Operational Characteristics of a Cluster of Distributed Photovoltaic Systems

S. F. Heslop and I.F. MacGill

Abstract**-- As PV systems are an intermittent resource; their behaviour will need to be appropriately characterised before they can be reliably integrated into the electrical grid on a large scale. Relevant issues at the network level include voltage regulation and more variable power flows through network equipment. At the system level, high penetrations may have significant supply-demand and reserve implications. A particular challenge is that PV can be implemented in a highly distributed manner with potentially large numbers of small residential systems. This paper attempts to characterise the behaviour of a distributed system of small scale PV installations using high frequency (10 second) monitoring data from a number of existing PV installations in the Hunter region of Australia. Results highlight the significant variability of PV generation under some weather conditions, the benefits of distributed PV implementation in reducing aggregate short-term variability and the particular characteristics of this variability including its dependence on cloud cover and the time of day. Results are presented in a form which is hoped to allow easy incorporation into distributed generation models and which assists designers of smart grid systems.**

Index Terms**-- distributed power generation, solar power generation, photovoltaic systems, power system reliability**

I. Introduction

INCREASING levels of installed PV are being seen in a growing number of countries and regions including Europe, North America and Japan. The same increase on what are already significant penetrations (570 MW at 2010 with 383 MW installed last year [6]) are also expected of Australia, see "Fig. 1".According to [7] penetration levels are now estimated at around 1GW making it the third highest capacity renewable electricity source in Australia (after hydro and wind) and placing Australia as one of the top five solar nations in terms of installed capacity. With increased levels of installed PV will come increased levels of power system penetration. Penetration is defined in [5] as the ratio of the rated capacity of the installed PV and the peak load of the system. Note, however, that this is system level penetration, and other measures of penetration can be more relevant in the context of potential PV impacts within the electricity network. In particular, large individual PV systems or clusters of smaller systems can be significant at the feeder and substation level. Discussed in [3] is what is likely to limit the level of possible PV penetrations; these being voltage and frequency issues due to the inherent variability of PV generation. Voltage (and power quality) issues are likely to be seen first at the distributed level with frequency problems arising later, and only at higher penetrations, at the system level. At a system level, to offset this variable generation, in [2] it is claimed flexible dispatchable generation is required. Other approaches will be required at the distribution level. An example of this is given in [8] where battery storage and PV inverters running at a reduced power factor are being utilised to mitigate the impacts of fluctuations in PV output power. According to [1], once PV behaviour is understood its impact on regulation and load following operations can be quantified and a dispatch strategy developed. As Australia installed more small scale distributed PV systems than Germany in 2010 [10], impacts at the distributed level should certainly be the early focus for the main electricity grid. An important question here is the aggregated behaviour of clusters of PV systems within particular regions of the network. It can be expected that these systems will 'see' somewhat different solar insolation, and hence generation performance at particular times due to the highly variable nature of cloud formation and movement.

A key issue in analysis of PV behaviour at the system and distribution level is the timeframe of interest. System level analysis performed in [5] uses sampling rates of 1 minute, 5 minute and 1 hour, for [9] hourly samples are used. However, higher level voltage and power quality management as well as frequency regulation and contingency responses are generally managed with SCADA systems operating at time intervals of seconds. As high penetration distributed PV is expected to have impacts at this control level we will need analysis of PV behaviour at a similar resolution. In [4] the highest sampling rate is 1 minute and the need for higher resolution data for grid stability analysis is mentioned. Note also that this data set is a calculation of PV output based on solar radiation measurement. This paper presents the results of an analysis on the variability of 18 PV systems located in the Hunter Valley, the PV systems are all small scale (< 10 kW) and dispersed over an area approximately 24 km x 30 km. A key aspect of the analysis performed in this paper is not only the high sampling rate of 10 sec but also that it utilises actual PV output data and so more accurately represents PV behaviour from a grid perspective. The analysis is based on approximately 3 months' worth of high frequency (10 sec) PV output data from August to November in 2010.

The paper aims to present its results on the operational variability of individual and clustered PV systems using some standard data presentation formats, but also a number of new presentation approaches that better highlight the operational

This work was supported in part by Australian Solar Institute (ASI) research funding on solar forecasting and managing high PV penetrations

implications of clustered PV systems on the distribution network. The intent is also to permit easy incorporation into models for distributed generation and smart grids. The results are split into categories considered requisite for full understanding of the system. Results present the variability of the system overall, under aggregation, hour to hour and according to degree of cloud cover

Section I describes the method of analysis, the results of the analysis are presented in Section II and Section III discusses the results. Discussion of the work is presented in Section IV and conclusions established in Section V.

