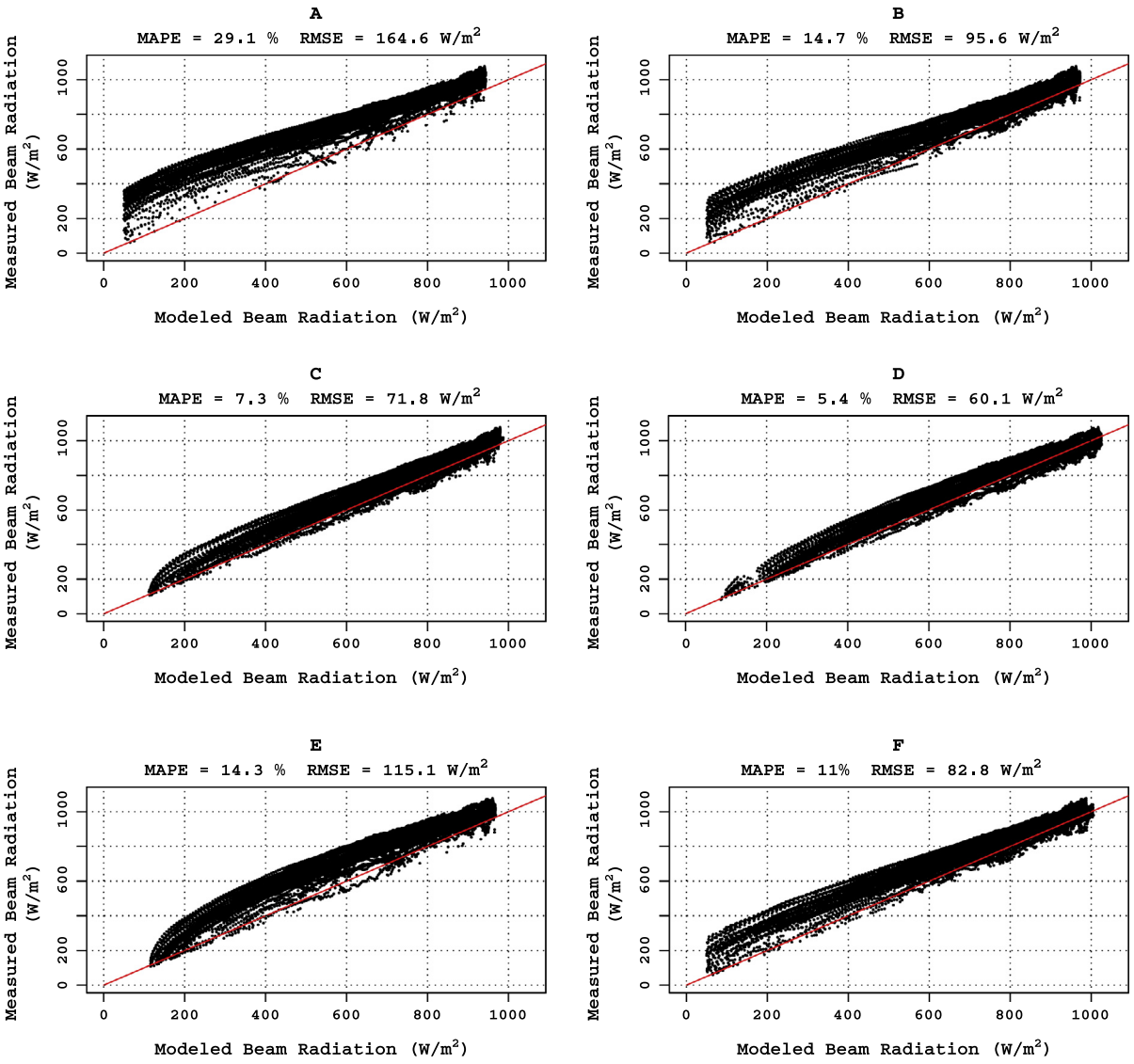

Fig. 1. Measured and modeled beam radiation from 82 mostly clear days at Wagga Wagga for the Ineichen (1983) (A), Ineichen and Perez (2002) (B), Bird and Hulstrom (1981) (C), ESRA (2002) (D), Molineaux et al. (1995) (E) and MAC (1981) (F) clear-sky radiation models. Perfect clear-sky predictions correspond with the red identity line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

their relatively similar performance, the ESRA model has been chosen for its simpler implementation.

3.1.2. Clear-sky radiation

The ESRA clear-sky model calculates normal beam radiation, E_{bnc}, from:

$$E_{bnc} = E_n \cdot exp(-8.662 \cdot T_L \cdot AM_r \cdot \delta_R) \quad (3)$$

where E_n is the normal component of extraterrestrial radiation, T_L is the Linke turbidity coefficient, AM_r is the altitude corrected relative optical air mass as defined by (Kasten, 1988) and δ_R is the path integrated Rayleigh optical thickness.

The horizontal diffuse component of radiation (E_{dhc}) is computed according to:

$$E_{dhc} = E_h \cdot T_R(T_L) \cdot F_D(\theta_z, AM_r) \quad (4)$$

where E_h is the horizontal component of extraterrestrial radiation, T_R is the diffuse transmissivity as a function of Linke Turbidity and F_D is the diffuse angular function as a function of air mass and solar zenith angle, θ_z.

The Linke turbidity coefficient, T_L, is obtained by fitting a curve to the monthly values retrieved from the SoDa dataset (Rigollier et al., 2001). The fitted polynomial is provided in Fig. 2.

3.1.3. Radiation available to the PV array

The two components of the radiation incident on a tilted PV system are the titled beam and tilted diffuse components. The first is computed as Duffie and Beckman (2006):

$$E_{btc} = E_{bnc} \cdot \cos(\theta_a) \quad (5)$$

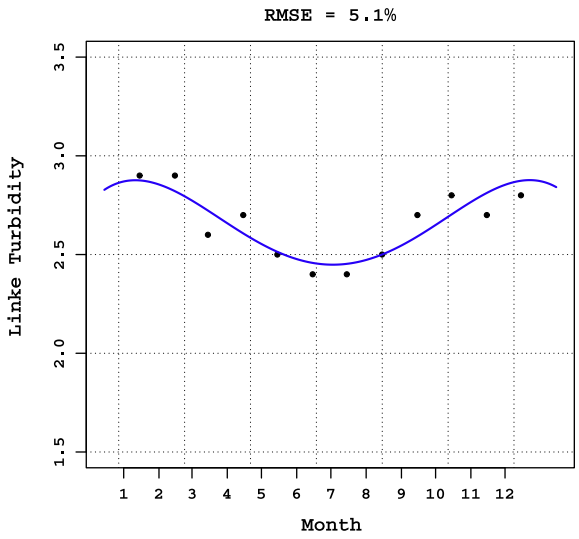

Fig. 2. Monthly values for Linke turbidity factor at Canberra. Points represent values retrieved from the SoDa database (Rigollier et al., 2000), with the line representing the corresponding fit used to obtain daily values.

with θ_a representing the angle of incidence for a given tilted surface, determined by its tilt (α_m) and azimuth orientation (ϕ_m) by:

$$\cos(\theta_a) = \cos(\alpha_m)\cdot\cos(\theta_z) + \cos(\phi - \phi_m)\cdot\sin(\alpha_m)\cdot\sin(\theta_z) \tag{6}$$

with ϕ representing the solar azimuth angle and θ_z the solar zenith angle.

