Towards a Tuning Method of PV Power Measurements to Balance Systematic Influences

Sven Killinger[1], Jamie M. Bright[1], David Lingfors[2], and Nicholas A. Engerer[1,3]

[1]Fenner School of Environment and Society, The Australian National University, Canberra, Australia
[2]Department of Engineering Sciences, Uppsala University, Lägerhyddsvägen 1, 752 37 Uppsala, Sweden
[3]Corresponding author: nicholas.engerer@anu.edu.au, http://nickengerer.org

Abstract

With rapid deployment and penetration rates of residential photovoltaic (PV) systems in the distribution grid, there is growing need for accurate assessment of the real-time power generation for grid management and energy market operations. Many of these installed PV systems report their live power generation to online databases and can be used as references to estimate the power generation of neighbouring systems. Upscaling approaches have demonstrated their capability of using the data from these reference PV systems to estimate the power output of target PV systems that do not report their power generation data. However, there is an inherent issue with the representativeness of these reference PV systems power data, e.g. due to quality issues or system specific influences such as shading. Three methods were developed by the authors in earlier work: (1) a parametrisation of PV system metadata and quality control of the measured power, (2) a tuning routine that detects diurnal influences from shading and tunes the PV power in order to reach the expected generation without any shading. And (3) a method which eliminates high variances in k_{pv} based upscaling. An extensive cross-validation with 308 systems in Canberra, Australia in this paper shows significant improvements as a direct result of the application of these three methods. Furthermore, we present the preliminary findings for developments in: the parametrisation of shaded/multi-azimuth reference PV systems, as well as a method to reduce inertia in the shade detection and tuning. Overall, we successfully improve the management of reference PV system power data for use in upscaling.

Keywords – Upscaling, Parametrisation, Quality Control, Tuning, PV power

1. Introduction

In recent years, there has been worldwide growth in the total installed capacity of photovoltaic (PV) systems from 5.1 GWp to 300 GWp between 2005 and 2016 (REN21, 2016; Wedepohl, 2017). For operational control of the electricity grid and for economic trading in the electricity generation market, the precise knowledge of the current PV power generation is of increasingly high value (ENA, 2017).

A growing number of PV systems continuously report their power generation, enabling their data to be taken as a reference. These so called "reference PV systems" can be used to estimate the power of the remaining target PV systems through "upscaling" approaches. An extensive literature review covering upscaling approaches is provided by Bright et al. (2017). Upscaling approaches are impacted influences such as individual characteristics of the PV systems, their spatial distribution, the temporal and spatial aggregation level, interaction of individual and collective system behaviour and module orientation, as well as how variable small-scale weather conditions overlay each other in complex ways.

It is a direct consequence of these influences that not all reference PV systems reflect the power generation of neighboured systems. Earlier research present methods that improve upscaling approaches by increasing the representativeness of reference PV systems:

In Killinger et al. (2017b), a twofold routine is presented that parametrises metadata of reference PV systems (tilt β, azimuth γ and loss factor LF) and then quality controls their PV power output measurements. By eliminating measurement errors and other atypical data from the pool of reference PV systems, the representativeness can be increased. Killinger et al. (2017a) extends the concept of making refrernce PV systems appropriate for upscaling through the development of a tuning routine that can detect and balance diurnal patterns in PV power generation due to shading. The outcome of the tuning routine is an increase in representativeness of reference PV systems. This is because tuned reference PV systems better describe the current meteorological conditions and are a better source to estimate the power generation of other neighboured systems. Such is the success of reference PV system tuning, commercial solar forecasting company Solcast (2017); Engerer et al. (2017) are adopting the approach to drastically improve their forecasts. The preliminary use case for the tuning is demonstrated in Bright et al. (2017).

doi:10.18086/swc.2017.21.03 Available at http://proceedings.ises.org

Lingfors et al. (2017b) evaluated the tuning approach's capability at correctly detecting shading events and compared it with a LiDAR based approach (Lingfors et al., 2017a). The comparison highlighted important features in the tuning that had room for improvement, thus these improvements are the feature of this paper.

This paper has two main objectives. Firstly, the latest advancements in upscaling shall be evaluated in an extensive cross-validation with 308 systems in Canberra, Australia. These methods are briefly presented in section 2. In section 3, important data sources and models are announced. The evaluation is finally realised and discussed in section 4. Secondly, preliminary methods to further improve upscaling are developed and tested in section 5. The paper ends with a short conclusion in section 6.

2. Methodology

The methodological steps are presented in this section. An overview is illustrated in the flow chart of Fig. 1.