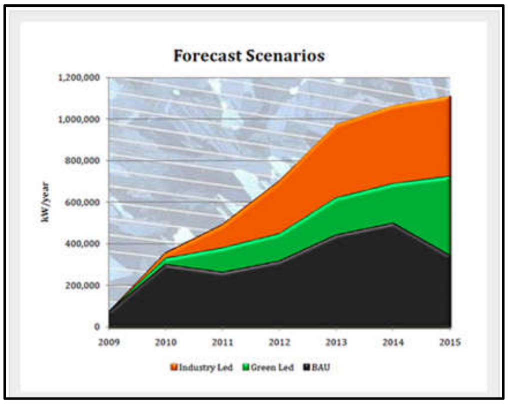

Fig. 1. Some forecast scenarios for future PV deployment in Australia under a number of possible market scenarios (Morris and Johnston, 2011). Courtesy: Australian PV Market Forecast 2010-2015, Prepared by Nigel Morris (Solar Business Service) and Warwick Johnston (SunWiz Consulting), Dec 2010

II. METHODOLOGY

Data for all the sites came as a comma delimited file with 2 columns, timestamp and PV output. The output current of the inverter is measured through a current transformer (CT) and sampled through an A/D converter every second; the A/D converter is connected to a Linux box via a serial link. 10 second averages are logged to the Linux box and then stored on a server located at the Commonwealth Science and Industrial Research Organisation (CSIRO) Energy Centre in Newcastle via the internet. As "Fig. 2" reveals, no site logged data continuously over the 3 months with the vertical steps indicate a gap in the sampling. The dashed line is how a plot would look if there was continuous sampling from the earliest sample recorded to the latest

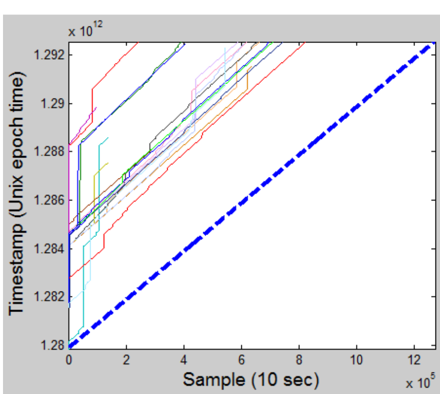

Fig. 2. Plot showing the continuity of sampling for each site. A vertical jump indicates a time gap in the sampling

This data had to be manipulated to group days and align timestamps before it was ready for analysis – that the data manipulation process was correct was always verified. "Fig. 3" confirms the correctness of a timestamp alignment process for a group of sites with the dips in PV output correlated. The particular value of the data provided by CSIRO is that the high resolution sampling rate enabled the analysis to pick up on the impact of cloud transients. Data manipulation and analysis is executed using Matlab.

A. *Description of Data Manipulation Techniques*

Described below is a description of the data manipulation techniques applied to the dataset

1) *Split into Individual Days*

The dataset for each site was provided as one continuous set across the 3 months and needed to be split into individual days. Start of day was identified as solar output showing a 10 sample moving average > 0.35 and end of day as 1000 samples of zero output. 99 individual days was the result.

2) *Change in Output (Variability)*

Change in output was determined by shifting a normalised PV output dataset and then taking the absolute difference between the shifted and non-shifted dataset. As each sample is 10 sec, 1 shift equates to a 10 sec time shift comparison, 6 shifts 1 min, 30 shifts 5 min etc. The shift distance is appended onto the front and end of the data set so the first and last section is being compared to zero.

3) *Grouped into Matching Days*

Timestamps for each individual day were compared and matched so days could be grouped. All sites had days with no data with some days being discarded from the analysis due to lack of samples. "Table 1" shows the number of days for each site

Site	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Days	12	81	77	86	18	73	74	8	90
Site	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
Days	76	84	36	83	88	34	45	18	8

Table 1. Number of days where PV output recorded for each site

4) *Timestamp Aligned*

For each group, the site with the earliest time for which solar output occurs is discovered; all other sites then have zeros padded to their output data down to this time. The same is repeated for the latest showing of solar output. See "Fig. 3" for an example day after timestamps have been aligned for the group.