Modeling the amount of diffuse radiation available to a tilted surface (transposition) is the subject of much discussion within the literature as the determination is not straightforward (Loutzenhiser et al., 2007; Hay and McKay, 1988; Kambezidis et al., 1994; Gueymard, 2009). Most debate focuses on correctly attributing diffuse radiation to clouds. As we are only considering clear-sky conditions, our choice is much more straightforward.

We based our selection of a diffuse radiation transposition model on the findings of Gueymard (2009), who compared the performance of 10 transposition models. The study found that the Reindl transposition model performed quite well under clear-sky conditions, and particularly so at lower sun angles as exhibited by comparing dual-axis tracking surface predictions with observation. For these reasons, and for its computational simplicity, we chose the Reindl model:

$$E_{dtc} = E_{dhc} \cdot \left\{\left[(1 - A_i)\cdot(1 + \alpha_m/2)\right]\cdot\left[1 + f\cdot\sin^3(\alpha_m/2)\right] + A_i\cdot R_b\right\} \tag{7}$$

where A_i is the anisotropy index defined as $A_i = E_{bn}/E_n$ and R_b is the geometric factor $R_b = cos(\theta_a)/cos(\theta_z)$.

The global radiation resource incident on a PV array on a clear day is computed:

$$E_{gtc} = E_{btc} + E_{dtc} \tag{8}$$

3.2. System performance modeling

In order to provide an estimate of the clear-sky power output of a given PV array, both the Photovoltaic Array Performance Model (PAPM) (King et al., 2004) and the Inverter Performance Model (IPM) (King et al., 2007) developed by Sandia National Laboratories (SNL) were forward integrated over the daily clear-sky curves generated as outlined in Section 3.1. The PAPM relies on empirically derived coefficients (32 for each module) obtained during field testing of individual modules. The IPM is similar and was developed under field testing of grid-tied inverters connected to appropriately sized PV arrays, and contains 12 empirically derived coefficients. Studies which have tested the accuracy of these models against measured power output have absolute errors within 5% (Fanney et al., 2006) and relative errors as low as ±3% (Cameron et al., 2008; Cameron et al., 2011). The success of these models has led to their incorporation in the System Advisor Model (SAM) (Gilman et al., 2008) developed jointly by the U.S. National Renewable Energy Laboratory and SNL. SAM is now being further developed as a part of The Photovoltaic Performance Modeling Collaborative (PVPMC) (Stein, 2012).

3.2.1. Photovoltaic array performance model

For each of the systems in the PVOutput.org dataset, the owner provides information about the brand and rating in watts of the PV modules used in their installation. Using this information, modules at each site were matched with the most similar module available in the Sandia Module Database (available within distributions of SAM). If an exact match was not available, the most similar was selected based on material type (e.g. mono vs. poly-crystalline), peak power rating, short circuit current and open circuit voltage. The matches chosen are provided in Table 1. Power output for a single module was estimated using the maximum power formulae (King et al., 2004):

$$P_{mp} = I_{mp}\cdot V_{mp} \tag{9}$$

$$I_{mp} = I_{mp0}\cdot(C_0\cdot E_e + C_1\cdot E_e^2)\cdot\left[1 + \alpha_{Imp}\cdot(T_c - T_0)\right] \tag{10}$$

$$V_{mp} = V_{mp0} + C_2\cdot N_s\cdot\delta(T_c)\cdot\ln(E_e) + C_3\cdot N_s\cdot\left[\delta(T_c)\cdot\ln(E_e)\right]^2 + \beta_{Vmp}(E_e)\cdot(T_c - T_0) \tag{11}$$

Here, E_e is the solar irradiance available to the module which includes module specific corrections to the radiation input. Once modified for our clear-sky case it is:

$$E_e = f_1(AM_a)\cdot\left[E_{btc}\cdot f_2(\theta_a) + f_d\cdot E_{dtc}\right]/1000 \tag{12}$$

where f_1 corrects for variations in the solar spectrum as a function of the air mass and f_2 accounts for optical influences as a function of θ_a. Their formulations can be found in (King et al., 2004).

Array level maximum power P_{mpA} is computed according to

$$P_{mpA} = (V_{mp}\cdot M_s)\cdot(I_{mp}\cdot M_p) \tag{13}$$

where M_s is the number of modules in series and M_p is the number of strings in parallel. For further details, the reader is referred to King et al. (2004).

3.2.2. Inverter performance model

Users of pvoutput.org also submit information about their inverter system, including the brand and rating. As in the previous section, these inverters were matched with the most similar entry in the Inverter Module Database (also included with SAM distributions) based on brand, rating and maximum allowable power. These matches are provided in Table 1. Direct current array peak power (P_{mpA}) is converted to alternating current power, less inverter level losses, by:

$$P_{ac} = [P_{ac0}/(A-B) - C\cdot(A-B)]\cdot(P_{mpA}-B) + C\cdot(P_{mpA}-B)^2 \quad (14)$$

where A, B and C are functions of DC voltage as defined in (King et al., 2007) and P_{ac0} is the inverter's rated capacity.

3.3. The clear-sky index for PV

The clear-sky index for PV is defined as:

$$K_{PV} = \frac{PV_{MEAS}}{PV_{CLR}} \quad (15)$$

where PV_{MEAS} is the measured power output from a system and PV_{CLR} is the simulated clear-sky power output.

3.3.1. Demonstrating the K_{PV} calculation

Figs. 3 and 4 provide examples of K_{PV} calculations for clear and partly cloudy days, respectively, in September 2011 for the five PV sites in Canberra. In each figure, first PV_{MEAS} is plotted, followed by PV_{CLR} with the ratio of the two (K_{PV}) in the third. During clear-sky periods, the K_{PV} values are approximately one, as is expected, with small variations occurring due to system level variations, possibly from inefficiencies in the modules, inverters or wiring (see Fig. 5).

Most differences among the station power output in Fig. 3 are due to orientation and capacity. Stations 1 and 2 have northwest and southwest facing azimuths, while stations 3, 4 and 5 have northeasterly azimuths. These differences affect the timing of the peak power output and result in significant differences in power output relative to capacity. For example, stations 3–5 produce more power in the first half of the day as compared to the westerly facing stations, with the opposite true in the afternoon/evening. However, in the K_{PV} calculation the effects of orientation, tilt and system size are removed.

3.3.2. Shading detection

Completely clear-sky conditions, as depicted in Fig. 3, provide a unique opportunity to identify the sources of deviations from theoretical clear-sky performance that are not due to clouds. The most obvious example is the sharp drop in power production at higher zenith angles for stations 3 and 4 in the morning, and station 5 in the evening. These occurrences of shading (surrounding obstructions were observed to be present in local aerial imagery) are not as apparent in the power output time series as they are in the K_{PV} plot. Furthermore, upon closer inspection, the K_{PV} plot reveals much more subtle and gradual instances of shading present for all of the given stations. This is shown by the gradual downward trends in K_{PV} values in the early morning and late evening periods as increasing amounts of diffuse and direct sunlight are obscured by, for example, topography and vegetation.

