Conference Proceedings

Solar World Congress 2015
Daegu, Korea, 08 – 12 November 2015

Real-time Simulations of 15,000+ Distributed PV Arrays at Sub-Grid Level using the Regional PV Simulation System (RPSS)

Nicholas A. Engerer[1], James Hansard[2]

[1] The Fenner School of Environment and Society, The Australian National University (Australia)

[2] College of Engineering and Computer Science, The Australian National University

Abstract

The Regional PV Simulation System (RPSS), recently jointly developed by The Australian National University and National ICT Australia (NICTA), has been created to assist in the integration of high penetrations of small-scale, distributed solar PV energy systems. This system is capable of simulating many thousands of small-scale PV systems with input from only a small number (~1%) of the total installed fleet. This distributed solar PV simulation system is the first step towards delivering forecasts of their collective power output, and will serve as a very useful tool to electrical utilities and energy markets. Through cooperation with the local distribution network service provider, ActewAGL, the RPSS has been deployed in the Canberra, Australia region delivering real-time simulations of distributed PV power output as grouped by transformer. Canberra, the capital city of Australia, and has some of the highest penetrations of solar PV in the country, and has set a renewable energy target of 90% by 2020, making it a prime candidate for the application of these real-time simulations. Herein, the methodology used in this real-time version of the RPSS are presented, as well as a case study which highlights its usefulness under strong variability in distributed PV system power output.

Keywords: *Solar Energy, Distribution, Grid Integration, Simulation, Solar PV*

1. Introduction

As the number of distributed photovoltaic arrays installed worldwide continues to grow, so does the need for real-time simulations systems that are capable of assessing their impact on the electrical grids to which they connect. In the case of micro-generators, which are defined here as small-scale PV systems with rated capacities less than 100 kW, this is often extraordinarily challenging, due to the lack of information available about their instantaneous power output. In Australia, over 1.4 million individual systems have been installed, the vast majority of which are unmonitored, and therefore their interval-level contributions to the electrical grid are unknown. Their total installed PV capacity now exceeds 4.5 GW as of June 2015, with an average system size of 4.5 kW (APVI 2015). This represents substantial growth from only 41 kW of capacity at the start of 2009. As a result of this impressive uptake, 20% of Australian households now have a solar PV system installed at their residence (ABS 2014). Looking forward, the installed cost of PV is approximately $1.75/W for systems >10 kW, and thus continued strong uptake is anticipated, particularly in commercial site applications (APVI 2015). Moderate uptake scenarios from Australian solar analytics company SunWiz project installed distributed PV capacity to exceed 7 GW by 2020 (SunWiz 2015). For perspective, Australia's largest electricity market, the National Electricity Market (NEM), has a total large-scale generation capacity of approximately 50 GW.

What these statistics represent is that distributed small-scale solar installations represent a fast growing and increasingly robust energy generation source in Australia. Its penetrations of PV are among the highest in the world, with cost parity being reached throughout the continent (Chen and Franklin 2011). However,

doi:10.18086/swc.2015.06.02 Available at http://proceedings.ises.org

these developments place Australia on the forefront of the challenges associated with high-penetration distributed PV integration into the electricity grid. At the time of writing, Australian utilities had already begun to ban grid export or limit the total installed capacity permitted per individual system (Ergon Energy 2015, Horizon Power 2012). This represents a reactive, rather than proactive response to the intermittency challenges faced by distributed solar resources, which have been documented to induce intolerable voltage fluctuations within Australian distribution networks (Noone 2013).

One of the key barriers to increasing the penetrations of distributed PV in these areas is lack of information regarding the real-time power output of these distributed PV generators. Without this information, utilities will remain unable to impose mitigative solutions via emerging technologies such as distributed energy storage systems, because they will remain naïve on where and when large changes in solar power output will occur. This has necessitated the creation of a simulation system that is able to produce real-time estimates of the contributions of distributed PV systems to the electrical grid, entitled the Regional PV Simulation System (RPSS), whose initial version was developed in the doctoral thesis of the first author (Engerer 2015a).