5) *Hourly Split*

Using a dataset which has been *split into individual days* and *timestamp aligned*, this dataset is split further into hourly sections, from 5am to 8pm, giving 15 datasets. To explain further, assuming 10 sites, each with 10 days' worth of data; after an *hourly split* the 5-6am dataset would consist of 100 (10 sites x 10 days) arrays, each containing an hours' worth of data

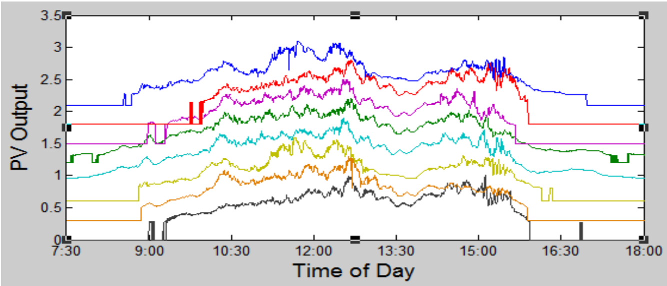

Fig. 3. PV output for 8 sites with timestamps aligned

6) Sunny, Partly Cloudy and Majority Cloud Split – Method 1

Each day was also categorised as sunny, partly cloudy or majority cloud using the fraction of cloud cover classification used by the Bureau of Meteorology (BOM). The BOM classifies cloud cover as the fraction of the sky covered by cloud in units of eighths (8^{th}). The measure ranges from 0 for clear sky up to 8 for full cloud cover. The closest measurement point for cloud cover is Newcastle University where a measurement is taken at 9am and 3pm every day. How each day was categorised as either sunny, majority cloud or partly cloudy using the BOM eighths method is as follows

- Sunny. Mean of 9am and 3pm measure < 1
- Partly cloudy. Mean of 9am and 3pm measure > 1 and < 7
- Majority cloud. Mean of 9am and 3pm measure > 6

7) Sunny, Partly Cloudy and Majority Cloud Split – Method 2

Each day for every site was categorised as sunny, partly cloudy or majority cloud. "Table2" below defines how each day was split. This split was done as it was expected that the operational characteristics of the PV system would differ substantially according to the degree of cloud cover. The conditions were determined through a process of data fitting. The following was assumed:

- Sunny day would have above average output and below average variability
- Day with Majority Cloud would have below average output and below average variability

Sunny	
Mean output for day > Mean output for site	Variance for day < (Mean variance for site - 1 STD of variance)
Majority Cloud	
Mean output for day < (Mean output for site - 1 STD of site output)	Variance for day < Mean variance for site
Partly Cloudy	
Outside both sunny and majority cloud conditions	

Table 2. Sunny, Majority Cloud, and Partly Cloudy split - method 1

B. Categorisation of Variability Analysis

Variability analysis of the datasets can be categorised into 4 sections: Overall Variability of the System, Aggregated Variability, Variability Hour to Hour, and Variability According to Cloud Cover

1) Overall Variability of the System

This analysis takes all data irrespective of day or site after it has been *split into individual days*. The *change in output* for this data is then calculated.

2) Aggregated Variability

Using a dataset which has been *grouped into matching days* and *timestamp aligned*, this dataset has its PV output for all matched days summed. The c*hange in output* for this summed data is then calculated, giving 99 days

3) Variability Hour to Hour

After the *hourly split*, the c*hange in output* for each of these 15 datasets is performed.

4) Variability According to Cloud Cover

After a *sunny, partly cloudy and majority cloud split* using method 1 and method 2 a change in output is performed on the 3 datasets (sunny, partly cloudy and majority cloud). A *change in output* after an *hourly split* of the 3 datasets is also performed for both methods.

C. Presentation of Results

The manner in which data describing the behaviour of PV is presented needs to be in a form useful to those who might utilise it for modelling of variable distributed generation and its impact on voltage and power quality. From [10], "Fig. 4" shows the frequency of ramp rates for a single PV system as a percentage of the total capacity of the PV array and could be considered a type of probability distribution function (PDF). The majority of the analysis results are presented in a manner similar to this. The 15 min sample is the mean of the 1 min sample.