Using an approach similar to that of Reno et al. (2012), we were able to detect instances of shading present within

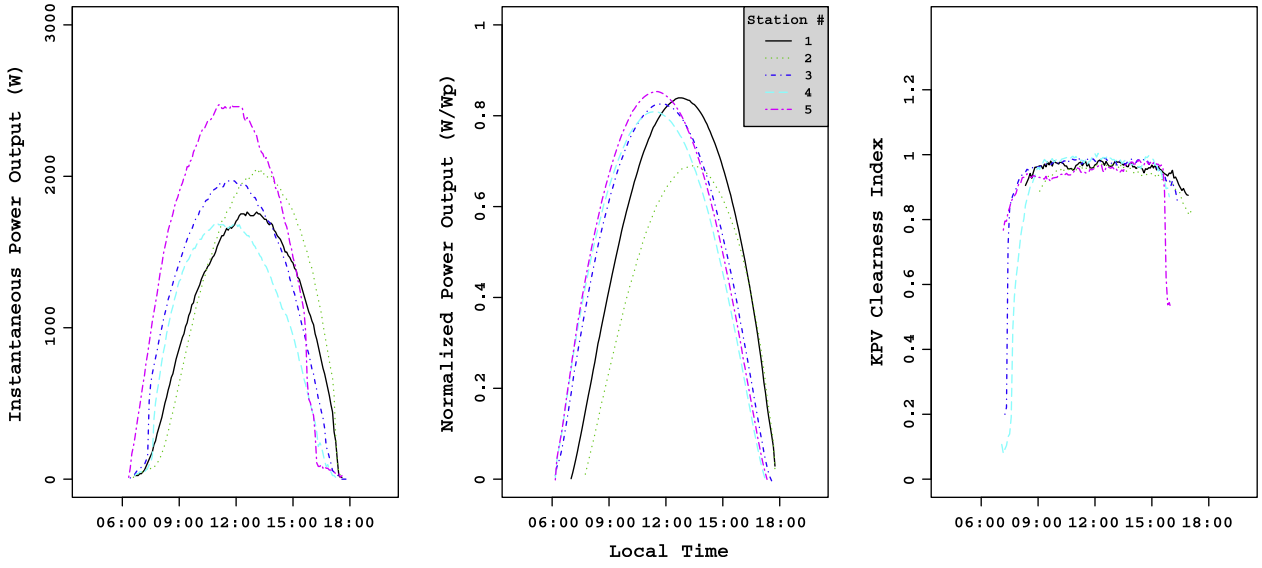

Fig. 3. Clear-sky K_{PV} calculations. From left to right, time series of five-minute average PV system power output, the simulated normalized clear-sky performance and the clear-sky index K_{PV}, are presented for the 5 systems listed in Table 1 on 18 September 2011; which was a clear day.

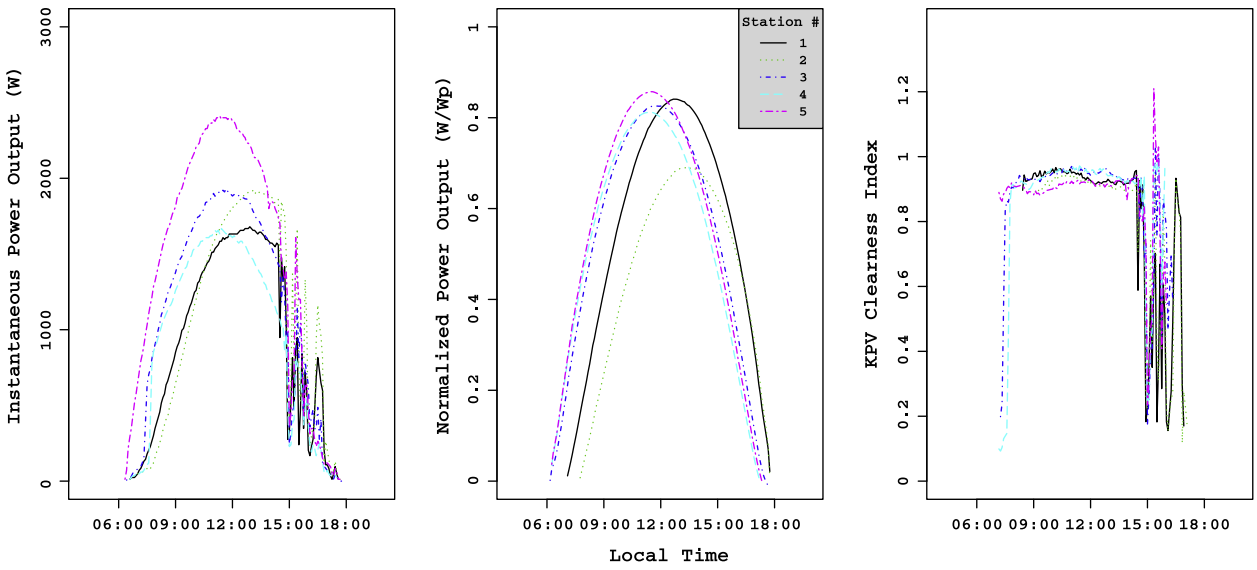

Fig. 4. Partly cloudy K_{PV} calculations. From left to right, time series of five-minute average PV system power output, the simulated normalized clear-sky performance and the clear-sky index K_{PV}, are presented for the 5 systems listed in Table 1 for a day with fair weather cumulus clouds in the afternoon, 17 September 2011.

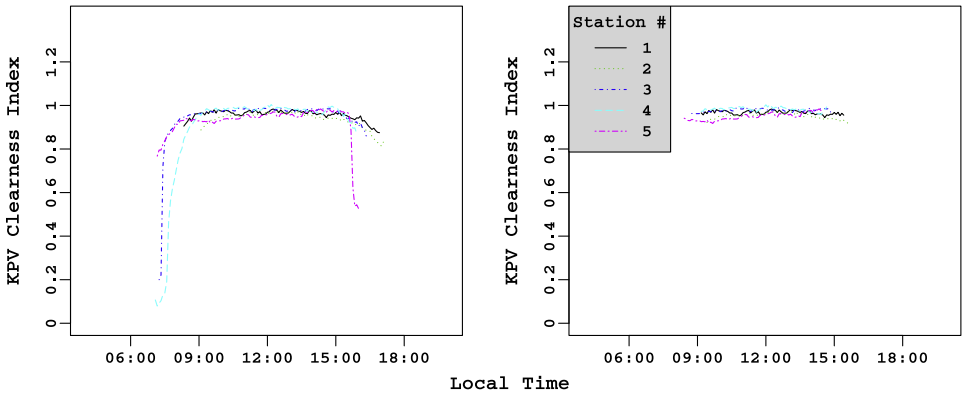

Fig. 5. Clear-sky K_{PV} calculations from 18 September 2011 (as presented in Fig. 3), with the original calculation on the left, and with data experiencing shading removed on the right.

time series of the clearest days. Periods of shading were detected by identifying portions of the time series which had variability exceeding 10% of the clear-sky curve, and by identifying portions of the time series for which values of the slope of power output differed from the clear-sky curve by more than 5%, and finally, by eliminating any K_{PV} value less than 0.8. In order to be truly considered an instance of shading, detected events were required to display systematic behavior over multiple clear days; meaning that they must occur at approximately the same solar hour angle, so as to prove they were not caused by isolated clouds. By recording which solar hour angles were known to experience shading at a particular site during a selection of the clearest days spanning an entire year, one is able to remove all periods of shading from the dataset by excluding from analysis any portions of the time series known to occur at those hour angles. In this manner, instances of shading were removed from the analysis as shown in Fig. 6. In future studies, a corrective model will be developed and employed for these instances of shading.

