ROOFTOP PV MODEL TECHNICAL REPORT

Victor Depoorter Ruelle
Matthew Jeppesen
Michael Brear

The University of Melbourne

22 July 2016

Contents

1. Executive summary

This document presents the methodology used to develop a model of the generation from rooftop PV systems in the National Electricity Market (NEM). It was developed by the University of Melbourne and AEMO, and is intended to improve the modelling of rooftop PV across multiple working teams in AEMO. Further information on model validation and performance will be published in the near future.

The main features of this model are as follows.

- It calculates distributed PV generation for each half hour at different levels of aggregation from single systems through to postcodes, connection points, states and the NEM.
- It calculates the generation uncertainty as 10% and 90% Probability of Exceedance (POE) levels.
- It was developed using historical generation traces from more than 40,000 installed rooftop PV systems across the NEM.
- It is able to calculate historical or project future PV generation depending on whether historical or forecast weather data is used.

The overall performance of this model can be summarised by two measures.

- The Normalised Mean Absolute Error (NMAE) for half hourly generation is approximately:
  - 8% for individual systems,
  - 6% for postcodes and connection points, and
  - 4% for states and the NEM.
- The annual Normalised BIAS (NBIAS) of the calculated generation is approximately 1.5% for postcodes and connection points and around 1% for states and the NEM.

2. Introduction and overview

The National Electricity Market (NEM) has experienced a rapid uptake of rooftop PV systems in recent years, with its installed capacity growing from 23 MW in 2008 to over 4,300 MW in 2016 [1]. These systems number more than one million individual installations and are estimated to have produced over 5,000 GWh in 2015–16 financial year, or 3% of the NEM's annual operational electricity consumption. This uptake is expected to continue strongly.

Whilst rooftop PV generation reduces greenhouse gas emissions, it also creates several challenges. Rooftop PV generation is not dispatchable, does not contribute inertia to the system, and can create local network issues [2]. In addition, there are limitations with the accuracy in short term PV generation forecasting; and the generation from almost all PV systems is not directly measured and therefore less certain than for other forms of generation. Related to this, for most PV installations only the net energy flow to the grid (export or import) is measured for the site at which a given PV system is installed. This practice, known as net metering, measures the sum of the local PV generation and the local consumption, thus precluding measurement of the PV generation on its own. All of these challenges will have greater impact on market operations, planning, and forecasting as rooftop PV becomes a larger generator in the NEM [3].

There is therefore an increasing need for an accurate and validated rooftop PV generation model that can be used to understand historical generation and also for producing forecasts. This document details such a model.

3. Methodology

This section describes the methodology used to develop the rooftop PV model. The model has three main components:

- **PV system configuration and performance model**, which calculates the configuration and performance (capacity, tilt, azimuth, shading, and derating) of installed rooftop PV systems in Australia. These parameters are calculated for each system individually, using its generation data and matched weather data. This information is then stored in a database and used in the second component of the model. This analysis is important because the configuration and performance of PV systems have a significant impact on their generation, and information about PV systems in Australia and internationally is often limited and approximate [4].
- **Distributed PV generation model,** which calculates the aggregated rooftop PV generation for each half hour in each NEM postcode, connection point, and state, as well as the associated uncertainty modelled as 10% and 90% Probability of Exceedance (POE) levels. The model is able to calculate the PV generation for any location using historical or forecasted weather data. To date, it has been used to calculate historical rooftop PV generation from 2000 to 2016.
- **Single system model,** which calculates the generation from an individual PV system given its configuration and weather data, and is a core part of both the above models. The single system model has been developed using the PV_LIB software library [5].

An overview of the rooftop PV model is shown in Figure 1.

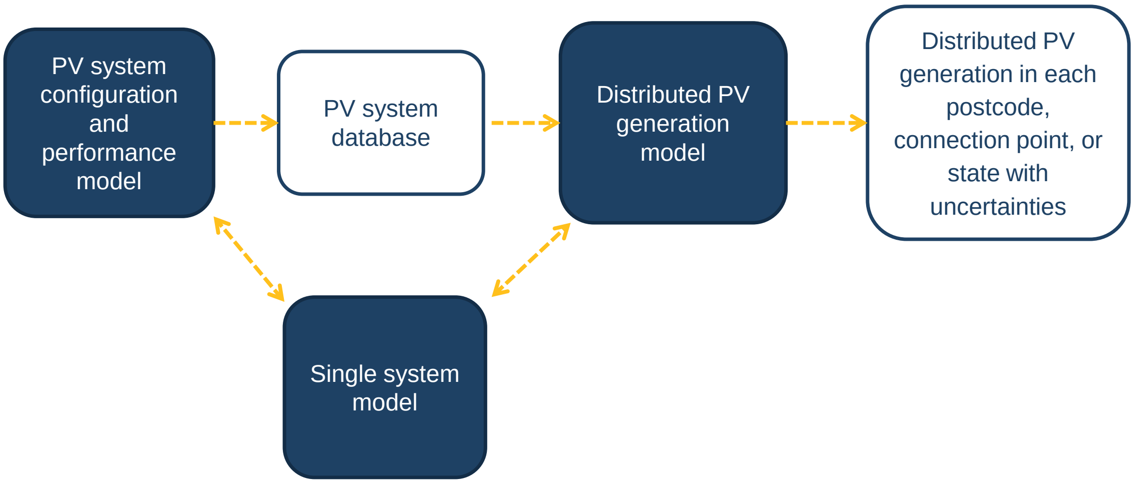

Figure 1. Overview of the rooftop PV model.

3.1 PV system configuration and performance model

The PV system configuration and performance model calculates the capacity, tilt, azimuth, and shading of PV systems using the historical generation data from that system and its matched weather data. The model was applied to generation data from over 40,000 systems to create a PV system database. The configuration and performance model was applied to each system individually. The calculation is divided in two sections: the first section calculates the configuration (installed capacity and orientation) and the second section calculates the shading pattern for the system. An overview is shown in Figure 2.