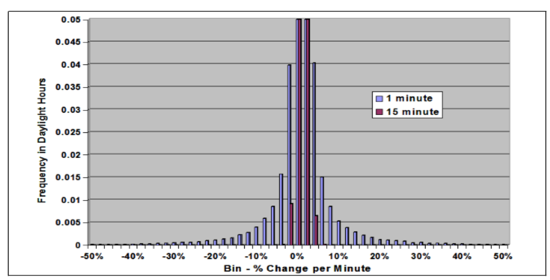

Fig. 4. Type of probability distribution function used as template. Frequency of different ramp rates for a PV system as a percent of the total capacity of the PV array [4]

III. Results

A. Overall Variability of the PV System

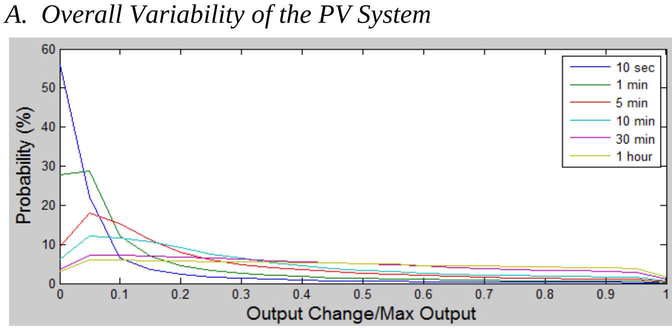

Fig. 5. System PDF for average change in output for all sites considered separately. Change in output calculated for time differences of 10 sec, 1 min, 5 min, 10 min, 30 min and 1 hour

"Fig. 5" shows a plot depicting the average variability for the system (average for all sites) for 6 time frames. As the PDF was performed on normalised data, the x-axis can essentially be viewed as a percentage change (%), this is the case for all plots of this type in this paper. Each of the bins has a width of 0.05 (5%) apart for the "0" bin which covers changes in output from 0 – 2.5%, "Table 3" gives the ranges for each bin. The variability is smallest at the 10 sec time frame with the majority of change in output being less than 0.1 or 10%. With an increase in time frame the plot is "flattened" with an increase in the probability of larger changes in output.

Bin	Range of Change in Output
0	0-2.5%
0.05	2.5-7.5%
0.1	7.5-12.5%
0.15	12.5-17.5%
0.2	17.5-22.5%
0.25	22.5-27.5%
0.3	27.5-32.5%

Table 3. Range of change in output for each bin

B. Aggregated Variability

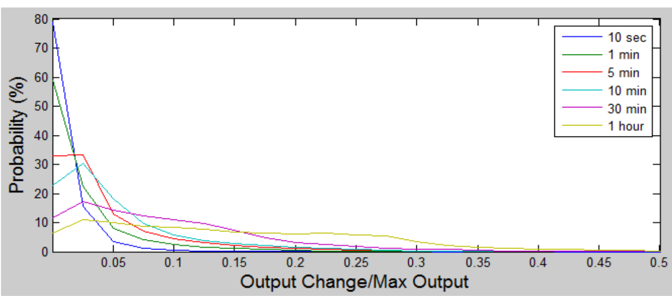

Fig. 6. System PDF for change in output after aggregation. Change in output calculated for time differences of 10 sec, 1 min, 5 min, 10 min, 30 min and 1 hour

"Fig. 6" shows a plot depicting the variability for the system after aggregation for 6 time frames. As expected the variability is noticeably less than that of "Fig. 5", the "squashing" (as opposed to "flattening") of the curves towards the left is an indication of a greater probability of smaller changes in output. Aggregation having a smoothing effect on the overall output variability of multiple PV systems is well appreciated and discussed in a number of existing papers, [11] and [12] for example. However, there has been very limited work at high sampling rates as shown here. Considering PV variability in aggregate is important; from the grids perspective, at feeder level and up, PV output will be experienced in aggregate.

C. Variability Hour to Hour

"Fig. 7" and "Fig. 8" depict the variability of the system when broken down by hour. Again, the variability is less for when the output change is calculated across a smaller time difference, this is shown by there being a greater probability of a change in output of less than 2.5% for the 1 min chart; averaging around 70% and around 50% for the 10 min chart.

Focusing on the 1 min chart, the plot also shows variability increasing towards the middle of the day with a concave shape present for output changes < 2.5% and a convex shape for output changes > 2.5%.