The RPSS is a real-time capable, scalable, robust computational system that is capable of producing power output estimates from many thousands of distributed small-scale PV systems. This is accomplished through use of monitored PV system power output as the sole operational input, using the K_{PV} methodology of Engerer and Mills (2014). This methodology has been demonstrated to produce excellent results (Tan et. a. 2014; Engerer and Mills 2014; Engerer 2015b). The initial version of the RPSS was shown to produce accurate city-scale simulations across the Canberra region with active reporting from less than 1% of the total PV generators (Engerer 2015a). These simulations were generated for 12,000+ PV systems using suburb-level PV data provided by local distributor ActewAGL, and presented under several different strong collective ramp events.

A further iteration of the RPSS is the subject of this manuscript. Whereas, the previous system computed only aggregate estimates using suburb-level data, version 2 applies these same methods to simulation of 15,000+ micro-generators installed in Canberra, Australia, as organized by transformer (referred to interchangeably herein as a "node"). A beta version of this product will run in real-time from November 2015, with its official launch coinciding with the delivery of this manuscript. The remainder of this paper will detail its data sources and methodology, as well as explore a case study highlighting its usefulness.

2. Data

There are two primary sources of data used within this simulation system. Firstly, the local distribution network service provider, ActewAGL, has provided system information for 15,638 installed micro-generators. This includes the locations of each PV system, the nominal rated capacity and the node on which these systems are installed within ActewAGL's network. Together, these 15,000+ systems comprise 45 MW of distributed generation capacity. Secondly, five-minute interval PV system power output is now available from over 200 sites across Canberra, which are actively reporting on the webpage PVOutput.org in near real-time. These data are actively collected every five minutes by this simulation system through the PVOutput.org Application Programming Interface (API), before being quality controlled and ingested into the simulation system. PVOutput.org users also provide system meta-characteristics, including the module and inverter types, array system layout (e.g. numbers of modules, parallel strings) and the system tilt and azimuth. The collected data are processed through a quality control algorithm, which extracts the exact orientation of each array and removes any erroneous data (Engerer 2015b).

Fig. 1: PV systems in Canberra which are actively reporting their power output every five minutes to PVOutput.org, before being ingested into the Regional PV Simulation System. The black lines are the boundaries of Canberra suburbs

3. Methods

3.1. Modelling PV system characteristics

Given that only the rated capacity, suburb level location and node of each unmonitored PV system is provided, there is a built-in uncertainty in the simulation data. PV systems can have a wide-variety of azimuths and orientations, module and inverter types and system layouts. They are also subject to many different soiling and shading conditions. The lack of availability of the meta-characteristics data (or, rather, its lack of provision) necessitates that the azimuth, tilt and array layouts be selected statistically, requires that the system layouts be assigned randomly and means the de-ratings are simulated.

Before delving into the uncertainty issues at hand, it is worth noting that these issues will not necessarily be present in future modelling studies. Plans for these limitations to be removed in future versions of the RPSS are in place through continued cooperation from ActewAGL, and therefore it is completely possible to know these characteristics with certainty and remove these constraints on the modelling process.

3.1.1 PV system orientations

For the systems' orientation in space, the uncertainty was reduced significantly by sampling the PV systems installed throughout the Canberra region. This was accomplished through a partnership with SolarHub, a local solar installer, who provided extensive detail about 535 PV arrays that they installed in Canberra. Through this information, and that from an additional 74 unique locations gathered from PVOutput.org, it is possible to extract a distribution of tilts and azimuths that are common for Canberra rooftops, and use these as the basis for making assumptions about the remaining 15,000+ sites.