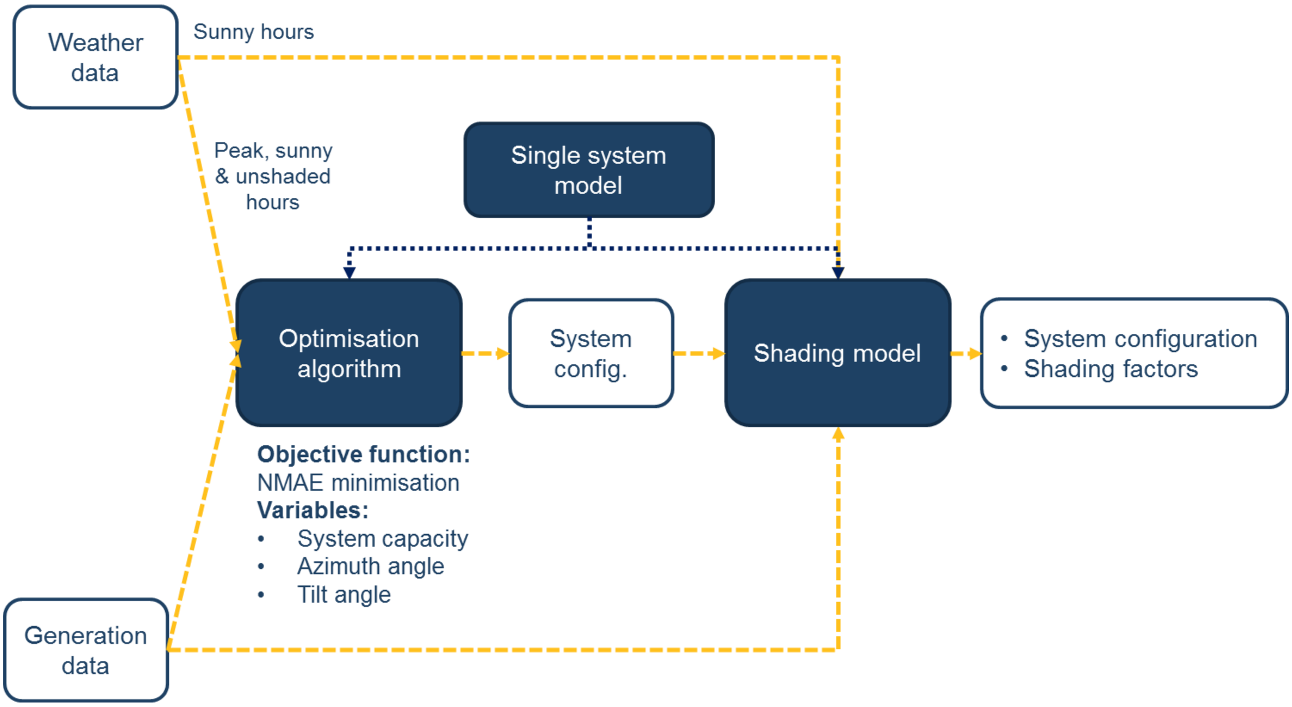

Figure 2. Overview of the PV system configuration and performance model.

3.1.1 System configuration

A customised optimisation algorithm is used to calculate the configuration of each system. This minimises the Normalised Mean Absolute Error (NMAE) between the simulated power output of the system (calculated using the single system model) and the measured power, using the installed capacity, tilt, and azimuth of the system as variables.

The optimisation algorithm uses a direct search method, which can find the minimum of a scalar function of several variables starting from an initial estimate [6]. This method is normally able to handle discontinuities (particularly if they are not near the solution), but it may find a local minimum instead of the global minimum. To avoid this, the NMAE is first calculated for 100 different combinations of the input variables, and then the best combination is used as an initial estimate for the direct search optimisation. In addition, the tilt angle is constrained to be between 0 and 90 degrees, the azimuth angle between 0 and 360 degrees, and the installed capacity between 90% and 130% of the maximum generation of the system.

The input data for the optimisation is filtered to prevent the algorithm from using times influenced by either clouds or shading. Filters were developed to select the sunny days at the location of the PV system, and then to select the times when the system is not affected by shading. The first filter selects a day as sunny day when the daily clear sky index is greater than 0.85 (i.e., the solar irradiance integrated over the day is greater than 0.85 of the clear sky irradiance integrated over the day). The second filter applies constraints to the concavity and convexity of the generation profile, and checks the number and magnitude of the profile peaks. Once the unshaded hours of the sunny days are identified, only the peak generation values of that day are used in the optimisation algorithm to limit its computation time.

3.1.2 Shading

Many rooftop PV systems are shaded at certain times from nearby hills, buildings, trees, or other objects. Shading effects are specific to the sun position (azimuth and elevation) relative to the system and are usually similar for all occurrences of a particular sun position. This is true since shading objects usually change over long (multi-year) time scales. The full effect of shading on PV generation is complex. Partial shading can produce mismatch losses which can be many times greater than the simple reduction in incident radiation [7]. A complete treatment of these effects would require detailed knowledge of each system, and is out of scope for this project. Instead, the model uses shading factors (SFs) that quantify the power drop compared to the power that the PV system would generate in unshaded conditions. Two types of SF are used in the model: SFs for direct radiation (SF_{direct}) and SFs for diffuse radiation (SF_{diffuse}):

$$DNI_{\text{shad}} = DNI \cdot SF_{\text{direct}}(\alpha, \beta)$$
$$DHI_{\text{shad}} = DHI \cdot SF_{\text{diffuse}}(SF_{\text{direct}}, k_c)$$

Shading factors are calculated separately for each sun azimuth (α) and elevation (β), with an angular resolution of 15 degrees and 3 degrees, respectively. The direct SFs are estimated by comparing the measured generation of a system to the calculated generation of a system with zero shading. The comparison is done only for sunny days, to ensure that cloudy weather conditions are not mistaken for shading. Using sunny days also ensures that the incident radiation is primarily direct radiation, and that the diffuse component can be neglected. However, for some systems, sunny times were not available for all sun positions. In this cases, the model used the average of the available shading factors with the same azimuth (i.e., above and below the missing data). These substitutions are based on the assumption that shading objects are similar along their vertical axis. Examples of direct SFs for both a lightly and a heavily shaded system are shown in Figure 3.

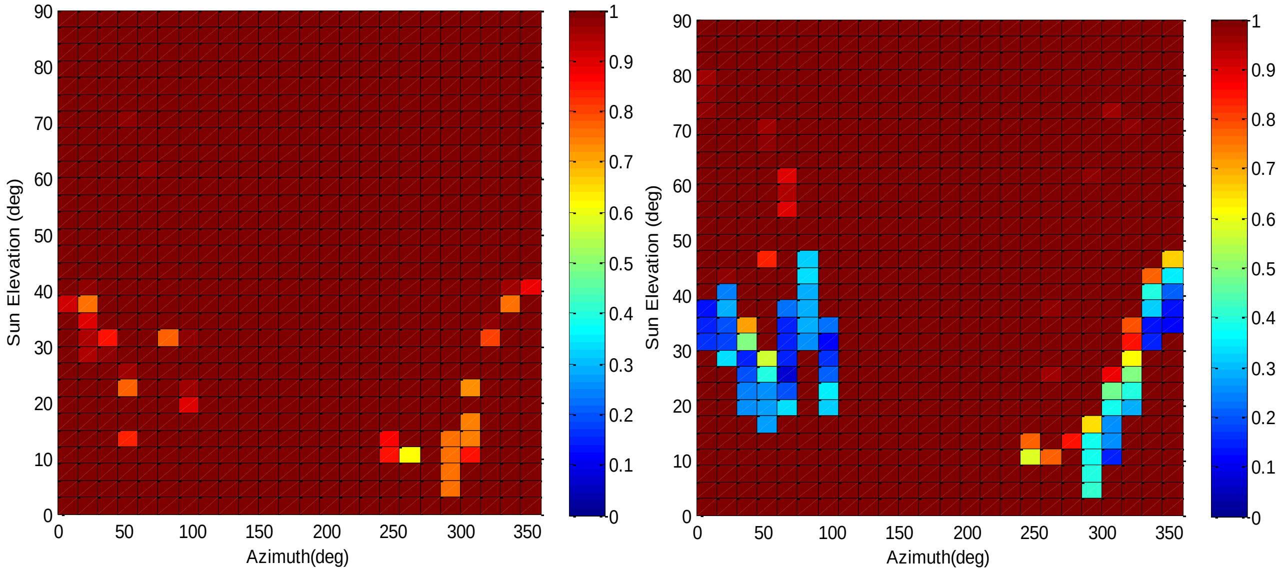

Figure 3. Examples of the shading patterns for an unshaded system (left) and a heavily shaded system (right).

An empirical relationship was determined for the diffuse SF, using the clear sky index (k_c) and the corresponding direct SF. The higher the clear sky index, the closer the value of SF_{diffuse} is to SF_{direct} (for $k_c = 1$, $SF_{\text{diffuse}} = SF_{\text{direct}}$). The relationship used was a fourth order polynomial:

$$SF_{\text{diffuse}} = SF_{\text{direct}} + (1 - SF_{direct}) \cdot \max\{0, ak_c{}^4 + bk_c{}^3 + ck_c{}^2 + \mathrm{d}k_c + \mathrm{e}\}$$

where a, b, c, d are determined empirically. Figure 4 shows this relationship for different values of SF_{direct} and k_c. Although this empirical relationship has reasonable performance, an important goal of future work should be to determine a physical shading model, using a separate treatment of the different components of the diffuse solar irradiance (circumsolar, isotropic and horizon brightening[1]) [8].