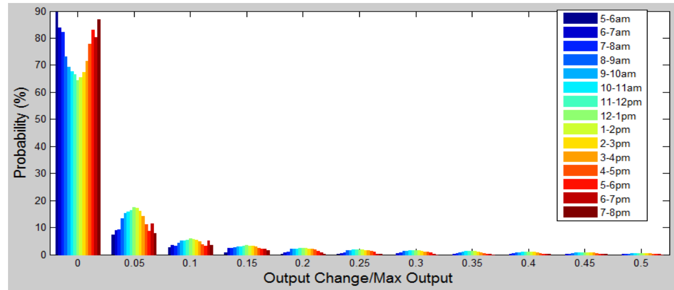

Fig. 7. System PDF for change in output broken down by hour. Change in output calculated for a time difference of 1 min

Looking closer at the 10 min chart we see small variation in output at the extremes of the day; this could be explained by there being very little output during these times. For the hours of 8-9am and 3-4pm we see variability is greater than for the middle of the day (indicated by the 2.5 – 7.5% bin), this could be explained by the ramp up in output as the sun rises; more pronounced on the 10 min chart than for the 1 min chart.

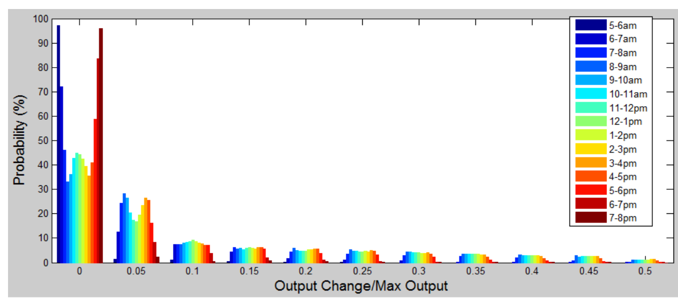

Fig. 8. System PDF for change in output broken down by hour. Change in output calculated for a time difference of 10 min

It is worthwhile comparing "Fig. 8" and "Fig. 5". "Fig. 5" shows a probability of approximately 12% for a 5% change in output over a 10 min interval. Looking at the 5% bin for "Fig. 8", for a large part of the day, the probability of seeing a 5% change in output is actually quite higher, almost 30% between 8am and 9am.

D. Variability According to Cloud Cover

The next figure, "Fig. 9", is an observation on the different behaviour to be expected for different levels of cloud cover. The top chart uses method 1 to split the days and method 2 is used for the bottom chart. Both charts show sunny days to be far less variable with a greater percentage chance of smaller changes in output. Interestingly, majority cloud and partly cloudy days exhibit similar variability.

One might expect partly cloudy days, with solar production showing regular drops and subsequent recoveries due to passing clouds in an otherwise sunny sky, to be more variable in comparison to a generally overcast day with no large changes in insolation. These charts only show a percentage change relative to the maximum daily output and not the magnitude. It may be that the changes in magnitude for the partly cloudy days are noticeably larger.

Method 1 splits the days according to the mean of two cloud measurements by the BOM (9am and 3pm) made at one location (Newcastle University). The average distance the PV sites are away from Newcastle University is 15 km, with seven sites over 20 km away. This distance may mean the cloud cover measured at the university is different to the cloud cover at the PV site. Also, two cloud cover measurements at 9am and 3pm may not be representative of the average cloud cover for the day. Considering this, the results from method 2, which was developed through a process of data fitting, has also been included. Looking at "Fig 9" below, the main difference between the two methods seems to be the categorisation of a sunny day.