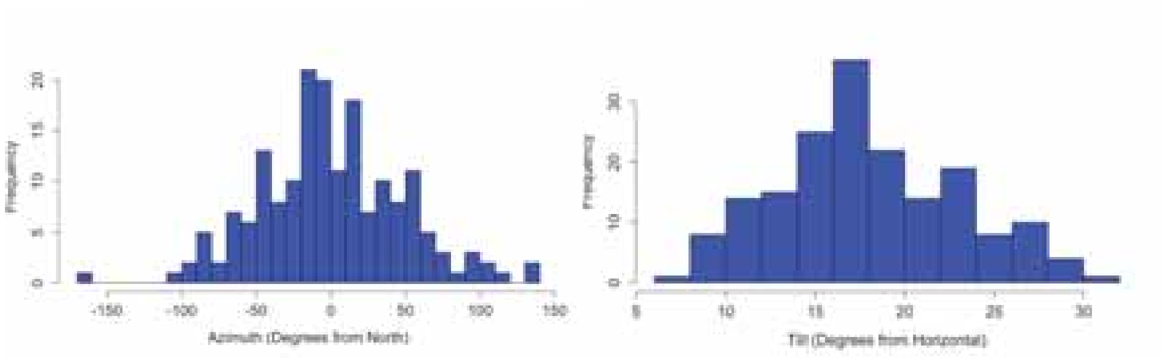

Fig. 2: Histogram of observed azimuths (left) and tilts (right) from over 600 Canberra homes

Figure 2 demonstrates a decidedly Gaussian distribution in the observed tilts and azimuths. This provides a solid statistical basis from which to model the azimuth and tilts of the remaining unmonitored PV systems. These data demonstrate a mean of 0.85° with a standard deviation of 48.5° for the azimuth and a mean of 18.3° and standard deviation of 4.9° for the tilts. From this, a Gaussian distribution of azimuths and tilts is modelled, from which the uncertain PV systems can be assigned an estimated pair of values through a probability weighted, random selection. This process allows for significant removal of uncertainty in system orientation from the modelling process.

3.1.2 PV system layouts

System layouts are more complicated, but are ameliorated indirectly through the limitations of the PV and inverter modelling software, and, again, through statistical sampling. The Sandia Performance and Inverter models are limited to databases of approximately 520 different PV modules and 460 inverter types, which cover a wide array of rated capacities. It is therefore possible, at this time, to simulate only modules that are present in these databases. The modules and inverters chosen from these databases in representation of each system were selected randomly with several constraints. First, only module combinations that were able to satisfy the reported rated capacity were permitted. This means, for example, a 1.4 kW array could not be simulated using 250 W modules, as no combination of 250W modules can create an array of 1.4 kW capacity. Second, from the possible combinations, a random solution was selected from a modelled Gaussian distribution based on the observed installed modules. This is to ensure that it is more likely to select 7 x 200 W modules to build the 1.4 kW array rather than to use 20 x 70 W modules, as the observed dataset has a mean module rating of 245 W and a standard deviation of 27 W. Inverters were chosen randomly from those whose ratings equalled or exceeded the installed capacity, but were limited to a doubling of the installed capacity.

3.1.3 PV system locations

Lastly, the locations of the PV systems were assigned randomly about each transformer node and constrained to lie within the reported suburb. Suburbs in Canberra have an average area of approximately 3 km^2 , but can reach as large as 11 km^2 or as small as 0.4 km^2 in some cases. The suburbs of Canberra are outlined by the polygons in Figure 1 above. There will some sensitivity to this random assignment of site locations, particularly within the timing of ramp events, however, these issues have not been quantified, based on the rationale that this limitation will be removed once the exact locations of the installed systems are known as will be the case in future collaborative data provided by ActewAGL.

3.2. Application of the K_{PV} Methodology

The primary input data to the simulation system is the recorded power output from the sub-set of monitored PV systems within the modeled network. The power output from the monitored sub-set of PV systems can be used to simulate the performance of all the nearby PV systems on the modeled network via the clear-sky index for photovoltaics, K_{PV} (Engerer and Mills 2014):

$$K_{PV} = \frac{PV_{meas}}{PV_{clr}} \qquad \text{(eq. 1)}$$

where PV_{meas} is the reported power output and PV_{clr} is the simulated clear sky power output.

If two sites are close enough to one another (nearby), the KPV values between sites can be assumed to be equivalent. This means that the estimation of a nearby PV system's performance can be computed via:

$$PV_{est_2} = \frac{PV_{meas_1}}{PV_{clr_1}} \bullet PV_{clr_2} \qquad \text{(eq. 2)}$$

where site 1 is the monitored site and site 2 is the nearby un-monitored site.