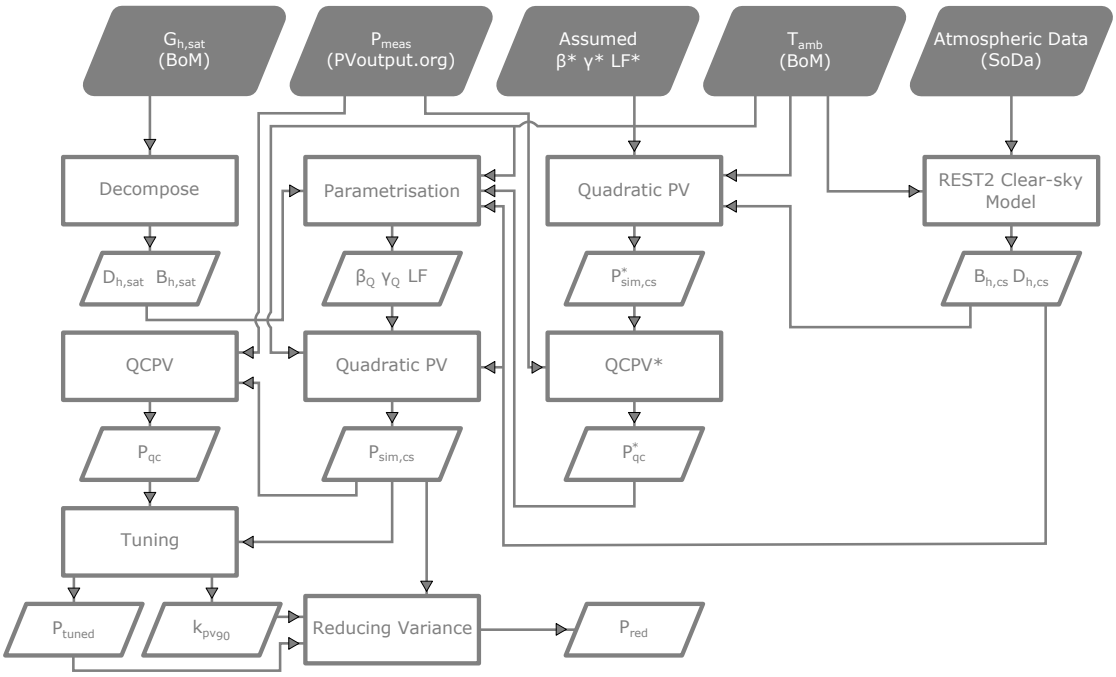

Fig. 1: Flow chart of the proposed methodology. Input data are in grey skewed boxes with data source provider in brackets. Model process steps are in rectangles and intermediate/final outputs are in white parallelograms.

2.1. Parametrisation and QCPV routine

An approach to parametrise the reference PV system metadata (tilt β, azimuth γ and loss factor LF) and to quality control the PV power measurements (QCPV routine) is presented by Killinger et al. (2017b). Both the parametrisation and QCPV routine were improved in Killinger et al. (2017a) and briefly summarised by Bright et al. (2017) and Lingfors et al. (2017b) Hence, only the basic principles of the methodology are detailed within this section.

1. To eliminate critical measurement errors from the parametrisation of the metadata, a modified version of the QCPV routine is applied to the measured power, P_{meas}, to find P^*_{qc} (Killinger et al., 2017a). All intermediate steps of the modified QCPV routine are marked with a * in Fig. 1. A central element in the routine to detect atypical behaviour is the simulation of the clear-sky power $P^*_{sim,cs}$. For the modified version of the quality control of PV power measurements (QCPV*), we assume the same metadata for all systems; north-facing azimuth (Canberra lies on the southern hemisphere), a tilt of 30°, and LF of 1. Two tests, the "maximum k_{pv}" and "cloud enhancements" criteria (section 5.1.3. and 5.2.1. in Killinger et al. (2017b), respectively) are excluded from the modified version QCPV*, as more detailed metadata would be required.

2. In the second step, the parametrisation of metadata is realised. The Non-linear Least-Square (NLS) solver takes satellite-derived beam ($B_{h,sat}$) and diffuse ($D_{h,sat}$) irradiance subcomponents (decomposed with

the Reindl decomposition model as inputs (Reindl et al., 1990a)), along with the ambient temperature (T_{amb}). These inputs are used to simulate PV power depending on the tilt, azimuth and LF. The three latter parameters are varied within the NLS solver until the squared error between the simulated power and the quality controlled measured power is minimal. Within the simulation chain, a quadratic PV power model (Killinger et al., 2017b) and transposition model (Reindl et al., 1990b) are applied. In a second step of the parametrisation, the derived LF is precisely adjusted based on detected clear-sky periods (see Killinger et al. (2017b) and Killinger et al. (2017a)).

3. Based on the parametrised metadata (tilt β_Q, azimuth γ_Q and loss factor LF), the PV power measurements are finally quality controlled. All tests are applied in this second run of the QCPV routine and atypical power values are set to NaN (not a number).

2.2. Tuning

The method for identification and correction of systematic influences that affect PV power output (particularly shading) at various time-steps throughout the day is known as the tuning routine, and is presented in detail by Killinger et al. (2017a). With the completion of parametrisation and QCPV routine, the quality controlled power output P_{qc} and the clear-sky power output at a reference PV system $P_{sim,cs}$ are used to calculate the clear-sky index for photovoltaics k_{pv} (Engerer and Mills, 2014):

$$k_{pv} = \frac{P_{qc}}{P_{sim,cs}}. \qquad \text{(Eq. 1)}$$

In theory, k_{pv} achieves a value of 1 when subjected to clear-sky conditions and higher values only when subjected to cloud enhancement events (Engerer, 2015; Killinger et al., 2017b). Low values of k_{pv} occur due to either increased thickness of cloud cover or reduced system performance. Conversely, in a case where a PV system never achieves clear-sky conditions in the form of $k_{pv} = 1$ for repeated particular time-steps in the course of the day, these periods can be considered to be influenced by systematic reduction, which is most predominantly caused by shading. For the reliable detection of shading, the k_{pv} values are analysed for a running window of time over 30 days. The 90$^{\text{th}}$ percentile ($k_{pv_{90}}$) is derived for each 10-min time-step of the day through an approach described by Lonij et al. (2012). The selection of the 90$^{\text{th}}$ percentile instead of other percentiles was a result of our analysis indicating it delivered a robust limit for representing clear-sky operations; in theory, k_{pv} values within the 90$^{\text{th}}$ are most likely to be clear-sky time periods. With the derivation of $k_{pv_{90}}$ and after the application of logical constraints, the tuning factor X can be calculated and allows for the tuning requirement of a system to be calculated for each time-step according to:

$$P_{tuned} = \frac{P_{qc}}{X}, \qquad \text{(Eq. 2)}$$

where P_{tuned} is the tuned power in $\mathrm{W/W_p}$ for a given system.

In essence, P_{tuned} represents the power profile of a reference PV system with all the systematic influences removed. By using P_{tuned} as opposed to P_{meas} or P_{qc} in the upscaling approach, power measurements recorded under shading of the reference PV system are corrected by balancing systematic influences (or tuned) before being used to estimate an unshaded neighbour.