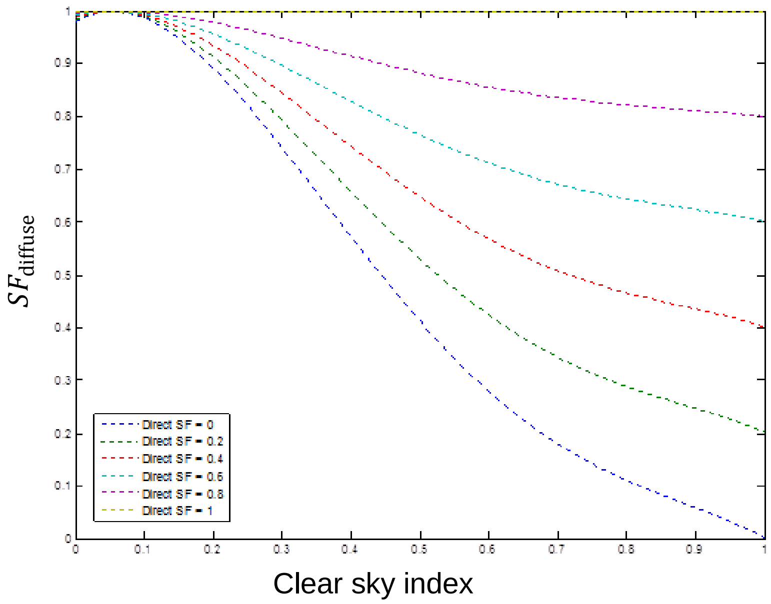

Figure 4. Diffuse SF as a function of the clear sky index and direct SF.

3.1.3 Derating factor

The model includes a derating factor which reduces the generation output calculated by the single system model. A derating factor is commonly used in PV modelling to account for effects such as soiling, wiring losses, and differing efficiencies compared to the reference system [9, 10]. The derating ensures that the calculated system capacity matches the rated capacity of the system. This is important since the distributed PV generation model relies on PV capacity data from the Clean Energy Regulator (CER). The derating was calibrated by matching the capacity calculated by the model to the official capacity listed on PVOutput. Only PVOutput systems were used because the official rated capacity was not available for the systems with generation data in AEMO's database.

Before this calibration was done, the listed capacity of PVOutput systems was validated against the maximum generation reported by the system. The listed capacity of a system on PVOutput is manually entered by the user when registering a system, so it is possible for errors to occur in data entry, or because the users have only an approximate knowledge of their systems. Systems where the maximum generation was greater than 110% of the listed capacity or less than 60% of the listed capacity were excluded from the calibration (around 2% of the total). From this calibration, a derating of 8% was calculated. Data from the calibration process is shown in Figure 5. The differences between the official listed capacity and the calculated capacity is due to both uncertainties in the configuration optimisation, and to different system efficiencies compared to the reference system used in the single system model (see section 3.3). It is difficult to separate these two effects, but the derating factor ensures that the official listed capacity and calculated capacity agree on average, which is of key importance when calculating distributed PV generation.

[1] The isotropic component is received uniformly from the entire sky dome. The circumsolar component results from forward scattering of sunlight through the atmosphere, and therefore propagates in approximately the same direction as the beam irradiance. The horizon brightening component is concentrated in a band near the horizon and is most pronounced in clear skies.

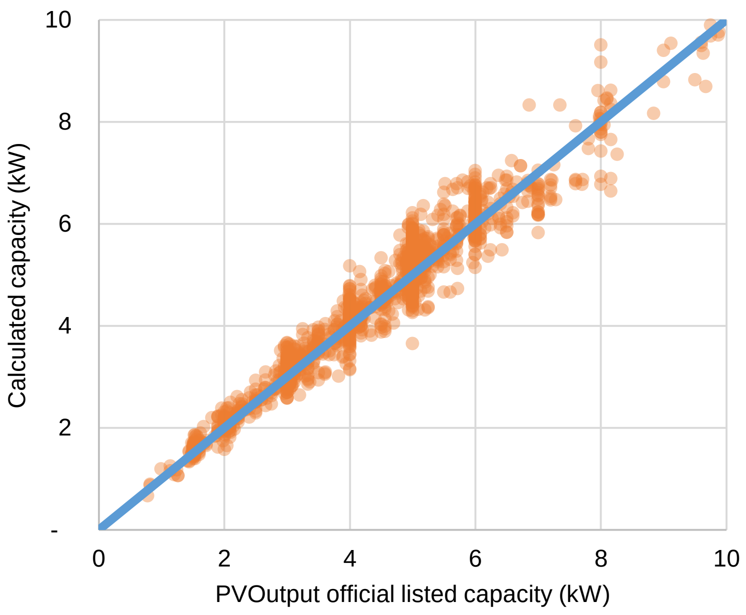

Figure 5. Calculation of the model derating.

3.2 Distributed PV generation model

The distributed PV generation model calculates the generation and uncertainty in each postcode, connection point, or state, on a half hourly basis, using input weather data and the PV system database. The model has been designed to calculate either historical or forecast PV generation, depending on whether historical or forecast weather data is provided as input. An overview is shown in Figure 6. A random sampling and aggregation method is used to calculate the PV generation in each area. An uncertainty model calculates the 10% and 90% POE levels, using the weather data and historical generation traces. An example of the model output is shown in Figure 7.

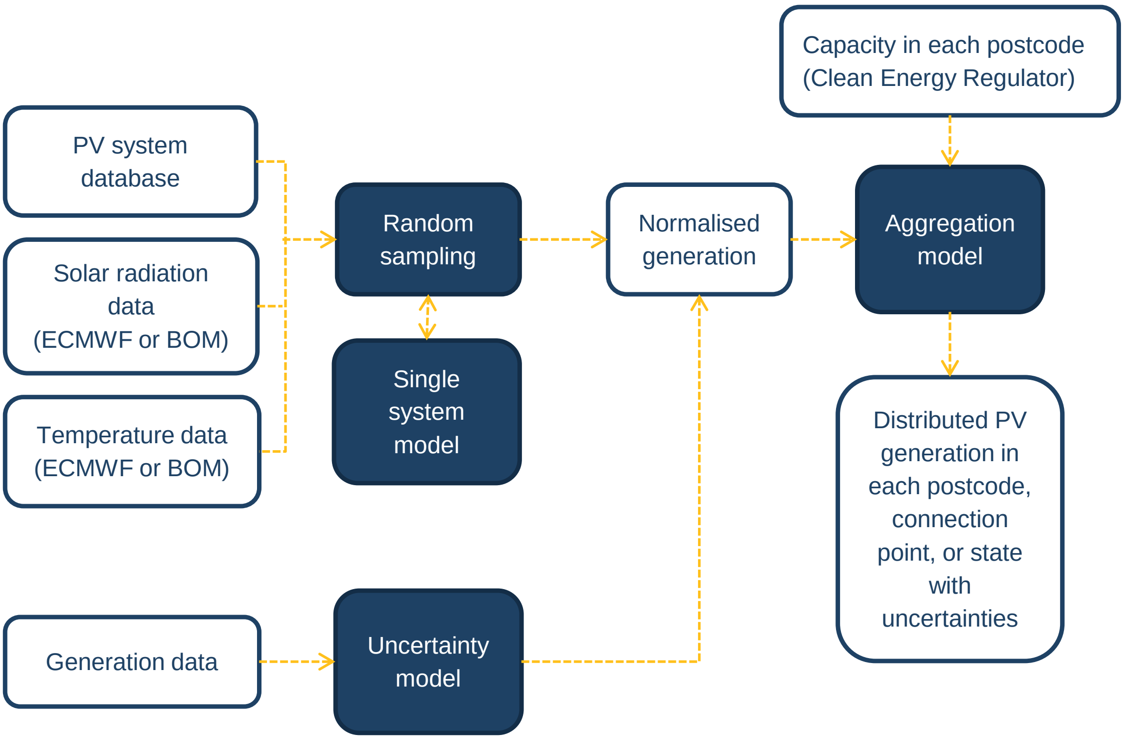

Figure 6. Overview of the distributed PV generation model.