3.3.3. Estimating power output of another PV system

With the instances of shading removed, it is now possible to use the power output time series of one system to estimate the performance of another nearby PV system. Since the K_{PV} value is independent of system size and orientation, as long as the clear-sky curve can be generated for the nearby system (which requires the approximate location, system size and orientation), an estimate of temporally coincident power output of the nearby system (denoted as number 2) is calculated by:

$$PV_{EST_2} = \frac{PV_{MEAS_1}}{PV_{CLR_1}} \cdot PV_{CLR_2} = K_{PV_1} \cdot PV_{CLR_2} \tag{16}$$

where PV_{CLR_2} represents the clear-sky power output estimate at the nearby station at the given time. The accuracy

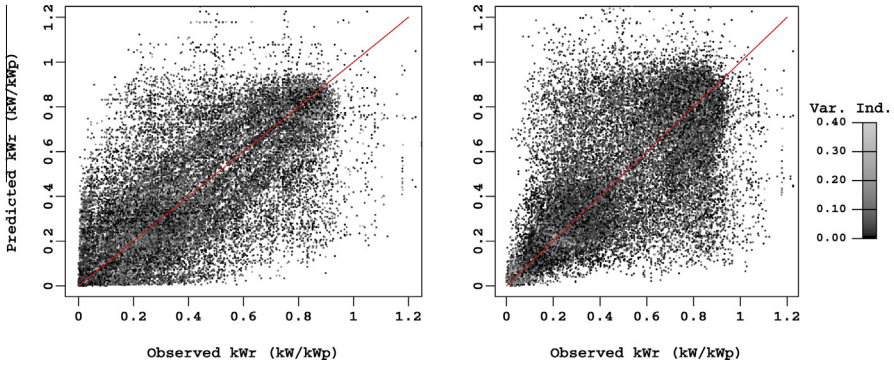

Fig. 6. Using 5-min average power output data from the first 6 months of 2012, estimates of PV system performance are plotted against observed values as computed for all sites. At left, the estimates were generated using the performance ratio. At right, K_{PV} based estimates are provided.

of this calculation is dependent on distance, cloud cover, cloud enhancement and system specific issues such as the presence of shading, soiling or system level inefficiencies. This is discussed further in Section 4.

4. Results and discussion

Using power output data collected from January 2011 through December 2012, K_{PV} based estimates of PV system performance were generated for each of the five sites using the remaining four sites (e.g. site 1 used to estimate sites 2–5, site 2 used to estimate sites 1, 3–5, etc.). The overall results from this process are presented in Tables 3,4, as broken down for each site. RMSE and MAPE values range from as low as 3.14% and 3.51% and as high as 9.77% and 10.8%, respectively.

Through a more thorough analysis of these calculations, we can assess the ability of the K_{PV} approach to accurately estimate the power output of a nearby PV system. To clarify, the terminology 'nearby' is limited to the present dataset. These sites have a mean separation distance of 4.75 km with no two sites being separated by more than 8.4 km. The accuracy of this method beyond these distances remains uncertain and will be the focus of future research. We will also attempt to estimate the performance of the K_{PV} approach under various sky cover conditions by using K_{PV} values as a proxy for cloud cover.

4.1. K_{PV} versus performance ratio based estimates

For the five sites, K_{PV} based estimates are compared to those based on a site's performance ratio (kW_r), which is the observed power output divided by the rated capacity of the system. Using a performance ratio based estimate could be considered the current 'standard approach' for estimating the power output at one site using the power output from another. A comparison of these two approaches as computed for the period of March to September 2012 is presented in Fig. 7. The points are color-coded according to the Variability Index (VI), which we have adapted from Davy et al. (2010) who first proposed it for wind power time series. They proposed a running standard deviation computed over 6 h time windows as an assessment of power output fluctuations from wind turbines. Here, we have modified the time window to be 30 min in order to capture the finer fluctuations that occur with scattered clouds. The shading is relative to the site from which the prediction is being made. Darker regions indicate lower VI values, lighter regions, increased values of VI. The kW_r based estimates reveal a much weaker correlation between observed and predicted power output. This is largely because the kW_r based estimates do not consider the orientation of the arrays. In contrast to this, the K_{PV} based estimates display a stronger linear relationship. There are three distinct groupings in the K_{PV} based scatter plot. The first are those clustered around high kW_r values, which tend to have reduced variability (darker). This clustering with low VI values shows the accuracy of the method under clear to mostly clear skies. The second is the kW_r values below 0.4, where again, a strong linear relationship is present with lower VI values. This grouping indicates the accuracy of the methodology under uniform, opaque cloud cover conditions. Lastly, the weakness of the correlation at mid level kW_r values (approximately 0.5–0.7) shows, unsurprisingly, that under non-uniform, partly cloudy conditions, the predictions are less certain. Here, the dark points that lie along the identity line indicate periods in which there are likely thin, more uniform cloud cover that covers both sites. Conversely, points where the VI value is larger tend to deviate away from the identity line. This is because the cloud cover conditions are different between the sites.

4.2. K_{PV} based estimates and cloud cover

The main cause of errors in K_{PV} based estimates are non-uniform cloud conditions. The errors resulting from this non-uniformity can be explored by using the K_{PV} value as a proxy for cloud cover, as is done with other formulations of the clear-sky index. This allows for the current cloud conditions to be inferred from the instantaneous and averaged K_{PV} values, respectively.

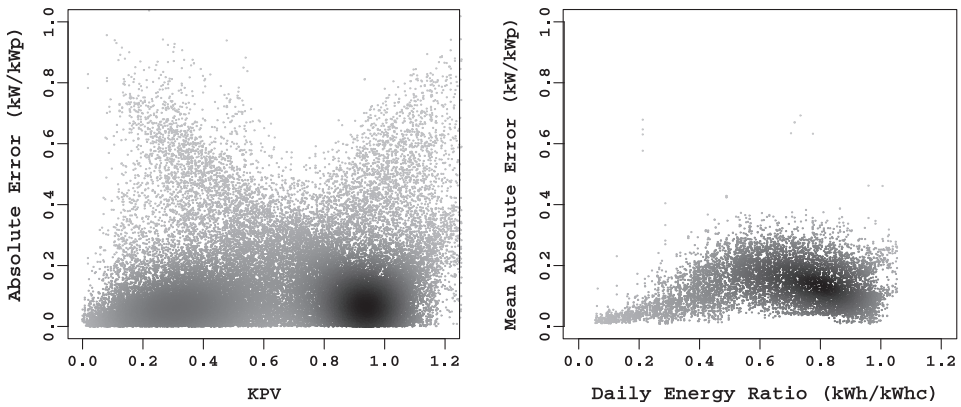

Fig. 7. In plot A, Absolute Error (kW/kW_p) is plotted against the values of K_{PV} for single site, single time-step estimates from 1st January – 31st March 2012. In B, daily Mean Absolute Error versus the Daily Energy Ratio (kWh/kWh_c). These are calculated by using each site to estimate the remaining four sites. The x-axis is indicative of the K_{PV} or DER value at the site making the estimate. Shading is also present in the Figure, with darker areas indicative of a higher density of points.

Fig. 8 shows the absolute error of K_{PV} based estimates between sites in terms of two proxies for cloud cover. The first is the instantaneous K_{PV} value and the second is the ratio of the total number of kilowatt-hours generated throughout a given day versus those that would have been generated on a clear day. This daily formulation is comparable to the daily clear-sky index first proposed by Liu and Jordan (1963), or the daily generation potential discussed in McKenney et al. (2008). Herein it is entitled the 'Daily Energy Ratio' (DER).