2.3. Reducing k_{pv} related variance

Variances in k_{pv} are mainly due to cloud cover changes, systematic influences (e.g. shading), or inaccurate simulation of $P_{sim,cs}$. The latter two reasons are reference PV system specific and the influence of their probability and magnitude grows with increasing zenith angles θ_z. These reasons are increasingly important to consider when performing upscaling approaches as filtering system specific systematic influences in order to improve the representativeness of the reference PV systems. Systematic influences (such as shading) are already balanced with the tuning approach, and small inaccuracies in the simulation of $P_{sim,cs}$ can hardly be avoided. For example, if $P_{sim,cs}$ is slightly underestimated for small power measurements occurring during high zenith angles θ_z (0.005 $\mathrm{W/W_p}$ instead of 0.01 $\mathrm{W/W_p}$), the resultant k_{pv} factor would be twice as high; the same is applicable to overestimations in $P_{sim,cs}$. Should the same inaccuracy occur at midday (0.5 $\mathrm{W/W_p}$ instead of 0.505 $\mathrm{W/W_p}$), the effect on the k_{pv} value would be relatively insignificant. Consequently, the impact of such inaccuracies is limited by removing power measurements P_{tuned} from reference PV systems where time-steps exhibit high k_{pv}-related variances. This filter is realised by analysing the standard deviation of $k_{pv_{90}}$ in parallel to k_{pv}, detailed by Killinger et al. (2017a). With the removal of particularly large variances in the morning or evening by setting atypical values NaN, PV power with reduced variances (P_{red}) is derived from P_{tuned}.

3. Data

The data sources and models used within this paper are mostly identical with the ones described in Killinger et al. (2017a) and are mentioned in Fig. 1. Power output data is significantly extended in this paper covering 308 PV installations in Canberra, Australia, where previously only 78 were considered. The power measurements P_{meas} are extracted from http://pvoutput.org and subsequently normalised by their reported installed capacity. The collected power output data comprises a total period spanning from 7th September 2016 till 31st July 2017 with missing data from 25th March 2017 till 22nd April 2017. In contrary to Killinger et al. (2017a), multi-azimuth systems are considered within this work. Ambient temperature T_{amb} values are obtained from the Canberra airport (Australian Bureau of Meteorology station 070351) and missing data replaced by values from the numeric weather model MERRA, derived from http://www.soda-pro.com.

Climatological atmospheric parameters of ozone, water vapour, aerosol optical depth and Linke turbidity for the clear-sky irradiance modelling are retrieved from the SoDa database as described in Engerer and Mills (2015); Killinger et al. (2017b). These parameters are used to derive the beam ($B_{h,cs}$) and diffuse ($D_{h,cs}$) horizontal clear-sky irradiance via the REST2 model by Gueymard (2008). The derivation of the satellite based irradiance is described in detail in Engerer et al. (2017).

As the satellite based irradiance has a temporal resolution of 10-mins, all datasets used within this paper (e.g. the PV power measurements, ambient temperature, etc.) are aggregated to a resolution of 10 minutes in order to correspond.

4. Results and discussion

This section will firstly define the procedure of the cross-validations, and subsequently present and discuss the results.

4.1. Definition of the cross-validations

In this section, four different cross-validations (CV) labelled $k = 1, 2, 3, 4$ are realised. Within these cross-validations, the estimated power output of target PV systems P_{est,CV_k} is simulated using k_{pv}-based upscaling techniques (see Engerer and Mills (2014), Killinger et al. (2017a) and Bright et al. (2017)). k_{pv} based upscaling estimates the k_{pv} value at a target PV system's location using a geospatial interpolation approach called Inverse Distance Weighting (IDW, defined in Shepard (1968) and using the recommended factor of 2) and weights the influence of a reference PV system as a function of its distance to a target PV system; the closer a reference PV system is to the target, the more influence it has. With the interpolated k_{pv} value at a target PV system's location alongside its simulated clear-sky power $P_{sim,cs}$, the power output can now be estimated following the principle of Eq. 1.

Detailed assumptions of the cross-validations are provided in Tab. 1. CV1 is the baseline scenario without any data processing that uses metadata directly from http://pvoutput.org. Should metadata be missing, a north-facing azimuth and a tilt of 30° are assumed (P_{meas}). For all other validations (CV2-CV4), the parametrised metadata of both the reference and target PV systems as well as quality controlled power data from the methodology described in section 2.1 are used. In CV2, only the QCPV methodology from section 2.1 is applied (P_{qc}). In CV3, the additional inclusion of the tuning routine from section 2.2 is applied (P_{tuned}). Lastly, CV4 extends the tuning routine through application of the variance reduction routine (P_{red}) from section 2.3.

Estimates from each cross-validation are compared to the quality controlled version of the measured power P_{qc} in order to have a common dataset that is (mostly) free of measurement errors; using P_{meas} holds no value as there exists unquestionably erroneous data prevalent throughout the dataset. The simulated power estimates and number of NaN values are different in CV1-4 due to increased strictness of filtering with each methodology. If simulated or measured data registers a NaN, it is excluded from the evaluation.

Tab. 1: Definition for different cross-validations (CV).

	CV1	CV2	CV3	CV4
QCPV Routine		✓	✓	✓
Tuning			✓	✓
Reducing Variance				✓
Module Orientation	pvoutput.org	Simulated	Simulated	Simulated
Loss Factor LF	0.9	Simulated	Simulated	Simulated
P_{tar}	P_{qc}	P_{qc}	P_{qc}	P_{qc}

The cross-validations are conducted for an increasing number of reference PV systems I (2, 5, 10, 20, 30, 40, 50) and a fixed number of $J = 250$ target PV systems are simulated. The purpose is to explore the impact that number of available reference PV systems has on the accuracy of power estimate. For each value of I, simulations are repeated 100 times in order to obtain a representative distribution of error metrics and test reproducibility. With each test, the reference and target PV systems are randomly re-assigned each time while maintaining I and J allowing for a rich number of combinations per CV. The comparisons explore the performance of all the power estimates P_{est,CV_k} from each cross-validation $k = 1, 2, 3, 4$ against the actual power output from the target PV systems P_{tar}. Estimated and quality controlled measured target PV system power values are aggregated at each time-step t and normalised by J. P_{est,CV_k} and P_{tar} are, therefore, time series of aggregated power output for all J target PV systems.