In order to apply this method at scale, each simulated system requires the computation of its clear-sky performance. These can be populated using the system level information provided by ActewAGL, following the clear sky simulation procedure outlined in Engerer and Mills (2014).

In order to estimate the power output of each simulated site, a K_{PV} value must be assigned based on the performance of nearby monitored systems. Rather than simply use a single nearby site, an average of the nearby K_{PV} values is used, computed using η closest sites.

$$PV_{est_2} = \frac{1}{\eta}\sum_{i=1}^{\eta} K_{PV_i} \bullet PV_{clr_2} \qquad \text{(eq. 3)}$$

Where η is chosen on-the-fly, in a manner that maximizes the simulation accuracy. This is determined by using the sub-set of monitored sites to estimate one another at each time step, to determine which number of neighboring sites (η) minimizes the RMSE error.

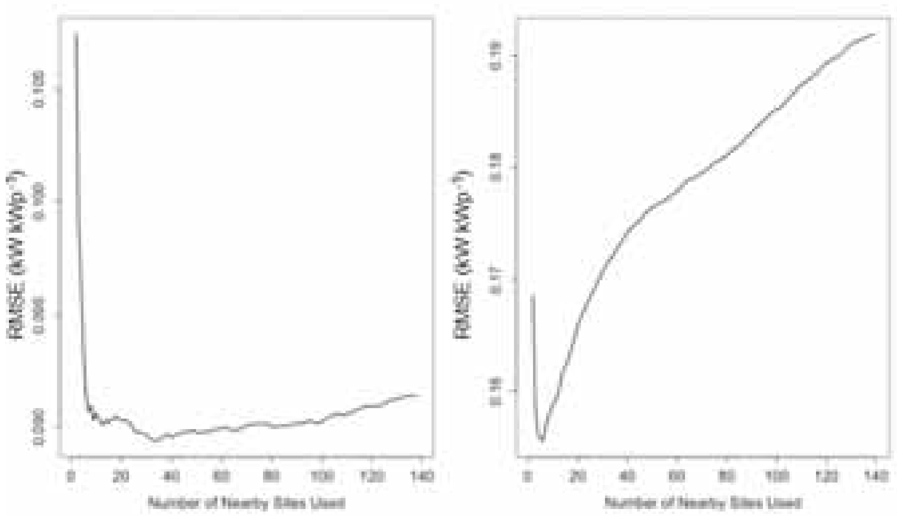

Fig. 3: RMSE values plotted against the number of neighbouring sites used in the K_{PV} based estimate. At left, an example from a clear sky day, at right, that from a partly cloudy day.

3.2.1 Simulated De-rating

An important step in the simulation process is the inclusion of de-rating from soiling, shading and other inefficiencies. These values are very difficult to quantify for non-monitored PV systems. Herein, the simulated behavior of the unmonitored sites will be based on the observed de-ratings of the monitored sites. This is based on the K_{PV90} calculation (Engerer 2015b), which is the running 90th percentile of K_{PV} from the past 30 days. Each unmonitored site is randomly assigned a K_{PV90} profile from one of the monitored sites for a given day. This allows for de-rating to be partially accounted for. At the time of writing, there is not any available research on this topic in the literature, and thus how to simulate PV system de-rating at this scale is not presently well understood. Future research work extending from this manuscript will explore this topic through a sensitivity analysis, simulating many thousands of possible shading scenarios to better quantify these uncertainties

4. Result: Real-time Operations

Through combination of the above data sources and methodologies, city-wide simulations of distributed PV power output across Canberra can now be created in real-time. The result is a live, operational version of the RPSS, which is now hosted and displayed in real-time at http://rpss.info. This online system produces city-wide simulations of PV system power output, mapped to ActewAGL's distribution network by transformer. This result, made possible by an up-scaling of the K_{PV} methodology, represents a significant, unique and promising tool for scientific, engineering and operational purposes. For the first time, it is now possible to quantify the power production of unmonitored distributed PV assets in near real-time, as grouped by distribution asset. The usefulness of this tool can be demonstrated through the following case study, from a particularly high variability day from 5 March 2014.