In Fig. 8a, the Absolute Error (AE) is plotted against the instantaneous values of K_{PV}. There is a strong relationship between higher K_{PV} values and reduced AE, with the exception of the scattering of points above $K_{PV} \approx 1$, which can be attributed to mostly sunny days with scattered clouds where clear conditions at one site are occasionally not representative of another nearby system. As K_{PV} values fall, the AE tends to rise. The strong linear trend in AE versus K_{PV} values is believed to be attributable to partly cloudy to mostly cloudy conditions where there is increasing cloud opacity, but within which breaks in the clouds remain. This results in large errors in the K_{PV} based estimations. As the breaks in clouds close with thicker cloud decks, the errors in K_{PV} tend to fall, but have a large spread attributable to non-uniform levels of cloud opacity.

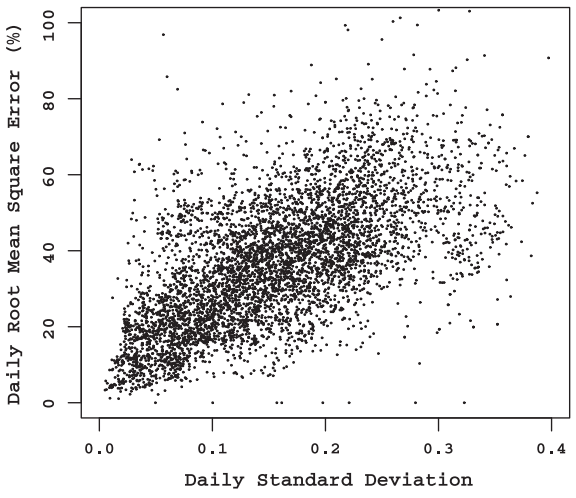

Fig. 8. Daily RMSE (%) for KPV based predictions (each site to every other site) plotted against the daily standard deviation as measured at the site producing the estimate as computed for data from 2012.

Fig. 8b provides the Mean Absolute Error (MAE) plotted against the Daily Energy Ratio (DER). The DER allows for assessing the performance of the K_{PV} based predictions over the course of an entire day in terms of its overall cloudiness. Low DER (<0.4) values are indicative of persistent opaque cloud cover, high values (>0.8), of mostly clear conditions, with the inner range of DER values providing slightly less information. Mid-range (0.4–0.8) DER values could arise from days with equal amounts of partial cloud cover through the day, or, for example, from days with strong ramp events. The MAE values in Fig. 8b shows an increase in error with increasing DER reaching a peak at around a value of 0.6, whereupon MAE begins to fall as overall cloud cover decreases and clear-sky conditions return.

It is also helpful to analyze the daily RMSE in terms of the variability experienced at a given site. Fig. 9 presents the daily RMSE as plotted against the observed standard deviation at the site creating the estimate. An strong linear, upward trend is present in the RMSE values with increasing overall daily variability (as represented by the standard deviation). The spread in the data also displays an increase with increasing standard deviation. These two observations reveal that with increasing variability in the time series (increasing variability in cloud cover) there is an overall reduction in the accuracy of the method and increase the uncertainty of the estimates. It is likely that averaging multiple sites to produce an estimate at a single site will reduce these issues, which we will explore in future work with this method.

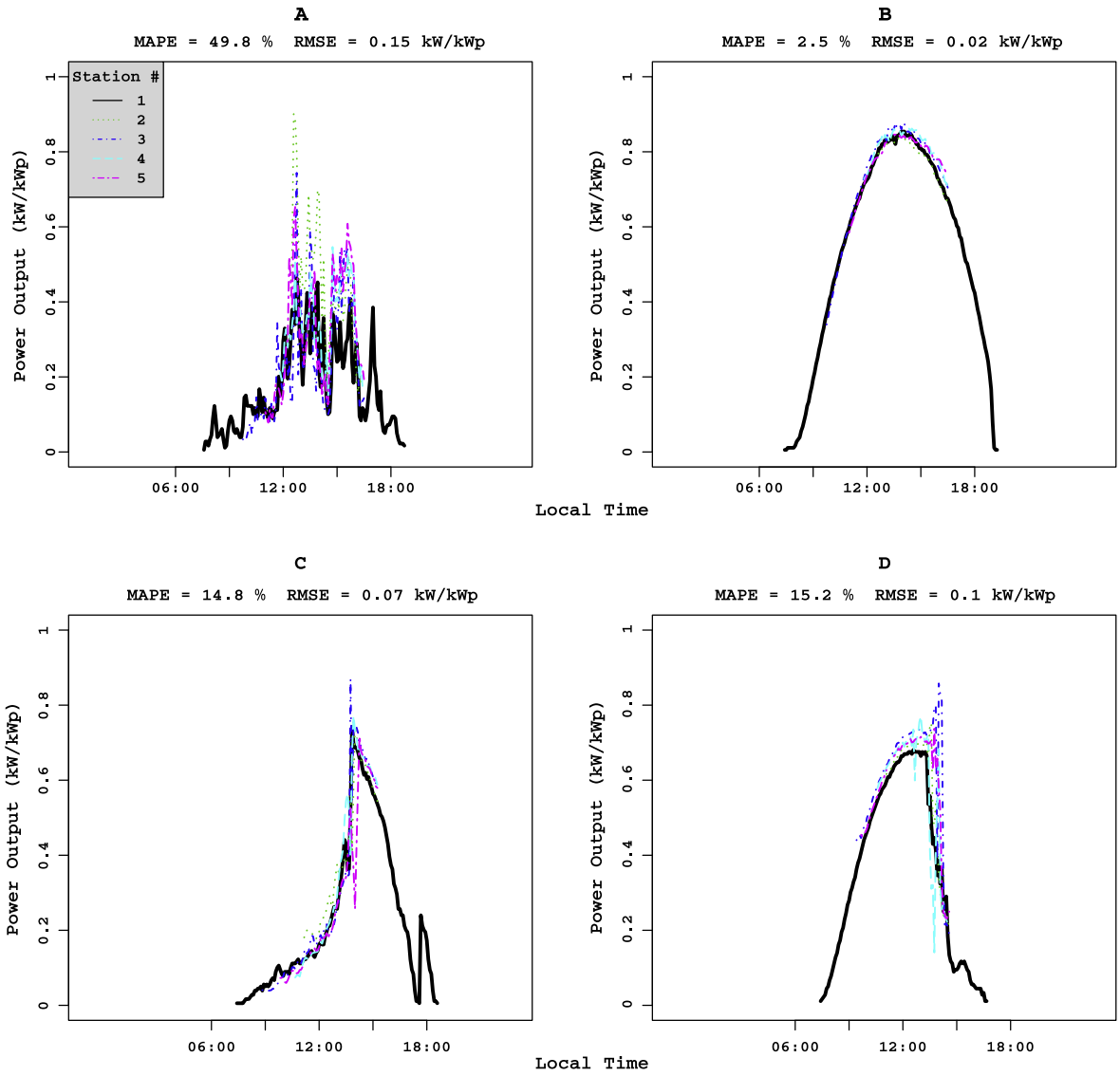

Fig. 9. K_{PV} based predictions for station 1 as generated from stations 2–5 are compared with the actual power output from station 1. Plot A displays a highly variable case of non-uniform opaque cloud cover from 27 February 2012. Plot B depicts a very stable day with no cloud cover from 24 February 2012. In C, a positive ramp event from 4 May 2012. In D, a negative ramp event from 1 April 2012.