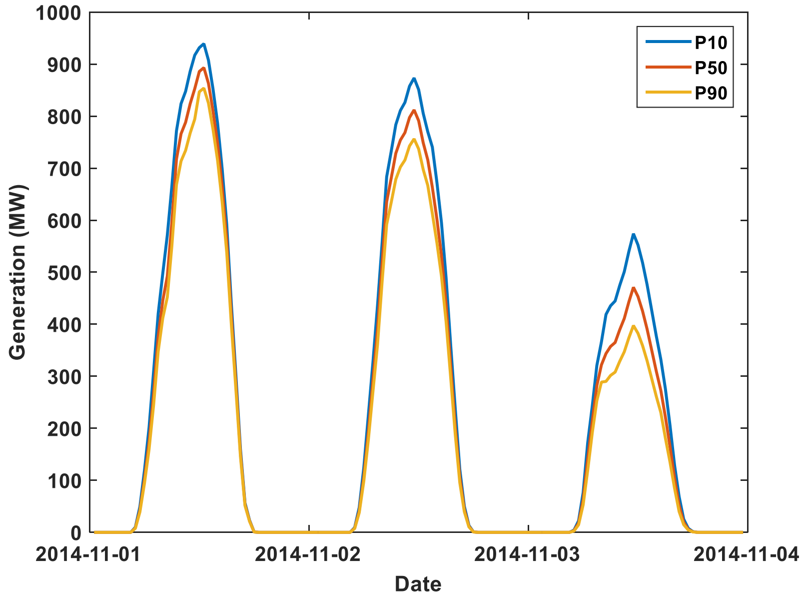

Figure 7. Example model output, showing aggregate generation for Queensland rooftop PV systems.

3.2.1 Random sampling

PV generation in any area depends on the local weather and the configuration and performance of the PV systems in that area. Configuration data for all the systems in the NEM is not available, and even if such data were available, it would not be computationally tractable to separately calculate the generation from the more than one million individual systems in the NEM. Therefore, similar to other studies, the model calculates the aggregate generation from the distributed PV systems in a particular area (such as a postcode) using a representative ensemble of systems [9].

To validate this approach, it was important to determine the ensemble size at which the model output and performance converges, i.e., is independent of further increases in size. To do this, an analysis was done for the postcodes for which a large number of systems with generation data were available (greater than 1,000 systems). For each of these postcodes, random ensembles of different sizes were taken from the system database. For each random ensemble, the generation of each system in the ensemble was calculated using its configuration data (capacity, tilt, azimuth, and shading). The aggregate generation of the ensemble was calculated, and the NMAE calculated with respect to the measured generation of all systems. As shown in Figure 8, the NMAE quickly reduced as the ensemble size was increased. Above 100 systems, the NMAE became independent of ensemble size, and converged to the NMAE of the aggregate calculated generation of all systems.

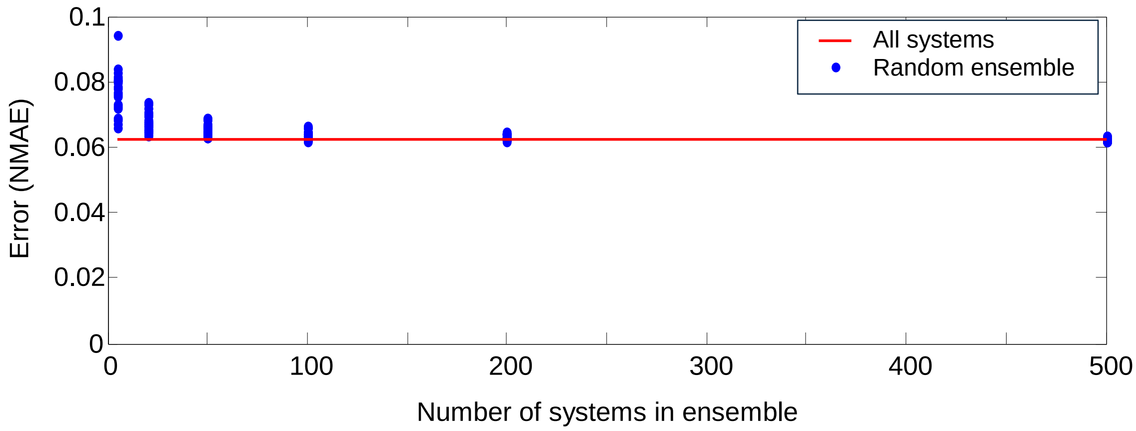

Figure 8. NMAE compared to the number of systems in each random ensemble.

3.2.2 Uncertainty model

The primary source of uncertainty in the calculation of the distributed PV generation is the input solar irradiance. As ground measurements of solar radiation are not available for all locations in the NEM, the use of satellite-based estimates and Numerical Weather Prediction (NWP) models introduce errors to the PV generation calculation. These errors, and hence rooftop PV model errors, are greater in overcast and cloudy sky conditions. There is therefore a relationship between the uncertainty of the PV generation and the metric which describes these conditions, the clear sky index. Other studies of PV generation have also used the clear sky index in combination with the sun elevation [11]. The clear sky index k_c is defined:

$$k_c = \frac{\mathrm{GHI}}{\mathrm{GHI}_{clear\ sky}}.$$

Model uncertainty is calculated as 10% and 90% POE levels, which were calibrated and validated directly from the system generation data. The uncertainty model is based on the relationship between the fractional error of the PV generation estimate and the clear sky index. Figure 9 shows how the fractional error varies as a function of the clear sky index.

The POE levels are calculated by grouping the clear sky index in bins and calculating the 10th and 90th percentiles of the fractional error for each bin. From this, the 90% and 10% POE levels for each half hour are calculated using the clear sky index (k_c) for that half hour:

$$90POE_t = \frac{F_t}{1 + P_{10}(k_c)}$$

$$10POE_t = \frac{F_t}{1 + P_{90}(k_c)}$$

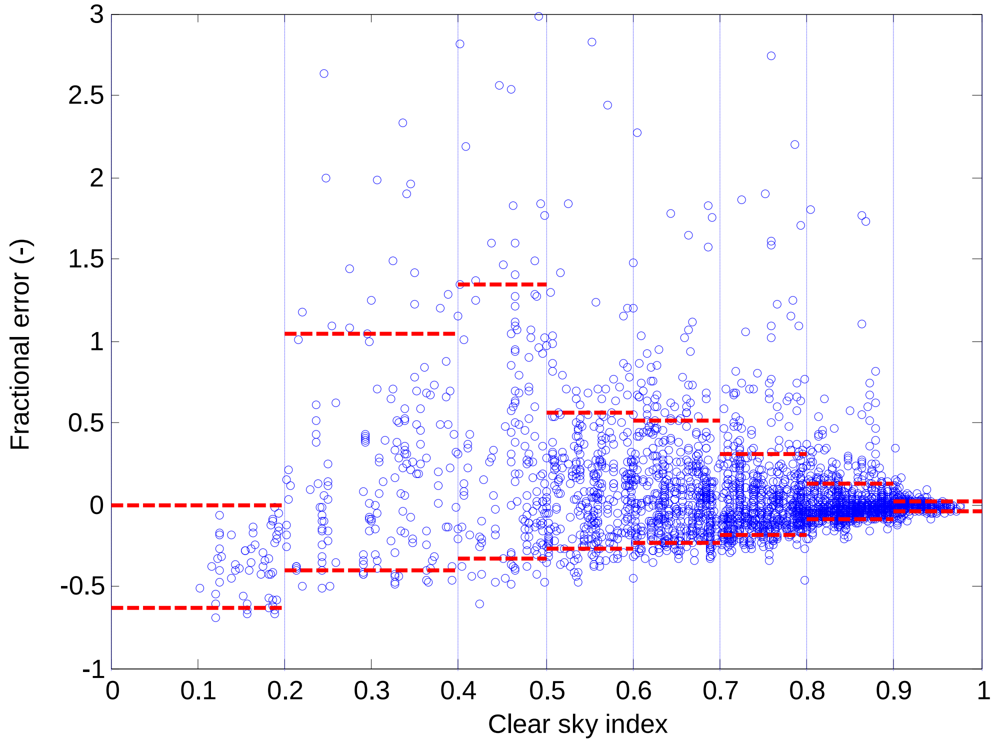

Figure 9. Uncertainty (fractional error) as a function of clear sky index.