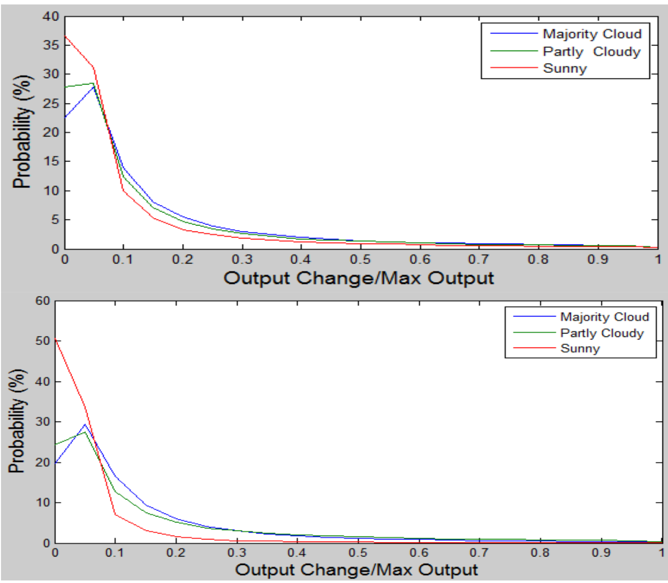

Fig. 9. System PDF for change in output for all sites after being split into majority cloud, partly cloudy and sunny days. Method 1 is the top graph and method 2 the bottom. Change in output calculated for interval of 1 min

Comparing "Fig. 9" and the 1 minute plot of "Fig. 6" it would still be said that despite the reduced variability on sunny days, it is still greater than that of the aggregated output. "Fig. 10" to "Fig. 15" shows 5 min variability by the hour for Majority Cloud, Partly Cloudy and Majority Cloudy days for both methods 1 (BOM) and 2 (data fitting). Again comparing the two methods, when broken up by the hour, the results for both methods are very similar. Days with majority cloud show the most variability around midday for bins 5%, 10% and 15% compared to partly cloudy days.

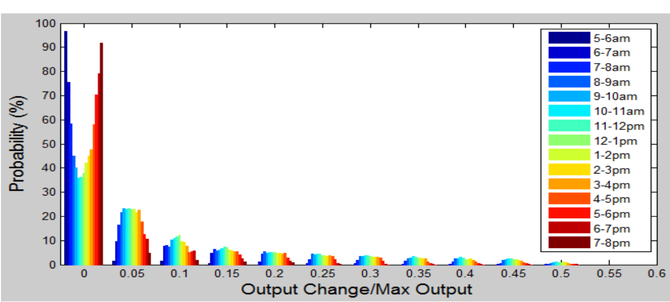

Fig. 10. Method 1 - majority cloud hourly split. Change in output calculated for a time difference of 5 min

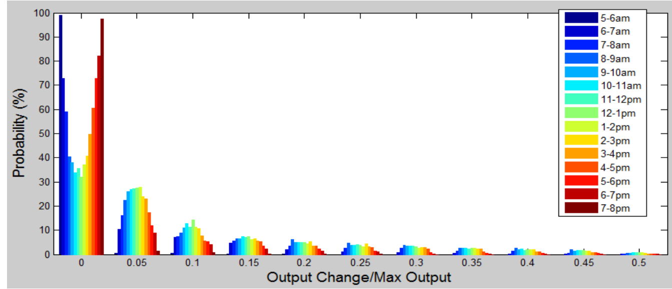

Fig. 11. Method 2 - majority cloud hourly split. Change in output calculated for a time difference of 5 min

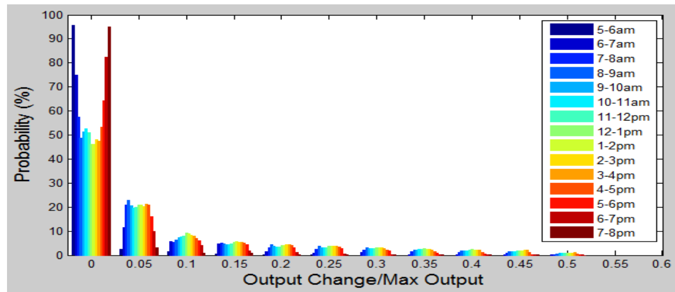

Fig. 12. Method 1 - partly cloudy hourly split. Change in output calculated for a time difference of 5 min

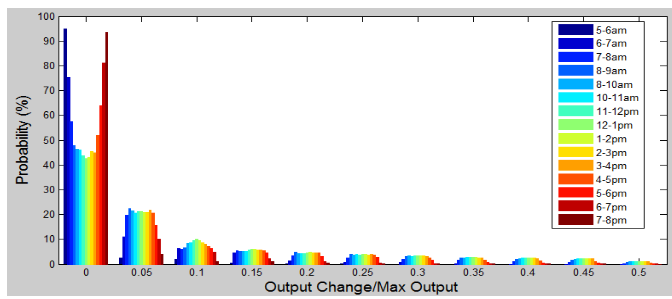

Fig. 13. Method 2 - partly cloudy hourly split. Change in output calculated for a time difference of 5 min