An example test (1 of 100) to derive results is described:

1. $J = 250$ target PV systems randomly selected
2. I number of reference PV systems randomly selected from the remaining 58 systems
3. Aggregated P_{est,CV_k} calculated for each cross-validation $k = 1, 2, 3, 4$
4. Aggregated P_{tar} calculated
5. Compare power aggregate time series using different error metrics

Steps 1 to 5 are repeated 100 times before changing the value of I.

In the following evaluations, all results are restricted to time-steps fulfilling the zenith angle $\theta_z < 85°$ for the same reasons as mentioned in section 4.1. in Killinger et al. (2017a). Furthermore, all k_{pv} values used within the CV evaluations that are greater than 2 are excluded from the examinations as they represent an unrealistic magnitude.

4.2. Evaluation of the cross-validations

Results of the cross-validations are shown in Fig. 2. Each boxplot contains a distribution of results from the 100 different simulations, each with a unique combination of reference and target PV systems. The boxplots of the relative root mean square error ($rRMSE$, top plot in Fig. 2) indicate how each additional method presented in this paper further improves the estimated power outputs from the reference PV systems. The largest improvements are reached between CV1 and CV2 with the inclusion of QCPV. Although $rRMSE$ values further improve with CV3 and CV4, a decreasing marginal utility is visible.

Results in the $rRMSE$ also improve for an increasing number of reference PV systems with CV2-4. A similar improvement for an increasing number of reference PV systems can be seen for the baseline scenario CV1. An exception from this tendency occurs for a small number of reference PV systems in CV1 with $I = 2$ refrence PV systems, which is lower in $rRMSE$ than for $I = 5$; it is expected that this is a statistical coincidence which would disappear for a higher number of samples. Please also note that the ranges of the boxplots and outliers decrease for a greater number of reference PV systems; a balancing impact occurs with more reference PV systems and reduces uncertainty in the power estimation.

The uncertainty in the power simulation is not only reduced by an increasing number of reference PV systems, but also for the methods presented in this paper. This can be seen through analysis of the boxplots of the relative mean bias error ($rMBE$, middle plot in Fig. 2). The smallest ranges of $rMBE$ occur for CV3 and CV4 with a tendency towards overestimations. This can be explained with the tuning routine that increases the power magnitude of reference PV systems, however at present, this increase is applied to target PV systems that may also be shaded, resulting in an overestimation. These overestimations are not constant as shading only occurs for some systems at unique time-steps. In order to reduce such an estimation, future versions of the tuning must consider shading, perhaps by statistical diurnal consideration of typical shading events that can be probabilistically applied. This bias is rather dynamic and is different for each time-step of the day, making correction difficult.

Through the careful selection of time periods in the QCPV routine, as well as the constraints in the tuning and reduction of the variance, the number of NaN values in the time series of the reference PV systems increases. In the lower plot of Fig. 2 the time-steps excluded for the cross-validations are visualised. A time-step is excluded if either P_{est_k} or P_{tar} are NaN. While CV1 shows a rather consistent share of less than 25%, CV2-4 are greater for a small number of reference PV systems. The share slightly increases in CV4 due to the additional filters. At time-steps when a reference PV system registers a NaN value, it is obviously no longer able to estimate the power of other systems. Hence, for a very small number of reference PV systems, there can be individual time-steps in which it might be necessary to renounce the use of one of the methods or better use alternative data to estimate the power of the target PV systems. For all other time-steps, it is strongly recommended to use all the methods presented in this paper.