3.3 Single system model

The single system model calculates the power output of an individual PV system using the weather data and system configuration as inputs. This is a core function used in both components of the rooftop PV model described in sections 3.1 and 3.2. It contains subroutines that allow accounting for shading effects, panel efficiency, inverter losses, etc. To build this function, several subroutines from the PV_LIB toolbox were used (see section 4.4.1). The scheme in Figure 10 shows the structure and input variables of the single system model.

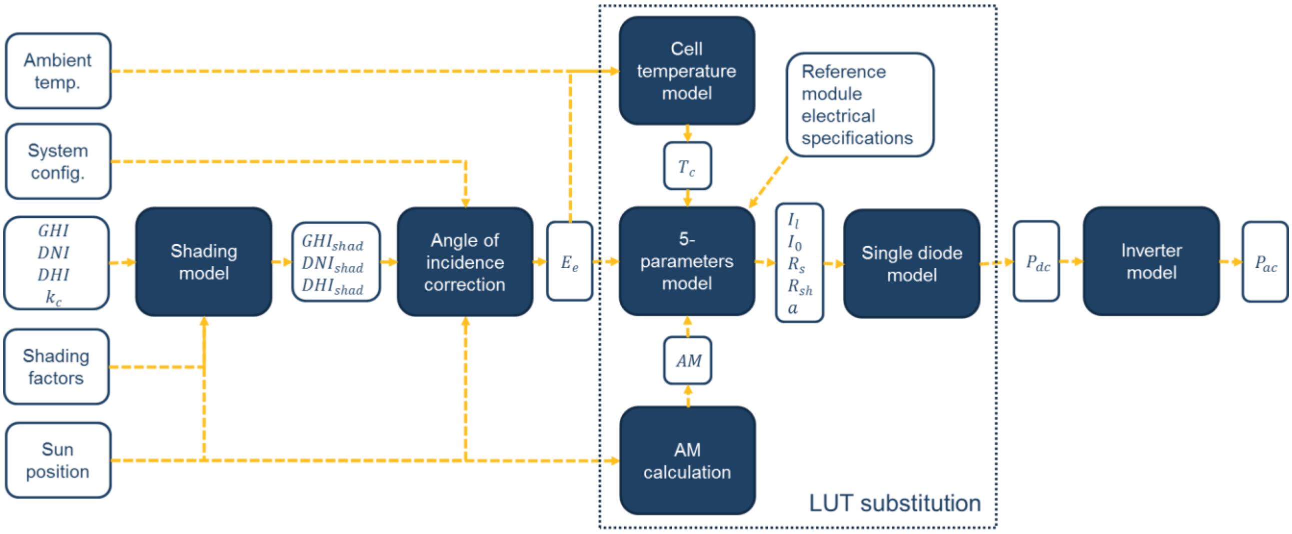

Figure 10. Scheme of the structure and the calculations of the single system model.

The initial subroutines of the single system model aim to calculate the incident irradiance on the plane of the array. First, the direct and diffuse SFs are used to correct the Global Horizontal Irradiance (GHI), Direct Normal Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI) at the system location (see section 3.1.2). Note that the irradiance is assumed to be uniformly distributed across the arrays. Any mismatch losses due to partial shading are included implicitly in the calculation of the shading factors. Then, the orientation of the system is used to calculate the incident beam irradiance according to the beam angle of incidence. In parallel, different models are used to estimate the sky and ground reflected components of the incident diffuse radiation.

The cell temperature also has a significant impact on the efficiency of PV panels. To account for this effect, the cell temperature is calculated using the simplified equation shown below, where T_a is the ambient temperature and *k* is a dimensionless parameter [12].

$$T_c = T_a + k \cdot E_e$$

More sophisticated methods to calculate the cell temperature are available in the literature [13], but they were found to introduce further complexity without significantly improving the final accuracy of the model.

Once the total incident irradiance (E_e) and cell temperature are calculated, the single system model uses the single diode model to calculate the power output of a PV panel [14]. This model uses the equivalent circuit represented in Figure 11 to simulate the current–voltage curve of a PV module. The five parameters (I_l, I_0, R_s, R_{sh}, and a) defining the equivalent circuit were calculated using the method introduced by Desoto et al. [15], known as the 5-parameters model. This algorithm estimates the five parameters from module specific electrical parameters (such as the open circuit voltage, the short circuit current, the maximum power current and voltage, or the temperature coefficients at open circuit voltage and short circuit current), the absorbed irradiance[2] (S), the cell temperature (T_c) and the absolute air mass (AM).

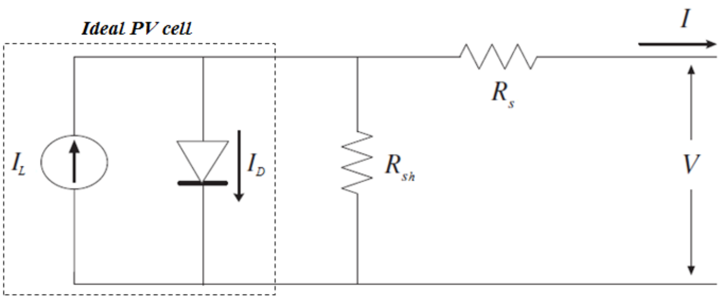

Figure 11. Equivalent circuit used in the 5-parameter or single diode model.

Details are not available about the solar cell technology, or the make and model of the panels of each PV system. Therefore, the electrical parameters of a polycrystalline module of representative efficiency were assumed for all systems. However, it may be possible for the CER to provide this information, and further analysis of this topic should be done as future work.

In addition, assuming the same type of module for all the systems made it possible to replace many of the computationally expensive calculations of the single diode model with a look up table (LUT). The LUT stores pre-calculated power output values as a function of three variables: the incident irradiance, the ambient temperature and the sun elevation. Then, values for specific query points are calculated using linear interpolation. This is important because the single system model is called a large number of times in the optimisation that calculates the system configuration and the calculation of the aggregate PV generation at each postcode. This change allowed reducing the computing time of the single system model by more than 10 times without a significant loss of accuracy.

Finally, a subroutine accounts for inverter losses occurring during the conversion of the DC power from the PV modules to AC power, with an efficiency related to the load of the inverter. The

[2] The absorbed irradiance is assumed to be equal to the incident irradiance.

specifications of a representative inverter were also assumed for all systems, and the inverter nominal capacity was scaled up or down according to the capacity of the each PV system.

4. Input data

4.1 Weather data

The model uses the GHI, DNI, DHI and ambient temperature as meteorological inputs to calculate the rooftop PV generation. The GHI and the ambient temperature are provided by a weather provider, while the DNI is estimated from the GHI using the "DIRINT" model [16] and the DHI is derived from the equation relating these three components (where θ_z is the solar zenith angle):

$$GHI = DHI + DNI \cdot \cos(\theta_z)$$

The rooftop PV model was originally developed using the weather data from the European Centre for Medium-Range Weather Forecasts (ECMWF) [17] and then was adapted to use Australian Bureau of Meteorology (BOM) solar irradiance data derived from hourly satellite observations [18].