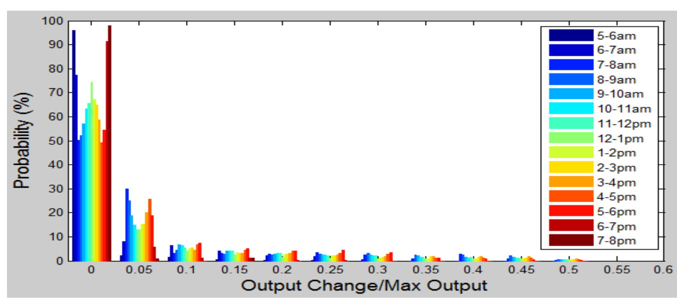

Fig. 14. Method 1 - majority sun hourly split. Change in output calculated for a time difference of 5 min

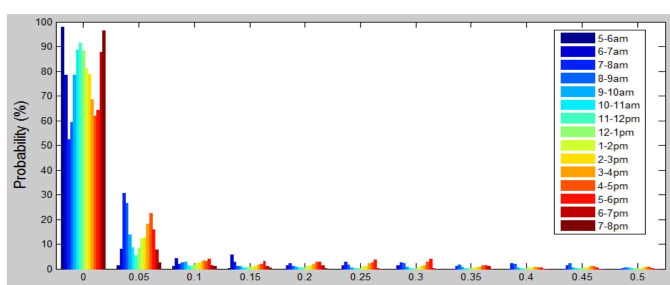

Fig. 15. Method 2 - majority sun hourly split. Change in output calculated for a time difference of 5 min

Days with Majority Sun show the greatest percentage of output changes in the 5% bin for the hours of 7-9am and 3-6pm, this would be due to the rapid ramp in output as the sun rises to and falls from its peak without the influence of cloud cover. Around midday, days with Majority Sun show very little variability, with the majority of output changes being less than 2.5%. Comparing “Fig. 14” with “Fig. 5”, if cloud type weren’t taken into consideration, one might expect a 15% probability for a 10% change in output over a 5 min interval. Looking at “Fig. 14”, if the day were sunny this prediction would be out by a margin of around 10% (greater if the chart of method 2 is referred to) for the entire day.

IV. Discussion

Understanding the operational characteristics of distributed solar is necessary to assist in the development of effective network management strategies to mitigate expected impacts that come with increased levels of PV penetration. As this study is small in scale its findings on variability would only be applicable at the distribution level for voltage and power flow management. When designing smart grids utilising PV, information of this type or similar could assist in the development of optimisation algorithms and for the sizing and operational management of network control elements, controllable loads, storage and conventional backup generation.

How this information is presented is also important. The use of charts displaying the probability of a percentage change in output provides insights into the possible variability of PV at different time intervals, when aggregated across multiple systems within a region, under different weather conditions and at different times of the day. It is intended that these findings can be incorporated into smart grid models and assist in the design of voltage and power flow control for feeders featuring high levels of PV.

For analysing the possible impacts on voltage of large penetration PV at the distribution level, high resolution data is required to pick up the large change in output due to cloud transients. Before large penetrations of PV are permissible on a feeder, it may be that analysis on local high resolution irradiance measurements occurs first to ensure the PV can be handled. It would be interesting to compare the behaviour of PV in different locations, weather patterns certainly vary according to geographical location and so should the behaviour of PV.

Thoughts on future studies

1. It is expected that despite the results showing partly cloudy and days of majority cloud showing similar operational characteristics when presented as percentage changes, the magnitude of these changes would be far greater for partly cloudy days. An analysis incorporating magnitude would be interesting
2. The results presented should ideally be useful to engineers within Distribution Network Service Providers. A survey of utility engineers to get their thoughts on useful presentation formats for solar behaviour would be informative
3. The development of a standard for determining what constitutes a day with majority cloud, part cloud or majority sun

V. Conclusions

From the analysis we can conclude that it was necessary to break the analysis up into categories to get a more accurate picture of the behaviour of the PV system. The following was revealed

1. PV generation is highly variable across all time frames.
2. Aggregated PV generation displays significantly less overall variability than the individual systems
3. Drilling down to an hour by hour analysis highlights how variation differs throughout the day – this has potentially significant network implications.
4. Breaking down the analysis into cloud type -Majority Cloud, Partly Cloudy and Majority Sun - highlights how expected variability will depend significantly on the particular weather patterns experienced on a day to day basis. Forecasts of future weather might, therefore, also be able to guide electricity industry participants in the potential future behaviour and variability of distributed PV. Results show how PV operational characteristics vary day to day according to cloud type

This change in the degree of variability according to cloud type and hour to hour has potentially significant implications for network operators where increased levels of PV are expected. The design of voltage and power flow management will likely not be able to take a broad view of expected variability of PV but, instead, a more context specific approach depending on the particular local network characteristics, load behaviour, weather patterns and variability and location of distributed PV.