4.3. K_{PV} based estimates under various conditions

Specific cases also have been analyzed to provide insight into the more nuanced behavior of the K_{PV} based predictions. In Fig. 10, four different cases are presented to show the performance of the K_{PV} based estimates relative to highly variable, mostly stable and suddenly changing conditions in the local radiation resource. For each of the four plots provided in Fig. 10, K_{PV} based estimates of the power output at station 1 were calculated by using stations 2–5.

4.3.1. Highly variable opaque cloud cover

In Fig. 10a, highly variable conditions were present, with nearly uniform cloud cover of highly non-uniform opacity. There are significant disagreements between the timing and magnitude of sudden swings in power output, resulting in a fairly high MAPE of 49.8%. This is mostly driven by a few large spikes in estimated power output that were not observed at station 1. Accordingly, inter-site RMSE was only 0.15 kW/kW_p.

4.3.2. Clear-sky, stable conditions

In Fig. 10b a nearly opposite situation is present, with clear, highly stable conditions from 24 February 2012. The estimates of power output perform very well, with a very small MAPE of 2.5% and an RMSE of 0.02 kW/kW_p. This is likely attributable to system level inefficiencies that are not captured accurately by the modeling framework. These could include soiling of the panels, poor wiring or thermal inefficiencies that were inaccurately handled by the performance model. Interestingly, these errors are larger in the afternoon period than in the morning, where the observed temperature in Canberra climbed to 30 degrees Celsius at the Canberra Airport. It should also

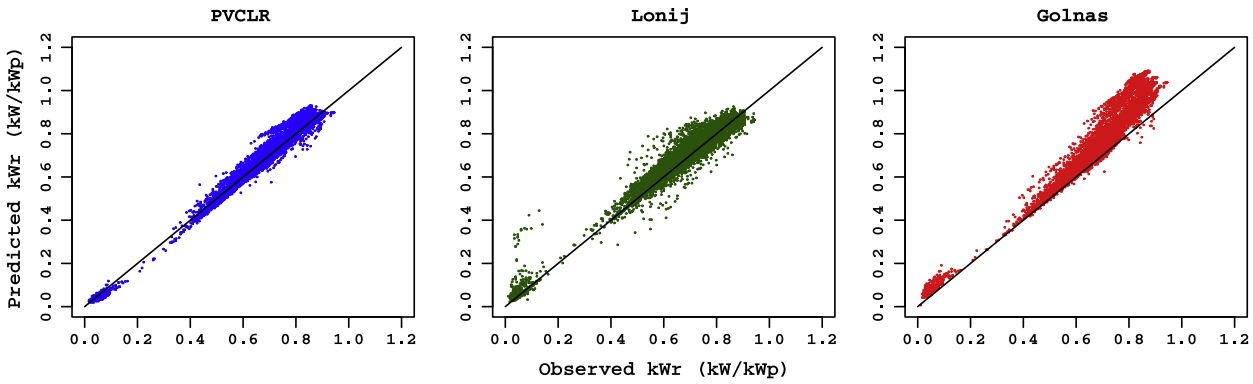

Fig. 10. PV system clear-sky power output estimates as produced by the PV_{CLR} method from the present study, Lonij (Lonij et al., 2012) and Golnas (Golnas et al., 2011). Estimates were computed using clear-sky periods extracted from 80 mostly clear days. Overall error metrics are reported in Table 2.

Table 2
Mean absolute percent error (%) as calculated for the 2011–2012 period. Table is presented by displaying the sites being estimated presented by the site being used to make the estimation.

Site used	Site estimated				
	1	2	3	4	5
1	–	9.77	8.55	4.79	7.0
2	9.20	–	3.63	7.49	4.72
3	6.84	3.74	–	5.76	3.14
4	4.76	9.05	6.53	–	5.12
5	8.24	4.68	3.45	3.86	–

Table 3
Root Mean Square Error (%) as calculated for the 2011–2012 period. Table is presented by displaying the sites being estimated presented by the site being used to make the estimation.

Site used	Site estimated				
	1	2	3	4	5
1	–	10.4	9.55	5.53	9.96
2	10.8	–	3.87	8.57	5.67
3	7.19	3.68	–	6.91	4.49
4	9.71	10.2	7.38	–	5.44
5	9.88	5.36	4.63	3.51	–

be noted that these errors are even lower than the overall errors as computed in the clear-sky estimates from Table 2 (MAPE of 4.1% and RMSE of 4.06%).

4.3.3. Ramp events

Within solar energy research, sudden changes in power output from a PV system are referred to as ramp events (Lave and Kleissl, 2010). In Fig. 10c and d, positive and negative ramp events are presented. The positive ramp event took place on 4 May 2012, wherein an overcast day saw a gradual decrease in cloud cover give way to mostly clear skies by mid afternoon. The uniformity of the cloud cover is apparent from the strong agreement among the stations in both the magnitude and timing of the ramp event. Conversely, in 10d, a negative ramp event from 1 April 2012 is presented. The sudden arrival of a uniform opaque cloud deck is apparent from the agreement in the magnitude and timing of the K_{PV} based predictions, which is only slightly less accurate than that in 10c, as shown by the slight increase in RMSE from 0.07 to 0.10 kW/kW_p. This is mostly due to the small disagreements in the timing of the event.

4.4. Comparison with previous methods for PV-based power output estimation

The Bird Performance Index (BPI) (Golnas et al., 2011) is defined as

$$BPI = \frac{PV_{MEAS-G}}{E_{gtc}/1000} \tag{17}$$

where PV_{MEAS-G} is the measured 15-min energy value for each system, E_{gtc} is the 15-min clear-sky plane-of-array irradiance estimated using the Bird Clear Sky Model (Bird and Hulstrom, 1981), and the factor of 1000 converts the irradiance to suns. After removing days with system outages, the clear-sky BPI (CSBPI) for each system is defined as the 93rd percentile of the set of BPI values in the training set.

The clearness index of Lonij et al. (2012) is defined as

$$K = \frac{E_{gt}}{E_{gtc}} = \frac{PV_{MEAS-L}}{PV_{CLR-L}} \tag{18}$$

where E_{gt} is the plane-of-array irradiance, E_{gtc} is the clear-sky plane-of-array irradiance, PV_{MEAS-L} is the measured power output normalized by a system's peak power, and PV_{CLR-L} is its clear-sky normalized power output. PV_{CLR-L} is defined as the 80th percentile of PV_{MEAS-L} at the specified time of day on the previous 15 days for a system Lonij et al. (2012). System-specific factors, like shading, were then removed by taking the 80th percentile of an ensemble of 80 systems spread across a $50 \times 50\ km^2$ region Lonij et al. (2012).