4.1.1 ECMWF

ECMWF is an independent intergovernmental organisation which provides medium-range solar irradiance forecasts. This data contains gridded forecasts of meteorological variables and surface parameters such as irradiance, wind, humidity, height, etc. The grid covers the whole territory within the NEM and has a horizontal resolution of about 25 km (0.25 x 0.25 degrees latitude longitude). The forecasts are released every 12h and have a 3-hourly resolution for the first 3 days, increasing to 6-hourly from day 4 to 10.

As the rooftop PV model is required to generate half-hourly calculations, an interpolation of the meteorological data was required. This interpolation was done using the clear sky irradiance calculated with the method presented by Ineichen and Perez [19]. Figure 12 shows how the aggregated values from ECMWF have been interpolated into half-hourly values using the clear sky GHI.

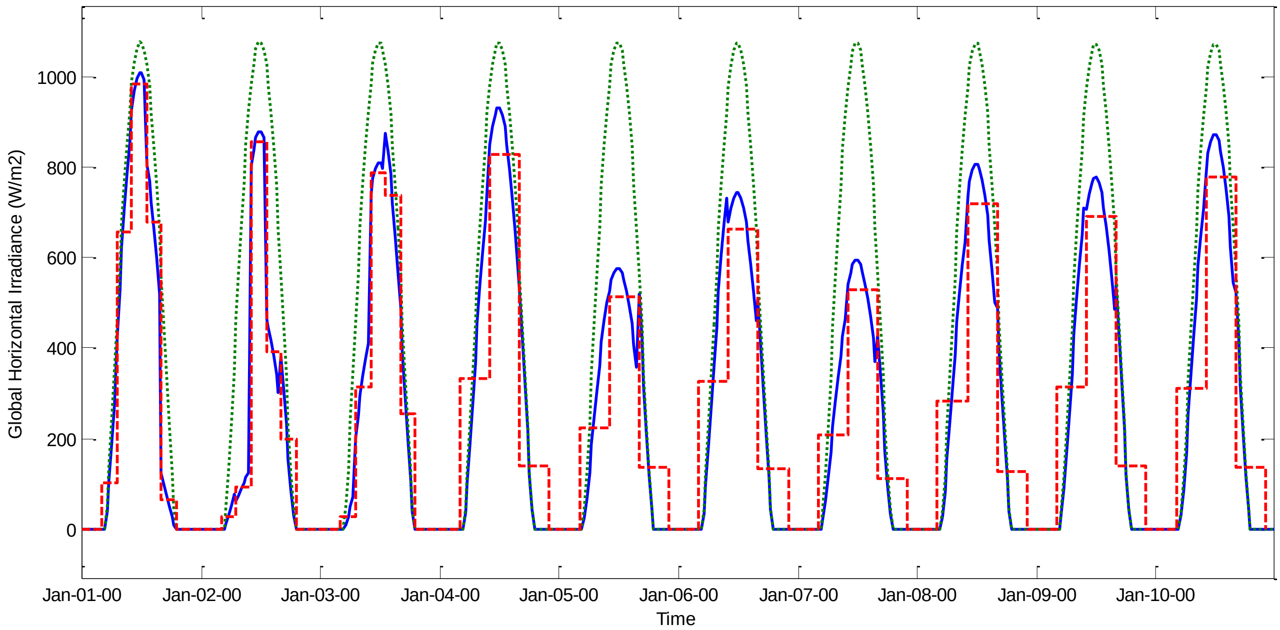

Figure 12. GHI interpolated value (blue), GHI aggregated value from ECMWF (red) and clear sky GHI (green).

The GHI interpolated value (GHI_{interp}) is calculated from a clear sky factor ($\lambda_{\text{clear\_sky}}$) which is then applied to the GHI aggregated value (GHI_{aggreg}) provided by ECMWF. The clear sky factor indicates the weighting of the clear sky GHI for the half-hour time period t compared to the aggregated time period x (which can be 3 or 6 hours according to the forecast horizon).

$$\lambda_{clear\_sky_t} = \frac{2x \cdot GHI_{clear\_sky_t}}{\sum_{t \in x}\left(GHI_{clear\_sky_t}\right)}$$

$$GHI_{interp_t} = GHI_{aggreg_t} \cdot \lambda_{clear\_sky_t}$$

4.1.2 BOM

In order to calculate the historical PV generation from 2000 to 2016, solar radiation and temperature data from the Australian Bureau of Meteorology (BOM) was also used, since the ECMWF data was not available prior to 2013. For this, historical GHI data derived from hourly satellite observations with a 5km × 5km spatial resolution was used [18]. An example of the BOM GHI data is shown in Figure 13. Hourly temperature data from the BOM was also used. Both the hourly GHI and temperature data were interpolated to a half hourly time basis to calculate historical PV generation.

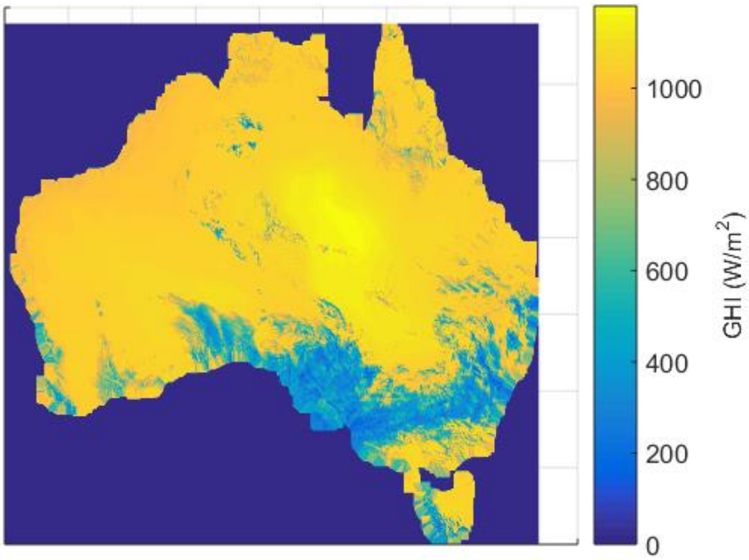

Figure 13. Example BOM GHI data.

4.2 Historical PV generation data

Historical PV generation traces were obtained from two sources: AEMO's database and the online service PVOutput. Most of the analysed systems are from the AEMO database, where gross metered generation traces for around 38,000 systems were available. However, most of these systems were located in NSW. PVOutput provided around 2,000 additional systems, which were more widely distributed across the NEM.

State	AEMO databases	PVOutput	All data sources
NSW (including ACT)	37,592	432	38,024
VIC	529	309	838
QLD	39	965	1,004
SA	12	226	238
TAS	4	62	66
TOTAL	**38,176**	**1,994**	**40,170**

Table 1. Total number of systems for which gross metered data is available in each state