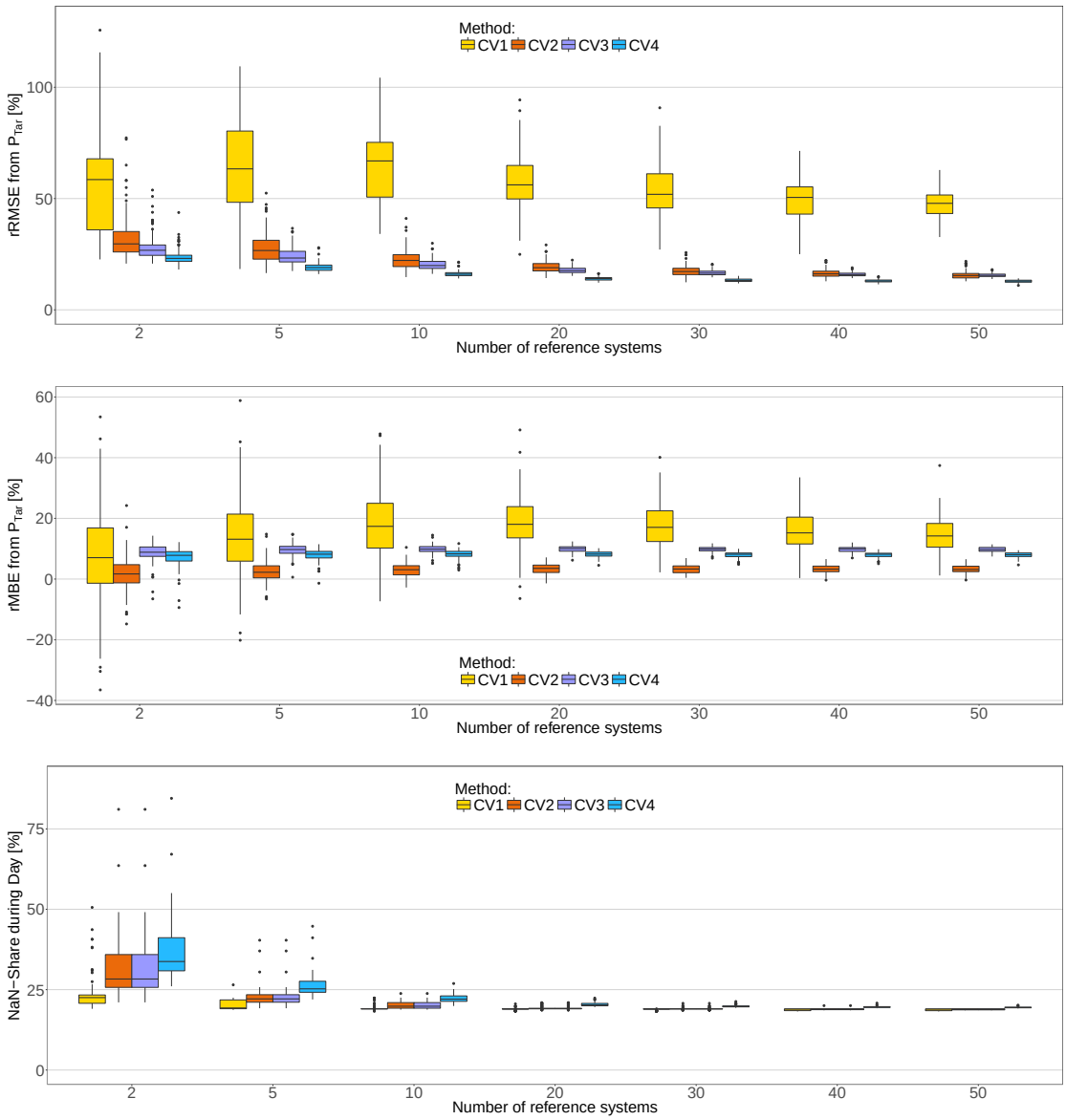

Fig. 2: Box plots indicating the four quartiles of error distributions. The lower, middle and upper lines of each box indicates the 25%, 50%/median, and 75% quartiles, respectively. Whiskers indicate the minimum and maximum values of the distribution. Outliers are represented as points and are defined as existing 1.5 times outside the inter-quartile range of the 25% and 75% quartiles. Each box plot represents the resulting error metric on the y-axis derived when comparing the estimate power value of a cross-validation $k = 1, 2, 3, 4$ (P_{est,CV_k}) against the target power value (P_{tar}). The error is plotted against the number of reference PV systems I. The colours represent the different cross-validations CV1-4. The top, middle and bottom plots indicate the relative root mean square error (rRMSE), the relative mean bias error (rMBE), and the fractional NaN-share, respectively.

5. Preliminary advancements

The analysis in the previous section showed that the strongest relative improvements are possible for a small number of available reference PV systems. Particularly for a small number of reference PV systems, it is important to better understand their individual behaviours. Hence, three different and preliminary methods are presented and evaluated within this section:

- The parametrisation of shaded systems in section 5.1
- The consideration of multi-azimuth systems in the parametrisation stage in section 5.2
- The reduction of the inertia in the tuning routine in section 5.3

The development of these methods are based on findings by Killinger et al. (2017a), Lingfors et al. (2017b), Bright et al. (2017) and from within this paper.

5.1. Parametrisation of shaded systems

Diurnal shading events can strongly influence the daily power profile of a PV system. Especially in case of strong shading in the morning or evening, it can be difficult for an observer to distinguish between shading events or a normal behaviour caused by an (extreme) azimuth angle of a system facing east or west. The parametrisation approach described in section 2.1 is facing exactly this challenge. Four PV systems are discussed in Lingfors et al. (2017b), which show significant errors in the parametrisation of the azimuth angles and strong influences of shading in the morning (see Lingfors et al. (2017b) ID 135, 157) or evening (ID 44, 67). The problem lies in the procedure of the NLS solver, which tries to find a combination of tilt (β_Q), azimuth (γ_Q) and loss factor (LF) leading to the smallest least square error. Even though this procedure hast a high rate of success for unshaded systems, there is a certain risk for erroneous parametrisation for shaded systems. As a consequence, all subsequent steps such as the QCPV and tuning routine are negatively influenced. Motivated by this challenge, an alternative procedure was developed and is presented next. The hypothesis behind this approach is that, should a system be shaded in the morning, the afternoon hours are more suitable to derive the correct module orientation, and vice versa. Hence, the approach uses different time periods throughout the days to run the parametrisation routine.

1. 00:00 - 23:55 ($\widehat{=}$ old procedure described in section 2.1)
2. 10:00 - 23:55 (if shaded in the early morning hours)
3. 12:00 - 23:55 (if shaded in the morning hours)
4. 00:00 - 13:00 (if shaded in the afternoon hours)
5. 00:00 - 15:00 (if shaded in the late afternoon hours)
6. 08:00 - 17:00 (if shaded in the early morning and late afternoon hours)
7. 10:00 - 14:00 (if shaded in the morning and afternoon hours)

The PV power is simulated with the derived parametrisation for each of the seven time periods. Then, the Mean Absolute Percentage Error (MAPE) is calculated between the simulated power and the quality controlled version of the measured power P^*_{qc}. The parametrisation leading to the smallest MAPE is then chosen.

Tab. 2: Tilt and azimuth of the studied PV systems identified by the old and new parametrisation approach. As a reference, reported tilt from SolarHub (SH, a reputable local solar installer in Canberra) and azimuths derived by inspecting aerial images (AI). The table includes in the last two columns the angle θ_n between the normal vector of the PV module (as derived from SH/AI) and the normal vector of the old and new parametrisation approach, respectively. The building ID corresponds to the study by Lingfors et al. (2017b).

ID	Tilt			Azimuth			θ_n	
	SH [°]	old [°]	new [°]	AI [°]	old [°]	new [°]	old [°]	new [°]
44	25	23	17	-57	-95	-83	18.9	12.1
67	20	28	17	55	10	70	18.6	5.5
135	22	22	9	5	46	1	16.9	12.5
157	18	30	15	-70	-18	-63	23.4	3.9

Tilt and azimuth of the studied PV systems identified by the old and new parametrisation approach are provided in Fig. 5. The tilt reported by SolarHub (SH) and the azimuth from an aerial inspection (AI) are given as references. It can be seen that the azimuth angles strongly improved with the new approach. This is also confirmed by smaller angles between the normal vector of the parametrised module orientation and the orientation from SH/AI. Presented in the table is θ_n, which is the angle between the normal vector of the reference PV system (calculated from from SH/AI metadata) and the parametrised-derived normal vector; an angle of $0°$ is an exact estimation. The use of θ_n is a better representation of the differences of orientation discrepancies than considering the azimuth and tilt separately. The new developments to the parametrisation routine show significant improvement for the four reference PV systems analysed with a 36%, 70%, 26% and 83% reduction in absolute deviation away from the normal for PV system IDs 44, 67, 135 and 157, respectively.