The numerators in BPI and K are equivalent to that in K_{PV}: PV_{MEAS-G} is energy while PV_{MEAS} is power and PV_{MEAS-L} is normalized by peak power while PV_{MEAS} is not. The differences among these three indices lie in the

denominators. The denominator in BPI ($E_{gtc}/1000$) does not include the performance of the PV system, although Golnas et al. (2011) addressed the impact of this on system-to-system estimations via the Conversion Profile, CP_{ij}:

$$CP_{ij} = \frac{CSBPI_i}{CSBPI_j} \tag{19}$$

for systems i and j. The $CSBPI$ also can be adversely affected by persistent cloud cover in the training data and by shading. Multiple years of training data may potentially be required for all sites in order to compute CP_{ij} for a full year. The denominator in K, PV_{CLR-L} implicitly includes the performance of the PV system, and K implicitly includes shading in both the numerator and denominator. K, thus, is a more accurate and robust estimator for system-to-system estimations than BPI. However, K can be adversely affected by persistent regional cloud cover in the training data, and system-to-system estimations require data for all sites to calculate PV_{CLR-L}. The denominator in K_{PV}, PV_{CLR}, is entirely derived from model simulations. System performance factors and environmental conditions are explicitly included in the simulations. Secular degradation in system performance can be included, and historical data are only required for system-to-system estimations in which the performance of at least one system is significantly different from nominal. PV_{CLR} is not affected by cloud cover over a site. Consequently, K_{PV} is a more robust estimator for system-to-system estimations than either BPI or K if the characteristics of all systems are known.

Given the equivalence of the numerators in BPI, K, and K_{PV}, the relative accuracy of these indices will be determined by the accuracy of the denominators, i.e., their ability to generate accurate clear-sky estimates. In order to compare these three methods for generating the clear-sky estimates, we hand-selected 80 clear-sky days in Canberra and extracted the clear-sky periods from them via the methodology of Reno et al. (2012) and use each method to generate clear-sky estimates. The results are compared in Fig. 10 with a tabulation of four error metrics (MBE, MAPE, RMSE, and Coefficient of Determination (COD)) in Table 4. The PV_{CLR} approach outlined in the present study demonstrates superior performance in all error measures. Of particular interest is the over-prediction of Golnas et al. (2011) under increasingly bright sunshine, which may be due to the lack of system level simulations (e.g., module losses at higher temperatures). The Lonij approach performs better than the Golnas approach and does not require knowledge of PV system characteristics, but it requires historical data and can be adversely affected by clouds, shading, outages, etc., during the historical reference period. The PV_{CLR} method can be applied to systems for which no data has been recorded and can give the most accurate estimates of clear-sky irradiance, Table 4, when PV system characteristics are known.

The similarity of the three indices – BPI, K, and K_{PV} – means their properties and utility are similar for spatial interpolation, solar forecasting, cloud motion studies, ramp rate analyses, and comparative generation estimates. Consequently, the advanced analysis techniques and multi-site estimation procedures discussed in Golnas et al. (2011); Lonij and Jayadevan (2012); Lonij et al. (2012, 2013) should work at least as well with K_{PV}. In fact, one expects the enhanced accuracy of the clear sky term in K_{PV}, Fig. 10 and Table 2, should improve the performance of those analysis techniques.

Table 4
Error measures as reported after evaluating three methods for generated PV system clear-sky power output: PV_{CLR} from the present study and Lonij (Lonij et al., 2012) and Golnas (Golnas et al., 2011). Estimates were computed using clear-sky periods extracted from 80 mostly clear days and are also plotted in Fig. 3.

	MBE (%)	MAPE (%)	RMSE (%)	COD
PV_{CLR}	1.09	4.1	4.06	0.99
Lonij	2.58	8.1	6.15	0.98
Golnas	10.5	12.3	12.3	0.96

The limitations of the three indices also are similar. In particular, all three rely on global plane-of-array irradiance (either simulated or measured), so there is an implicit requirement that all PV systems being considered should have reasonably similar orientations. If the orientations of two systems are not sufficiently similar, then it becomes necessary to consider separately the contributions of diffuse and direct irradiance to power generation under cloudy conditions. This cannot be done with BPI or K. K_{PV}, however, is based on model simulations of the radiation field and the PV system response, so it conceivably can be modified to provide this separation. This will be examined in future work.

5. Conclusion

Through the use of clear-sky radiation modeling, transposing the radiation components to a tilted surface and using the transposed modeled radiation as input to PV and inverter modeling routines, the theoretical clear-sky power output for a given PV system can be calculated, provided the characteristics of the PV system are known. Normalizing the measured power output from a PV system by this theoretical value, gives a PV clear-sky index, K_{PV}, which can be used to estimate the power output of nearby PV systems if their system characteristics are known. A single PV system power output time series becomes quite useful for estimating the performance of other nearby PV systems. The accuracy of these estimates are strongly related to cloud cover, with MAPE of a few percent under clear skies, growing to $\sim$ 50% or more under partly cloudy skies, and then falling again to 10–20% under mostly cloudy sky conditions. Power output during large scale changes, such as strong ramp events, can be estimated with inter-site RMSE of less than 0.10 kW/kW_p using a single system for several other sites.

K_{PV} has been shown to be more accurate than the Bird Performance Index, BPI, (Golnas et al., 2011) and the

clearness index, K, (Lonij and Jayadevan, 2012) because it uses a more accurate method to estimate a PV system's clear-sky power generation. K_{PV} also is expected to be a more robust estimator than either BPI or K, particularly in persistently cloudy conditions. K_{PV} has the potential to be modified in future work to handle PV systems with panels at significantly different orientations, while BPI and K cannot. These improvements are only possible if the PV system's characteristics are known.

Future research work will explore this estimation technique in greater detail by utilizing larger data sets and employing input information from meteorological time series. By investigating a larger data set, distance based errors will be calculated as well as using averages of multiple sites to reduce the scale of errors. It will also become possible to apply this technique to estimate the power output for collectives of non-monitored sites with known characteristics using a selection of monitored ones. This could provide a cost-effective method for estimating the power output from all PV sites in a given area. Furthermore, we believe this type of methodology will be very useful in PV system power output forecasting, particularly when applied to collective power output forecasts for tens, hundreds or thousands of distributed PV systems.

Acknowledgements

NAE would like to thank the United States National Science Foundation Graduate Research Fellowship Program and The Australian National University, which provided partial support for this project. We thank pvoutput.org and its members for supplying this data, as well as Sandia National Laboratories for the development of the modeling techniques which have made this research possible.

References

Badescu, V., Gueymard, C., Cheval, S., Oprea, C., Baciu, M., Dumitrescu, A., Iacobescu, F., Milos, I., Rada, C., 2012. Computing global and diffuse solar hourly irradiation on clear sky: review and testing of 54 models. Renew. Sustain. Energy Rev. 16 (3), 1636–1656.

Benghanem, M., Mellit, A., 2010. Radial basis function network-based prediction of global solar radiation data: application for sizing of a stand-alone photovoltaic system at Al-Madinah, Saudi Arabia. Energy 35 (9), 3751–3762.

Bird, R., Hulstrom, R., 1981. A Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces. Tech. Rep. SERI/TR-642-761, Solar Energy Research Institute.

Black, J., Bonython, C., Prescott, J., 1954. Solar radiation and the duration of sunshine. Quart. J. Roy. Meteorol. Soc. 80, 231–235.

Calbó, J., González, J., Pagès, D., 2001. A method for sky-condition classification from ground-based solar radiation measurements. J. Appl. Meteorol., 2193–2199.

Cameron, C., Stein, J., Hansen, C., 2011. Evaluation of PV performance models and their impact on project risk. In: PV Rollout Conference.

Cameron, C.P., Boyson, W.E., Riley, D.M., 2008. Comparision of PV system performance-model predictions with measured PV system performance. In: 2008 33rd IEEE Photovolatic Specialists Conference, pp. 1–6.

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

Davies, J., McKay, D., 1982. Estimating solar irradiance and components. Sol. Energy 29 (1), 55–64.