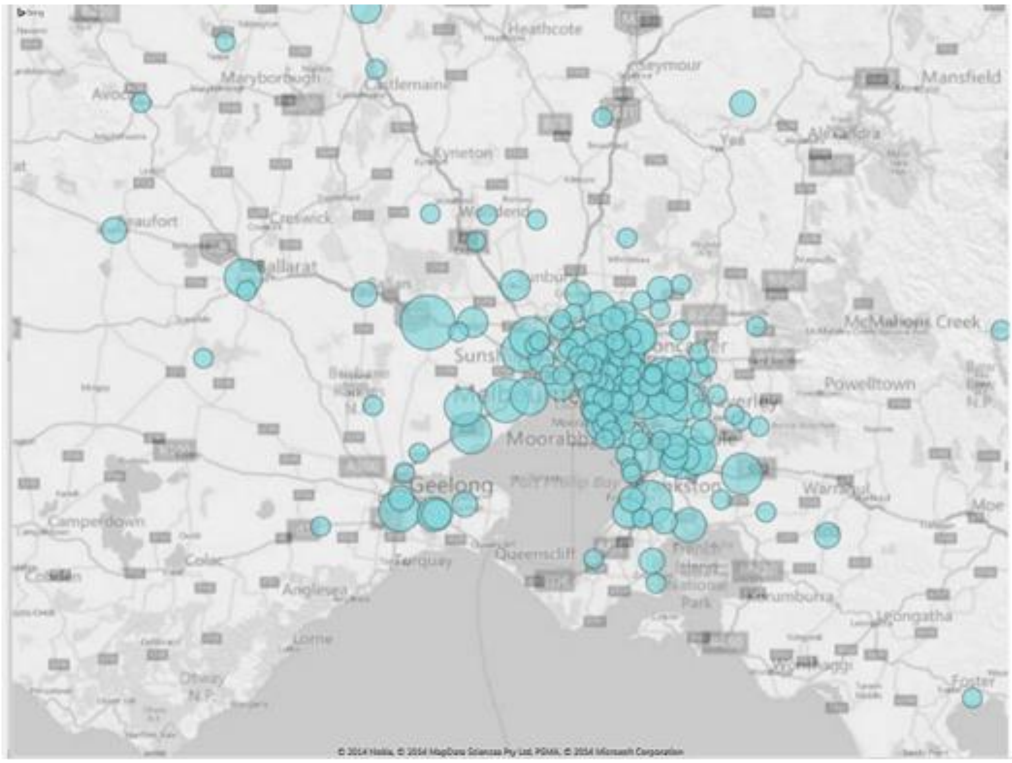

Figure 14. Example of the location of the systems with available historical PV generation traces (VIC)

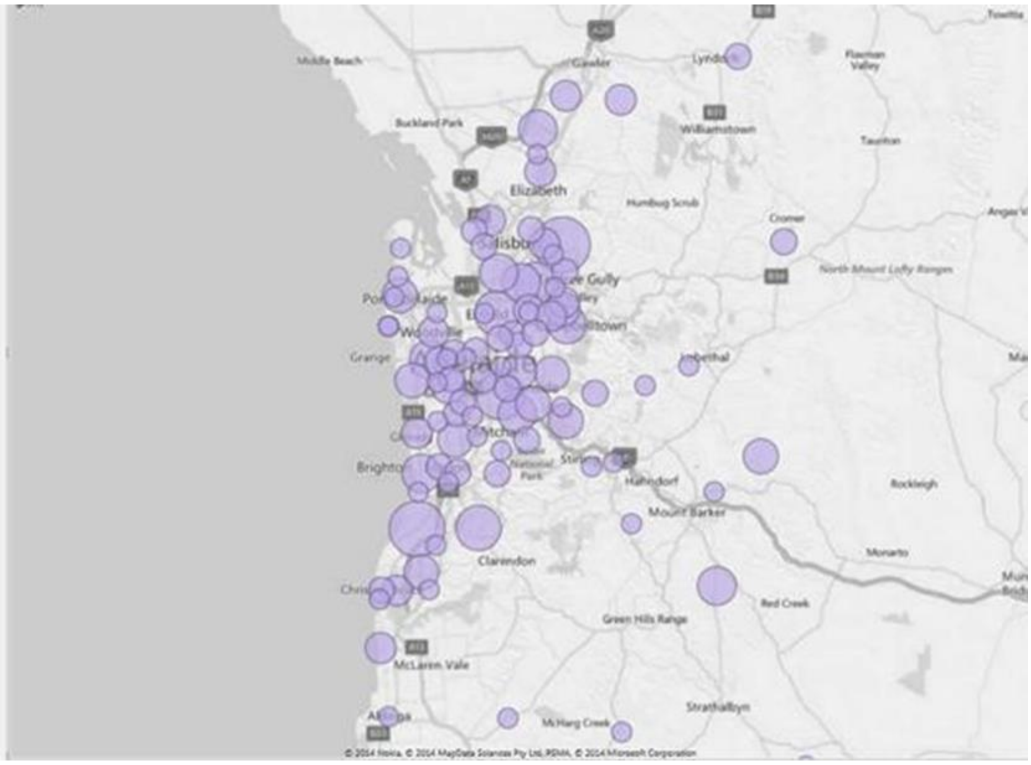

Figure 15. Example of the location of the systems with historical PV generation traces (SA)

4.3 Installed capacity data

The rooftop PV capacity installed in each postcode was taken from the CER database. As part of the Small-scale Renewable Energy Scheme (SRES) of the Renewable Energy Target (RET), the CER publishes every month an update on the number and capacity of rooftop PV systems, installed in each postcode in each month of the previous 12 months [20]. For each system, there can be a delay of up to 12 months between the installation of the system and its appearance in the CER data. This is due to the fact that participants in the scheme are permitted up to 12 months to register the system and claim the Small-scale Technology Certificate (STC) subsidy. For this reason, when using CER data published on a particular date, a scaling factor is applied to the capacity installed in the recent months prior to that date. These scaling factors were taken from an Australian PV Institute study of the CER reporting data [21], and are shown in Table 2. For example, it is assumed that only 50% of systems installed one month or less prior to the CER publication date have been reported. Five months after installation, it is assumed that effectively all systems have been reported.

Date prior to CER data publication date (months)	Fraction of systems assumed to have been reported	PV capacity scaling factor
1	50%	2.0
2	75%	1.3
3	83%	1.2
4	90%	1.1
5 and greater	~100%	1.0

Table 2. Scaling of CER postcode capacity data to account for systems installed but not yet registered

4.4 Software

4.4.1 PV_LIB toolbox

After reviewing the literature about the existent methods and tools used to model PV arrays, the PV_LIB Toolbox was selected and its accuracy tested against measurements. This is an open source library which contains a set of well-documented functions and example scripts for modelling PV systems. The toolbox was created jointly by Sandia National Laboratories and the PV industry. It contains a wide range of functions that allow the calculation of the incident radiation, cell temperatures, power output, inverter efficiency, mismatch losses, etc. It uses the system configuration and the weather data (Global Horizontal Irradiance, ambient temperature, wind speed and atmospheric pressure) as inputs. It contains an implementation of King's model [22], which uses a database of empirically derived PV module parameters, or the single diode array performance model [14]. In this project the single diode array performance model was used.

5. Glossary

- **Air Mass (AM)**: relative measure of the optical length of the atmosphere at sea level. It is expressed as the ratio of light's path length through the atmosphere relative to the path length which is vertically upwards. The relative air mass quantifies the loss of power that light experiences as it is absorbed by atmospheric gases (like ozone), water vapour and aerosols.
- **Absolute Air Mass:** refers to the relative air mass multiplied by the local atmospheric pressure and divided by the standard pressure at sea level, which accounts for weather effects and elevation.
- **Azimuth angle**: clockwise angle between the surface of the module and the North direction (N = 0 degrees, E = 90 degrees, S = 180 degrees, W = 270 degrees). The optimal angle for fixed solar PV systems is 0 degrees in the southern hemisphere, and 180 degrees in the northern hemisphere. Note that it may also refer to the sun's azimuth angle.
- **Clear sky radiation**: terrestrial solar radiation under cloudless conditions. This is a key variable in solar modelling for which several models have been presented in the literature [23]. In this case, the model from Ineichen and Perez was used to calculate the clear sky GHI [19].
- **Clear sky index (k_c)**: Ratio between the ground incident GHI and the clear sky GHI. Very clear days take values close to 1 and very cloudy days take values close to 0. It is calculated as follows:

$$k_c = \frac{\mathrm{GHI}}{\mathrm{GHI}_{clear\ sky}}$$

- **Derating factor**: all-rounder factor used in solar PV modelling which multiplies directly the calculated generation to take into account effect such as soiling, wiring losses, snow cover, inaccuracy of the manufacturer's nameplate rating, etc. It is often used in the calibration of models.
- **Diffuse Horizontal Irradiance (DHI)**: terrestrial irradiance received by a horizontal surface which has been scattered or diffused by the atmosphere. It is the component of the GHI which does not come from the beam of the sun.
- **Diffuse shading factor**: Factor between 0 and 1 multiplying the diffuse component of the solar irradiance at the PV array location. It is modelled as a function of the clear sky index and the direct shading factor. It is always greater or equal to the direct shading factor.
- **Direct Normal Irradiance (DNI)**: terrestrial irradiance arriving from the Sun's direct beam on a plane perpendicular to the beam. To compare the DNI to the GHI and the DHI, it is necessary to obtain its horizontal component by multiplying it by the cosine of the sun's zenith angle (θ_z).
- **Direct shading factor**: Factor between 0 and 1 multiplying the direct component of the solar irradiance at the PV arrays location. A value is assigned to direct SFs for each particular sun position (azimuth and elevation angles).
- **European Centre for Medium-Range Weather Forecasts (ECMWF)**: Weather data provider from which irradiance and temperature variables used in the model are obtained. Link: http://www.ecmwf.int/.
- **Fractional Error (FE)**: metric that quantifies the relative error by which the calculated power from the model ($\bar{P}_t$) differ from the measured power (P_t) at a particular time t. The fractional error for a time period t is calculated as the error, $\bar{P}_t - P_t$, divided by the measured value:

$$FE_t = \frac{\bar{P}_t - P_t}{P_t}$$

- **Global Horizontal Irradiance (GHI)**: total terrestrial irradiance falling on a horizontal surface. It can be measured or it can be calculated from the DNI and the DHI using the following equation:

$$GHI = DHI + DNI \cdot \cos(\theta_z)$$

- **Gross metering**: metering method in which the generation and the consumption are measured separately.
- **Net metering**: metering method in which consumption from the grid and PV generation are measured together. Only the energy flow (import/export) to or from the grid is recorded.
- **Normalised bias (NBIAS)**: metric that quantifies the deviation between the average calculated power from the model and the average measured values in a time set. This deviation, known as bias, is then normalised against the PV system capacity (P_0). This metric quantifies the persistent tendency of the model to underestimate or overestimate the measured power.

$$NBIAS = \frac{1}{N}\sum_{t=1}^{N}\frac{(\bar{P}_t - P_t)}{P_0}$$

- **Normalised Mean Absolute Error (NMAE)**: metric that quantifies the average magnitude of the errors between the calculated power from the model ($\bar{P}_t$) and the measured power (P_t). It is normalised against the PV system capacity (P_0). Note that only daylight hours are included in the calculation.

$$NMAE = \frac{1}{N}\sum_{t=1}^{N}\frac{|\bar{P}_t - P_t|}{P_0}$$

- **Probability of Exceedance (POE)**: metric that quantifies the probability, as a percentage, that a PV generation level will be met or exceeded in a given time step. The 10% and 90% POE levels together give an indication of the uncertainty related to the calculation. In the rooftop PV model this uncertainty is directly related to the clear sky index.
- **PVOutput**: free online service where users can upload the generation of their systems for sharing and comparing PV systems output data. It provides both manual and automatic data uploading facilities. Link: http://www.pvoutput.org/.
- **Shading pattern**: Matrix in which the shading factors are stored as a function of the sun position (sun azimuth and elevation angles).
- **Terrestrial irradiance**: Earth ground incident irradiance.
- **Tilt angle**: angle between the horizontal plane and the plane of the array. Ideally, the tilt angle should be modified along the year to adapt to the differences of the sun elevation between summer and winter. For fixed frame installations, to maximize the annual energy yield, it is accepted as a general rule to set a tilt angle approximately equal to the site's latitude.

6. References

[1] Australian Energy Market Operator, “National Electricity Forecasting Report,” 2016.

[2] Australian Energy Market Operator and Electranet, “Renewable Energy Integration in South Australia - Joint AEMO & Electranet Report,” 2014.

[3] Australian Energy Market Operator, “National Electricty Forecasting Report,” 2015.

[4] G. Colantuono, A. Everard, L. Hall and A. Buckley, “Monitoring nationwide ensembles of PV generators: Limitations and uncertainties. The case of the UK,” *Solar Energy,* vol. 108, p. 252–263, 2014.

[5] Sandia National Laboratories, “PV_LIB Toolbox,” PV Performance Modeling Collaborative, [Online]. Available: pvpmc.sandia.gov.

[6] . J. Lagarias, J. Reeds, M. Wright and P. Wright, “Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions,” *SIAM Journal of Optimization,* vol. Vol. 9 Number 1, pp. 112-147, 1998.

[7] C. Deline, “Partially Shaded Operation of a Grid-Tied PV System,” in *34th IEEE Photovoltaic Specialists Conference*, Philadelphia, 2009 .

[8] R. Perez, R. Seals, P. Ineichen, R. Stewart and D. Menicucci, “A new simplified version of the perez diffuse irradiance model for tilted surfaces,” *Solar Energy,* vol. 39, no. 3, pp. 221-231, 1987.

[9] Lorenz, et al., “Regional PV power prediction for improved grid integration,” *Progress in Photovoltaics Research and Applications,* vol. 19, no. 7, pp. 757 - 771, 2011.

[10] ROAM, “Wind and Solar modelling for AEMO 100% Renewables project,” 2012.

[11] Lorenz, et. al, “Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 2, no. 1, pp. 2-10, 2009.

[12] R. Ross, “Interface design considerations for terrestrial solar cell modules,” in *12th IEEE photovoltaic specialists conference*, Baton Rouge, LA, 1976.

[13] E. Skoplaki and J. Palyvos, “Operating temperature of photovoltaic modules: A survey of pertinent correlations,” *Renewable Energy,* vol. 34, pp. 23-29, 2009.

[14] J. A. Duffie and W. A. Beckman, Solar Engineering of Thermal Processes, 4th Edition, New York: John Wiley & Sons Ltd, 2013.

[15] W. De Soto, S. Klein and W. Beckman, “Improvement and validation of a model for photovoltaic array performance,” *Solar Energy,* vol. 80, p. 78–88, 2006.

[16] R. Perez, P. Ineichen, E. Maxwell, R. Seals and A. Zelenka, “Dynamic Global-to-Direct Irradiance Conversion Models,” *ASHRAE / Transactions - Research Series,* pp. 354-369, 1992.

[17] European Centre for Medium-Range Weather Forecasts (ECMWF), “Set I - Atmospheric Model high resolution 10-day forecast (HRES),” July 2016. [Online]. Available: http://www.ecmwf.int/en/forecasts/datasets/set-i.

[18] Bureau of Meteorology, “Australian Hourly Solar Irradiance Gridded Data,” [Online]. Available: www.bom.gov.au/climate/how/newproducts/IDCJAD0111.shtml.

[19] P. Ineichen and R. Perez, “A new airmass independent formulation for the Linke turbidity,” *Solar Energy,* vol. 2, pp. 151-157, 2002.

[20] Clean Energy Regulator, “Small-scale installations by postcode,” [Online]. Available: http://ret.cleanenergyregulator.gov.au/REC-Registry/Data-reports.

[21] A. P. Institute, “Interactive Australian PV Map Report,” 2013.

[22] D. King, W. Boyson and J. Kratochvill, “Photovoltaic Array Performance Model,” Sandia National Laboratories, Albuquerque, 2004.

[23] M. J. Reno, C. W. Hansen and J. S. Stein, “Global Horizontal Irradiance Clear Sky Models: Implementation and Analysis,” Sandia National Laboratories, Albuquerque, New Mexico & Livermore, California, 2012.