In Fig. 5 the simulated clear-sky curve with the old ($P_{cs,old}$) and new ($P_{cs,new}$) parametrisation method is visualised together with the measured power P_{meas} for an exemplary day. Both systems are facing strong shading events in the morning. As a consequence, the azimuth is overestimated by the old parametrisation, whereas the new method leads to a much more realistic clear-sky curve.

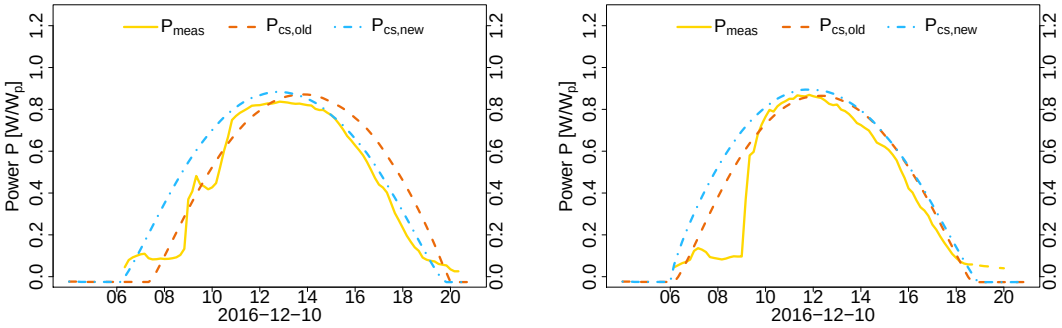

Fig. 3: PV systems (ID 135, left and ID 157, right) being influenced by strong shading events in the morning. Visualized is the measured power P_{meas} (yellow), the simulated clear-sky curve with the old parametrisation method from section 2.1 $P_{cs,old}$ (red) and the new approach from this section $P_{cs,new}$ (blue).

Even though the presented advancement of the parametrisation shows strong improvements for the four studied systems, first tests show problems for unshaded systems, leading sometimes to worsened results than the old parametrisation. Hence, the current challenge is to develop a method that helps to better decide which derived parametrisation option to use. Different error metrics were applied besides the MAPE, but no convincing strategy was found yet and is reserved to be the focus of future work.

5.2. Considering multi-azimuth systems

An additional challenge for the parametrisation approach are PV systems with modules facing different azimuth angles. Such multi-azimuth systems were not considered in Killinger et al. (2017b) and Killinger et al. (2017a) but included in this paper. The parametrisation approach was not adjusted in order to deal with PV power profile of such multi-azimuth systems. The NLS solver in this paper (see Eq. (1) in Killinger et al. (2017b)) estimates the parameters (tilt β_Q, azimuth γ_Q and loss factor LF) by minimising the least squares between the quality controlled version of the measured power P^*_{qc} and the simulated power based on the satellite derived irradiance $P_{sim,sat}$:

$$P^*_{qc} \sim P_{sim,sat}(\beta_Q, \gamma_Q, LF). \qquad \text{(Eq. 3)}$$

In order to deal with multi-azimuth systems, an alternative formulation was developed that allows the parametrisation of two different tilt and azimuth angles:

$$P^*_{qc} \sim \alpha \cdot P_{sim,sat,1}(\beta_{Q1}, \gamma_{Q1}, LF) + (1-\alpha) \cdot P_{sim,sat,2}(\beta_{Q2}, \gamma_{Q2}, LF) \qquad \text{(Eq. 4)}$$

The influence of the different module orientations is controlled by a factor α, being limited to a range between 0 and 1. In order to reduce the complexity, only one common loss factor LF is parametrised. To exclude other (disturbing) influences, this concept was tested based on simulated PV profiles. The findings from a multi and single azimuth system are presented next.

1. A multi-azimuth system (Tilt angles: $50°$ and $30°$; azimuth angles: $-45°$ and $90°$, $LF = 1$ and an $\alpha = 0.25$) was simulated with the transposition model from Reindl et al. (1990b) and the quadratic PV model from Killinger et al. (2017b). Then the NLS solver tried to derive the module orientation with Eq. 3 and alternatively with Eq. 4 based on the same irradiance and temperature data as input. The single azimuth approach lead to a tilt angle of $16°$, an azimuth of $68°$ and a $LF = 1$. Considering Fig. 4, the old parametrisation power estimate (P_{sing}, blue) seems to be a fair compromise for the multi-azimuth system and approximates the simulated PV profile (P_{sim}, yellow) well except for the early morning. The multi-azimuth approach from Eq. 4 (P_{multi}, red) is much more appropriate to approximate the simulated profile, as the morning pattern is excellently captured. The derived parametrisation was almost correct with tilt angles of $50°$ and $31°$; azimuth angles of $-47°$ and $90°$, $LF = 1$ and an $\alpha = 0.25$.

2. A single azimuth system (Tilt angle: 30°; azimuth angle: 0°, $LF = 1$) was simulated with the same models as above. This time the single azimuth approach from Eq. 3 finds the correct parametrisation and is able to perfectly approximate the simulated PV profile (P_{sim}, yellow), as is shown in the right plot of Fig. 4. However, the multi-azimuth approach from Eq. 4 shows small deviations in the late evening. The derived parametrisation in this scenario provides tilt angles of 35° and 17°; azimuth angles of −5° and 33°, $LF = 1$ and an $\alpha = 0.76$. The multi-azimuth approach obviously tried to find a combination of two module orientations describing the profile in the best way in order to satisfy Eq. 4.