Davy, R., Woods, M., Russell, C., Coppin, P., 2010. Statistical downscaling of wind variability from meteorological fields. Bound.-Layer Meteorol. 135 (1), 161–175.

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

Engerer, N.A., 2012. Estimating collective rooftop photovoltaic power output. In: Australian Solar Council Solar 2012. Melbourne, Australia. <http://solarnick.info>.

Fanney, H., Davis, M., Dougherty, B., King, D., Boyson, W., Kratochvil, J., 2006. Comparison of photovoltaic module performance measurements. J. Sol. Energy Eng. 128 (2), 152.

Gaetan, M., Latour, M., Rekinger, M., Theologitis, L., Papoutsi, M., 2013. Cloud cameras at the Pierre Auger Observatory. Tech. Rep., European Photovoltaic Industry Association.

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.

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.

Gueymard, C., 2003. Direct solar transmittance and irradiance predictions with broadband models. Part II: Validation with high-quality measurements. Sol. Energy 74, 381–395.

Gueymard, C., 2009. Direct and indirect uncertainties in the prediction of tilted irradiance for solar engineering applications. Sol. Energy 83 (3), 432–444.

Hay, J., McKay, D., 1988. Calculation of Solar Irradiances for Inclined Surfaces: Verification of Models Which Use Hourly and Daily Data. Tech. Rep., International Energy Agency.

Hoff, T.E., Perez, R., 2010. Quantifying PV power output variability. Sol. Energy 84 (10), 1782–1793.

Ineichen, P., 1983. Quatre annees de mesures d ensoleiilement a Geneve. Ph.D. thesis, University of Geneva.

Ineichen, P., 2006. Comparison of eight clear sky broadband models against 16 independent data banks. Sol. Energy 80 (4), 468–478.

Ineichen, P., Perez, R., 2002. A new airmass independent formulation for the linke turbidity coefficient. Sol. Energy 73 (3), 151–157.

Jamaly, M., Bosch, J.L., Kleissl, J., 2013. Aggregate ramp rates of distributed photovoltaic systems in San Diego County. IEEE Trans. Sustain. Energy 4 (2), 519–526.

Kambezidis, H.D., Psiloglou, B.E., Gueymard, C., Energy, F.S., Canaveral, C., 1994. Measurements and models for total solar irradiance on inclined surface in Athens, Greece. Sol. Energy 53 (2).

Kasten, F., 1988. Elimination of the virtual diurnal variation of the linke turbidity factor. Meteor. Rdsch. 41, 93–94.

King, D., Boyson, W., Kratochvill, J., 2004. Photovoltaic Array Performance Model. Tech. Rep. SAND2004-3535, Sandia National Laboratories.

King, D., Gonzalez, S., Galbraith, G., Boyson, W., 2007. Performance Model for Grid-Connected Photovoltaic Inverters. Tech. Rep. SAND2007-5036.

Klise, G., Stein, J., 2009. Models Used to Assess the Performance of Photovoltaic Systems. Tech. Rep. SAND2009-8258, Sandia National Laboratories.

Lave, M., Kleissl, J., 2010. Solar variability of four sites across the state of Colorado. Renew. Energy 35 (12), 2867–2873.

Lave, M., Kleissl, J., Arias-Castro, E., 2011a. High-frequency irradiance fluctuations and geographic smoothing. Sol. Energy 86 (8), 2190–2199.

Lave, M., Kleissl, J., Stein, J.S., 2013. A wavelet-based variability model (WVM) for solar PV power plants. IEEE Trans. Sustain. Energy 4 (2), 501–509.

Lave, M., Stein, J., Ellis, A., Hansen, C., Nakashima, E., November 2011b. Ota City: Characterizing Output Variability from 553 Homes with Residential PV Systems on a Distribution Feeder. Tech. Rep. SAND2011-9011, Sandia National Laboratories.

Liu, B., Jordan, R., 1960. The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Sol. Energy 4 (3), 1–19.

Liu, B., Jordan, R., 1963. The long-term average performance of flat-plate solar energy collectors. Sol. Energy 7 (2), 53–70.

Lonij, V., Brooks, A., Koch, K., Cronin, A., 2012. Analysis of 80 Rooftop PV systems in the Tucson, AZ area. In: 38th IEEE Photovoltaic Specialists Conference (PVSC).

Lonij, V., Jayadevan, V., 2012. Forecasts of PV power output using power measurements of 80 residential PV installs. In; 38th IEEE Photovoltaic Specialists Conference (PVSC).

Lonij, V.P., Brooks, A.E., Cronin, A.D., Leuthold, M., Koch, K., 2013. Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors. Sol. Energy 97, 58–66.

Loutzenhiser, P., Manz, H., Felsmann, C., Strachan, P., Frank, T., Maxwell, G., 2007. Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation. Sol. Energy 81 (2), 254–267.

Mathiesen, P., Kleissl, J., 2011. Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States. Sol. Energy 85 (5), 967–977.

McKenney, D.W., Pelland, S., Poissant, Y., Morris, R., Hutchinson, M., Papadopol, P., Lawrence, K., Campbell, K., 2008. Spatial insolation models for photovoltaic energy in Canada. Sol. Energy 82 (11), 1049–1061.

Mills, A., Wiser, R., 2010. Implications of Wide-Area Geographic Diversity for Short-Term Variability of Solar Power. Tech. Rep., Lawrence Berkeley National Laboratory.

Molineaux, B., Ineichen, P., Delaunay, J., 1995. Direct luminous efficacy and atmospheric turbidity – improving model performance. Sol. Energy 55 (2), 125–137.

Pages, D., Calbo, J., Gonzalez, J., 2003. Using routine meteorological data to derive sky conditions. Annal. Geophys. 21, 649–654.

Reindl, D., Beckman, W., Duffie, J., 1990. Evaluation of hourly tilted surface radiation models. Sol. Energy 45 (1), 9–17.

Reno, M., Hansen, C., Stein, J., March 2012. Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis. Tech. Rep. SAND2012-2389, Sandia National Laboratories.

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. Sol. Energy 68 (1), 33–48.

Rigollier, C., Lefevre, M., Wald, L., 2001. Heliosat version 2. Integration and Exploitation of Networked Solar Radiation Databases for Environment Monitoring (SoDa Project). Tech. Rep. <http://soda-is.com>.

Sfetsos, A., Coonick, A., 2000. Univariate and multi-variate forecasting of hourly solar radiation with artificial intelligence techniques. Sol. Energy 68 (2), 169–178.

Skartveit, A., Olseth, J., 1992. The probability density and autocorrelation of short-term global and beam irradiance. Sol. Energy 49 (6), 477–487.

Stein, J., 2012. The Photovoltaic Performance Modeling Collaborative (PVPMC). Tech. Rep. SAND2012-4531, Sandia National Laboratories.

SunWiz, 2013. PV Market Insights. Tech. Rep., SunWiz Consultants.

Woyte, A., Belmans, R., Nijs, J., 2007. Fluctuations in instantaneous clearness index: analysis and statistics. Sol. Energy 81 (2), 195–206.

Yang, D., Jirutitijaroen, P., Walsh, W.M., 2012. Hourly solar irradiance time series forecasting using cloud cover index. Sol. Energy 86 (12), 3531–3543.

Zarzalejo, L.F., Polo, J., Martín, L., Ramírez, L., Espinar, B., 2009. A new statistical approach for deriving global solar radiation from satellite images. Sol. Energy 83 (4), 480–484.