3.1 Operational Example: 5 March 2014

The meteorological conditions of 5 March 2014 were dominated by passing convective cloud embedded in westerly wind flow, which consequently produced large solar variability throughout the daylight period. The day began with partly cloudy conditions that persisted through 10:00 AEDT. At this time, there was a brief, but significant clearing of cloud cover, with a corresponding 12 MW increase in collective power output over only 10 minutes. From 10:20 AEDT, a very rapid drop in collective power output is observed as a uniform, opaque cloud deck moved in from the southwest, resulting in a 16 MW fall in collective power output over the 60 minute period to 11:20 AEDT. During the period between 10:05 AEDT and 11:00 AEDT, the apparent leading edge of the cloud shadow can be observed within the simulation (Figure 4).

Here, three nodes with particularly high penetrations (>40%) have been selected and highlighted, in order to demonstrate the temporal separation in ramp events across the geographic region (Figures 4 and 5). Node 2498 experiences the negative ramp event first, followed by node 2440, which is part-way through the negative ramp event, as can be seen in the right hand image of Figure 4. Node 2078, however, still experiences relatively clear conditions at this time, and has not yet ramped down its power output. This separation in the timing of these negative ramp events can be observed more clearly in Figure 5, which displays a 20 minute difference between the negative ramp at node 2498 and 2078. Given the separation distance of 20 km this implies a cloud speed of approximately 16.5 m s^{-1} (60 km $h^{-1)}$. Additionally, a strong positive ramp follows at approximately 12:05 AEDT through 1:15 AEDT. Again, a sequential pattern is observed, with node 2498 reaching a peak first, followed by node 2440 and then 2078.

This transformer level breakdown is also useful for comparing the ramp rates experienced by these individual nodes to that experienced collectively across the network. Using the most extreme 10 minute negative ramp events for the 10:00 - 11:00 time period, the collective ramp rate across all sites was 0.03 kW kWp^{-1} min^{-1}, while the three nodes experienced ramp rates of 0.04 kW kWp^{-1} min^{-1}, 0.072 kW kWp^{-1} min^{-1} and 0.057 kW kWp^{-1} min^{-1}. So, perhaps unsurprisingly, the ramps experienced by individual nodes can be shown to be more extreme than that experience by the collective system. We expect that by tracking these types of events over time, the most extreme events can be characterised, and the distribution systems engineered to withstand them. Plausibly, this simulation system could also be used to test the theoretical limits of distribution networks, by adding additional simulated PV systems to transformer nodes under strong ramp events like this one until the variability exceeds the capability of the given transformer to respond.