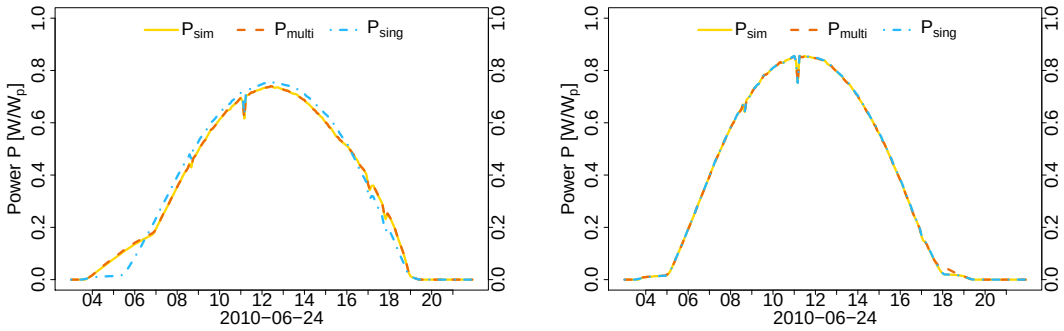

Fig. 4: Exemplary parametrisation for two different PV systems. In the left plot a multi-azimuth system (Tilt angles: 50° and 30°; azimuth angles: −45° and 90°, $LF = 1$ and an $\alpha = 0.25$) was simulated. In the right plot a single azimuth system (Tilt angle: 30°; azimuth angle: 0°, $LF = 1$) was simulated. The plots show the simulated profile (P_{sim}, yellow), the profile approximated by the multi-azimuth solver (P_{multi}, red) from Eq. 4 and the profile approximated by the single azimuth solver (P_{sing}, blue) from Eq. 3.

The above evaluations show that both the single and the multi-azimuth approach have their value and differences between the approaches are smaller than expected. Hence, future work is needed to quantify potential improvements of a multi-azimuth systems by analysing the presence and probability of such systems in reality. Similar to the approach in the last section, another challenge is when to use which approach without knowing if the parametrised system has a single or multi-azimuth angle. A selection based on error metrics between the actual (here: P_{sim}) and approximated (here: P_{sing} or P_{multi}) profiles were tested for some systems, but differences were insignificant. Further research is therefore needed to develop other selection criteria.

5.3. Reducing the inertia in the tuning routine

The discussion in Killinger et al. (2017a) revealed an inertia in the tuning approach that is mainly caused by the determination of $k_{pv_{90}}$ with a rolling window of time over 30 days (see section 2.2). As a consequence, time periods with shading cannot be completely compensated by the tuning approach as can be seen in the left plot in Fig. 5. Different approaches were developed and tested to overcome the inertia. One promising method considered $k_{pv_{90}}$ depending on the sun position as opposed to the time. However, the inertia still persisted. This is because the analysis had to be realised over a longer time period in order to differ between cycling shading influences and meteorological influences. Hence, the following advancement was developed and works in addition to the previous tuning routine.

1. For each time-step t within the time series of a PV system, the mean $k_{pv_{90}}$ values for the 30 minutes before and after the considered time-step are calculated. If one (or both) mean values are < 0.8, shading occurs during the 30 minute time period and criteria 1 is fulfilled. The idea behind this criteria is that, if shading occurs before or after a time-step t, the probability is increased that the current time-step is also affected. To further check if this is the case, two other criteria were developed.

2. Criteria 2 checks if the k_{pv} value of a system at time-step t is < 0.8. If this is the case, criteria 2 is fulfilled and an indication provided, that the system is influenced by shading.

3. In criteria 3, the mean k_{pv} value over all reference PV systems is calculated for time-step t and compared to the k_{pv} value of the analysed system at the same time-step. In case the k_{pv} value of the analysed system is no less than 0.2 smaller than the mean value, criteria 3 is fulfilled.

If criteria 1, 2 and 3 are all fulfilled, the $k_{pv_{90}}$ value is lowered for the analysed time-step t and the system is assigned the smallest mean $k_{pv_{90}}$ value from the 30 minutes before and after. The procedure with its three

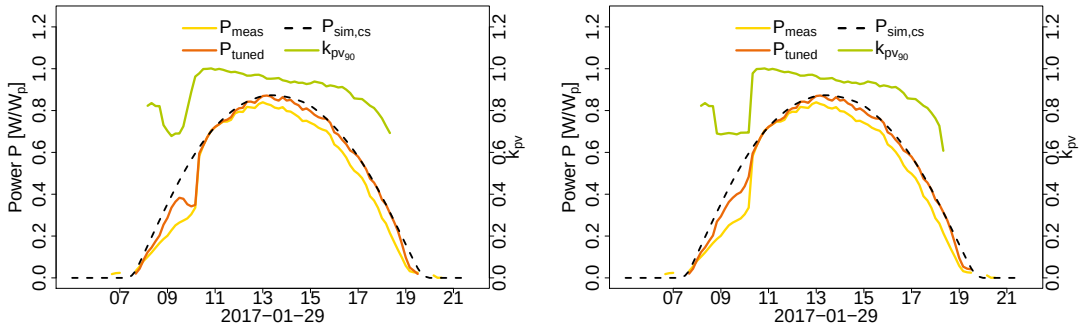

Fig. 5: Outputs for PV site 32 with the tuning method being described in section 2.2 in the left and an alternative approach being presented in section 5.3 in the right plot. The different series represent measured power output (yellow), tuned power output (orange), unsoiled clear-sky power output (black dashed line) and the running 90$^{\text{th}}$ percentile k_{pv} value $k_{pv_{90}}$ (green).

criteria is very cautious and tries to avoid an erroneous tuning. Results are promising as can be seen in the left plot in Fig. 5. The tuning method loses its inertia and is much more capable to balance the influence of shading. In future work this approach will be applied to more systems and carefully examined if the criteria are well suited.