These implications extend to designers of smart grids featuring PV as well. If a broad approach to variability is taken then ‘smart grid’ models won’t be accurate resulting in the possibility of insufficient storage and/or backup generation, a design failure for a smart grid designed to be self-sufficient

VI. Acknowledgement

The authors gratefully acknowledge T. Moore, S. Lindsay, G. Platt and D. Rowe of the Commonwealth Science and Industrial Research Organisation (CSIRO) Energy Centre, Newcastle for their assistance and for providing the PV output data

VII. References

[1] S. Achilles, S. Schramm, and J. Bebic, "Transmission System Performance Analysis for High-Penetration Photovoltaics," GE Global Research, Subcontract Report NREL/SR-581-42300, Feb. 2008.

[2] J. Bebic, “Power System Planning: Emerging Practices Suitable for Evaluating the Impact of High-Penetration Photovoltaics," GE Global

Research, Niskayuna, New York, Subcontract Report NREL/SR-581-42297, Feb. 2008.
[3] C. Whitaker, J. Newmiller, M. Ropp, and B. Norris, "Distributed Photovoltaic System Design and Technology Requirements," Sandia National Laboratories, Report SAND2008-0946 P, Feb. 2008.
[4] D. Renne, R. George, S. Wilcox, T. Stoffel, D. Myers, and D. Heimiller, "Solar Resource Assessment," National Renewable Energy Laboratory, Technical Report NREL/TP-581-42301
[5] X. Bai, K. Clark, G. A. Jordan, N. W. Miller, and R. J. Piwko, "Intermittency Analysis Project: Appendix B Impact of Intermittent Generation on Operation of California Power Grid," GE Energy Consulting, Report CEC-500-2007-081-APB, Jul. 2007.
[6] M. Watt, R. Passey, W. Johnston, "PV in Australia in 2010," Australian PV Association, May. 2011.
[7] W. Johnston, "Australian Solar's Gigawatt Valley of Death: Worth Dying For?" www.renewableenergyworld.com, Jul. 2011.
[8] M. Braun, "Integrating PV in Local Distribution Systems – Germany," IEA PVPS Task 14 Meeting, Dec. 2010.
[9] [9] D. Lew, N. Miller, K. Clark, G. Jordan, Z. Gao, "Impact of High Solar Penetration in the Western Interconnection," National Renewable Energy Laboratory, GE Energy, Technical Report NREL/TP-5500-49667, Dec. 2010.
[10] W. Johnston, "Australian Solar Industry Installs More Small PV than Germany in 2010," www.renewableenergyworld.com, Mar. 2011
[11] T. Oozeki, K. Otani, T. Takashima, Y. Hishikawa, G. Koshimizu, Y. Uchida, K. Ogimoto, "An Evaluation Method for Smoothing Effect on Photovoltaic Systems Dispersed in a Large Area," Research Centre for Photovoltaic, National Institute of Advanced Industrial Science and Technology, 2009
[12] T. E. Hoff, R. Perez, "Quantifying PV Output Power Variability," Clean Power Research, Jul. 2010

VIII. Biographies

S. F. Heslop received his BE in Electrical Engineering from the University of Newcastle in 2000 and ME (Energy Systems) from the University of New South Wales in 2010. He is currently completing a research masters at the University of NSW

I. F. MacGill is an Associate Professor in the School of Electrical Engineering and Telecommunications at the University of New South Wales, and Joint Director (Engineering) for the University's Centre for Energy and Environmental Markets (CEEM). Iain's teaching and research interests at UNSW include electricity industry restructuring and the Australian National Electricity Market, sustainable energy generation technologies, distributed energy resources, energy efficiency options, energy and climate policy and environmental regulation.