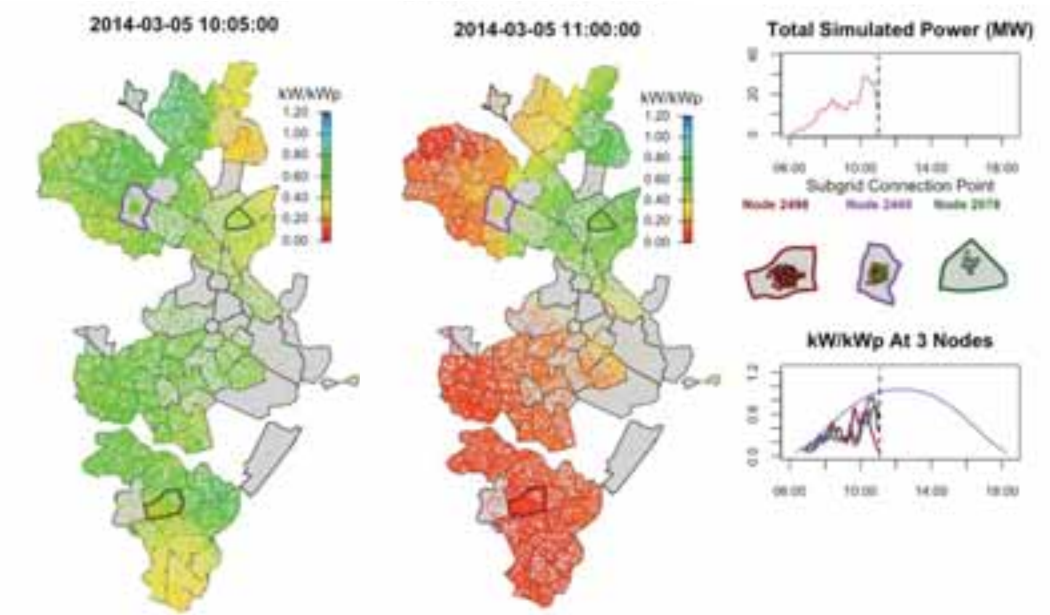

Fig. 4: These two images display the simulated power output from 15,000+ PV systems at two different times (10:05 and 11:00 AEDT) during a high variability day on 5 March 2014. Each point on the map represents a simulated PV system, with colour varying from red through to blue as power output (kW kW$_p^{-1}$) increases. Grey polygons display the boundaries of the major Canberra suburbs. At right, the power output is displayed as a time series, in both total power (MW) output and by three selected transformers (kW kW$_p^{-1}$). In the top time series plot, the blue line depicts the clear-sky curve for the collective output from all simulated PV systems. This simulation can be viewed online at bit.ly/ARENA_DNSPs.

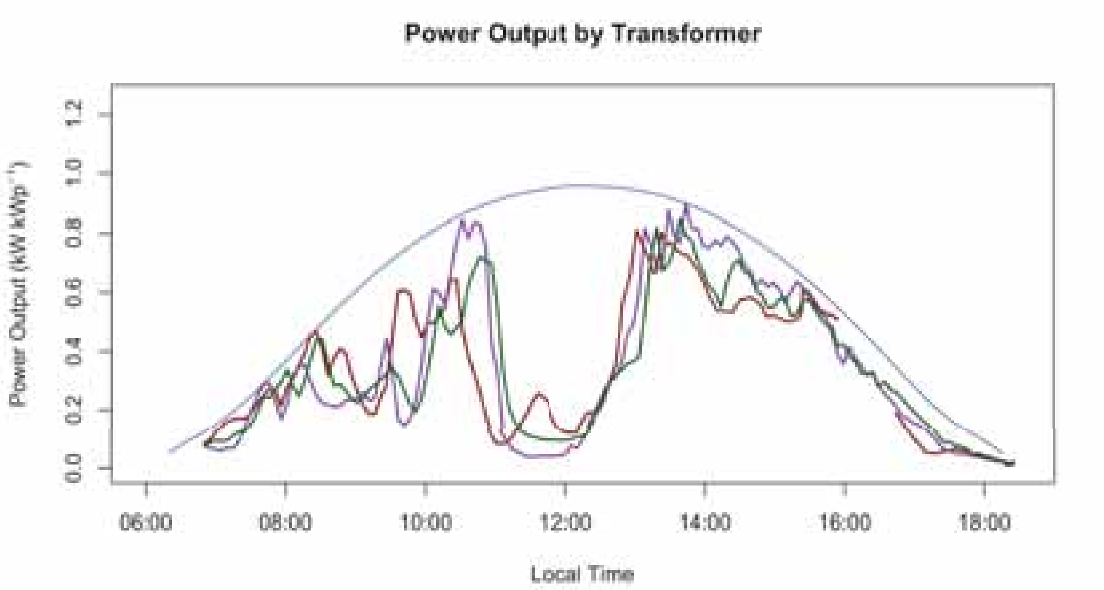

Fig. 5: The simulated power output (kW kW$_p^{-1}$) is presented for each of the three selected transformers (red 2498, purple 2440 and green 2078) during the high variability day of 5 March 2014.

5. Future Work

With the presented version of the RPSS in place, future efforts will work to increase the amount of monitored PV data ingested into the system, blend this surface data with satellite inputs and incorporate short-term forecasting methods to bring the system to true real-time status. Efforts are currently underway, which will double the current monitoring network, through deployment of tens of data-logging devices through funding from the Australian Renewable Energy Agency (ARENA), as well as to include data from inverter companies SMA and Fronius. Through these cooperative efforts, we expect to increase the total number of monitored sites in Canberra to 500-700 over the next two years. This will allow for more extensive validation of the present modeling methods.