Whilst the main objective of the three criteria was to find quantitative evidence that a time-step is subjected to shading, criteria 3 has an additional positive aspect. The across system analysis helps to consider the power generation of neighboured systems. If the power generation of these systems is reduced, e.g. due to the influence of an overcast sky, the calculated deviation to the analysed system is rather small and criteria 3 not satisfied. In cases where the analysed system is also under overcast skies, the procedure is advantageous because shading would be strongly reduced for an overcast sky. Hence, the consideration of across system analysis is an interesting feature for future developments of the tuning approach. However, there are also indisputable risks that some systems would be adjusted incorrectly by such an across system tuning routine and further developments would be necessary to handle them.

6. Conclusions

This paper aimed to address the latest suggestions and developments to the handling of reported power data from reference PV systems in the context of upscaling approaches. A central element of these upscaling approaches is the clear-sky index of photovoltaics k_{pv}, defined as the ratio between the current power generation of a system and its generation under clear-sky conditions. The key developments are (1) a parametrisation method to derive the module orientation (tilt and azimuth) as well as a loss factor for each system. This method is completed by a quality control routine to detect erroneous and atypical time-steps in PV power measurements. (2) a tuning routine that detects diurnal influences from shading and tunes the PV power in order to reach the expected generation without any shading. And (3) a method which eliminates high variances in k_{pv} based upscaling.

The application of these three advancements was evaluated based on power measurements of 308 PV systems in the region of Canberra, Australia, with data spanning almost one year of data. 250 of these systems are chosen as target PV systems, whereas a changing number systems are used as reference PV systems (2–50). Throughout many simulations, significant improvements were found as a direct result of the application of these three methods. Further improvements are possible when the number of reference PV systems is small, demonstrating the need to further consider individual characteristics of the reference PV systems. This reasoning provided the motivation for the development of three methods firstly presented in preliminary format within paper. These methods focus on (1) the parametrisation of shaded systems, (2) the consideration of multi-azimuth systems in the parametrisation, and (3) the reduction of the inertia in the tuning routine. These methods show preliminary merit in their application, however, have substantial potential for further improvements before they are of operational quality.

7. Conflict of interest

The authors declare no conflicts of interest.

Acknowledgement

The authors would like to thank the Australian Renewable Energy Agency (ARENA) for supporting this work (Research and Development Programme Funding G00854).

References

Bright, J.M., Killinger, S., Engerer, N.A., Lingfors, D., 2017. Improved satellite-derived PV power nowcasting using power data from real-time reference PV systems. Submitted to the special issue of Progress in Solar Energy .

ENA, 2017. Electricity network transformation roadmap: final report. Technical Report. Energy Networks Australia. URL: www.energynetworks.com.au/electricity-network-transformation-roadmap.

Engerer, N.A., 2015. Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia. Solar Energy 116, 215–237. doi:10.1016/j.solener.2015.04.012.

Engerer, N.A., Bright, J.M., Killinger, S., 2017. Himawari-8 enabled real-time distributed PV simulations for distribution networks, in: 44th IEEE Photovoltaic Specialists Conference (PVSC), Washington D.C., USA.

Engerer, N.A., Mills, F.P., 2014. KPV: A clear-sky index for photovoltaics. Solar Energy 105, 679–693. doi:10.1016/j.solener.2014.04.019.

Engerer, N.A., Mills, F.P., 2015. Validating nine clear sky radiation models in Australia. Solar Energy 120, 9–24. doi:10.1016/j.solener.2015.06.044.

Gueymard, C.A., 2008. REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – Validation with a benchmark dataset. Solar Energy 82, 272–285. doi:10.1016/j.solener.2007.04.008.

Killinger, S., Bright, J.M., Lingfors, D., Engerer, N.A., 2017a. A tuning routine to correct systematic influences in reference PV systems' power outputs. Solar Energy 157C, 1082–1094.

Killinger, S., Engerer, N., Müller, B., 2017b. QCPV: A quality control algorithm for distributed photovoltaic array power output. Solar Energy 143, 120–131. doi:10.1016/j.solener.2016.12.053.

Lingfors, D., Bright, J.M., Engerer, N.A., Ahlberg, J., Killinger, S., Widén, J., 2017a. Comparing the capability of low- and high-resolution LiDAR data with application to solar resource assessment, roof type classification and shading analysis. Applied Energy 205, 1216–1230. doi:10.1016/j.apenergy.2017.08.045.

Lingfors, D., Killinger, S., Engerer, N.A., Widén, J., Bright, J.M., 2017b. Identification of PV system shading using a LiDAR-based solar resource assessment model: an evaluation and cross-validation Submitted to Solar Energy.

Lonij, V.P., Brooks, A.E., Koch, K., Cronin, A.D., 2012. Analysis of 80 rooftop PV systems in the Tucson, AZ area, in: 38th IEEE Photovoltaic Specialists Conference (PVSC), Austin, USA. pp. 549–553. doi:10.1109/PVSC.2012.6317674.

Reindl, D.T., Beckman, W.A., Duffie, J.A., 1990a. Diffuse fraction correlations. Solar Energy 45, 1–7. doi:10.1016/0038-092X(90)90060-P.

Reindl, D.T., Beckman, W.A., Duffie, J.A., 1990b. Evaluation of hourly tilted surface radiation models. Solar Energy 45, 9–17. doi:10.1016/0038-092X(90)90061-G.

REN21, 2016. Renewables 2016 - global status report: Key findings. URL: http://www.ren21.net/wp-content/uploads/2016/10/REN21_GSR2016_KeyFindings_en_10.pdf.

Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data, in: 23rd ACM National Conference, Las Vegas, USA. pp. 517–524. doi:10.1145/800186.810616.

Solcast, 2017. Satellite Irradiance estimates from the Himawari-8 meteorological. URL: solcast.com.au.

Wedepohl, D., 2017. Photovoltaik-Meilenstein: weltweit 300 Gigawatt installiert. URL: https://www.solarwirtschaft.de/presse/pressemeldungen/pressemeldungen-im-detail/news/photovoltaik-meilenstein-weltweit-300-gigawatt-installiert.html.