Additionally, Australia has recently begun to receive 1 km^2 satellite data every 10 minutes from the new Himawari 8 satellite, which is a substantial improvement over the 4 km^2 hourly data from its predecessor. This will create the opportunity to develop model blending techniques which will be used to combine the K_{PV} and satellite derived irradiances, as well as cross-validate these two methods. Recent advances in solar PV analytics technology have even shown that PV systems can be used to extract the diffuse and beam components of radiation (Engerer and Xu 2015), raising the prospect of satellite derived irradiance validation or spatio-temporal analyses of solar radiation using large PV monitoring networks like the one being used by the RPSS.

Furthermore, this system, despite the inclusion of additional sites and the satellite data, will still only operate in near real-time, without the inclusion of short-term solar forecasting. This is due to the inherent time-lagged nature of the reporting data and the time taken to complete the modeling process and deliver the result. For this tool to truly provide proactive decision making capability, it must be incorporate short-term solar forecasting technologies. In the near future, we intend to apply the latest in short-term, distributed solar forecasting technology to this problem, thereby advancing the RPSS to true real-time operational status.

6. Conclusions

This version of the Regional PV Simulation System, which is now capable of simulating 15,000+ PV systems at transformer level in Canberra, Australia, is a unique and promising tool for distribution network service providers and energy markets in Australia. For the first time, the contributions of small-scale, monitored solar PV assets can be quantified for distribution networks, as organised by distribution asset. This information will be a key enabler for future scenarios that increase the allowable penetrations of solar PV via combinations of accurate distributed PV forecasting and energy storage technologies. With the information provided by this simulation system, distribution networks will no longer be naïve with respect to the real-time behaviour of their distributed PV assets, removing one of the most significant sources of uncertainty within Australian electrical grids. This tool, made possible through an up-scaling of the K_{PV} methodology, is now running in real-time, and can be located at http://rpss.info. Soon, this operational system will begin delivering near real-time information to local utility ActewAGL, with plans to incorporate short-term solar forecasting and high-resolution satellite imagery into its operations. Future work will aim to increase the accuracy and robustness of the modelling methods, with emphasis on increasing the number of ingested monitoring sites and blending surface based K_{PV} estimates with those based on satellite-derived irradiance.

7. References

[ABS] Australian Bureau of Statistics, 2014. One in five households now use solar energy. 2014. Web. Accessed 20 July 2015.

Chen, L. and Franklin, E., 2011. Mapping the onset of solar grid-parity in Australia. Proceedings of the 49th Annual Conference of the Australian Solar Energy Society. Sydney, November 2011.

Engerer, N.A., 2015a. City-wide simulations of distributed photovoltaic array power output. Ph.D. thesis, The Australian National University. Submitted August 2015.

Engerer, N.A, 2015b. QCPV: A quality control algorithm for distributed photovoltaic array power output. Submitted to IEEE Transactions on Sustainable Energy. August 2015.

Engerer, N.A. and Mills, F.P., 2014. K_{PV}: a clear-sky index for photovoltaics. Solar Energy, 105, pp. 679–

693.

Engerer, N.A. and Xu, Y., 2015. A simple model for estimating the diffuse fraction of solar irradiance from photovoltaic array power output. 21st International Congress on Modelling and Simulation (MODSIM2015). Gold Coast, Australia, December 2015.

Ergon Energy. PV Industry Alert. Issue 20. 23 June 2015

Horizon Power, 2012. Technical requirements for renewable energy systems. Technical Report HPC-9FJ-12-0001-2012.

Noone, B., 2013. "PV integration on Australian distribution networks." The Australian PV Institute, UNSW, Australia.

[APVI] Pv-map.apvi.org.au. Australian Photovoltaic Institute Market Analyses, 2015. Web. Accessed 21 July 2015.

SunWiz, 2012. Solar PV forecast for AEMO 2012-2022. Technical Report.

Tan, J., Engerer, N .A., and Mills, F.P., 2014. Estimating hourly energy generation of distributed photovoltaic arrays: a comparison of two methods. Proceedings of the 52nd Annual Conference of the Australian Solar Energy Society. Melbourne, May 2014.