Customer-Led Network Revolution

# Technical Note: Solar Photovoltaic (PV) Installations

**DOCUMENT NUMBER**
CLNR-L095

**AUTHORS**
Robin Wardle, Rupert Way, Phoebe Su, Gavin Whitaker, Christian Barteczko-Hibbert, Durham University

**ISSUE DATE**
17/02/2015

# Contents

# 1 Document Purpose

The Customer-Led Network Revolution (CLNR) Solar PV trials investigated domestic customers' electricity usage and PV array output for three groups of customers; a control group with Solar PV panels but without any intervention (test cell 5) and two groups with interventions designed to test the demand flexibility of customers and their ability to use their own generation: either a manual in-home display to help the customer actively manage their generated electricity (test cell 20IHD), or automatic within-premises balancing of excess generation to a water heating system (test cell 20auto).

This technical note addresses two areas of interest which arose from the Solar PV trials (i) investigation of the installer-declared peak capacity of the PV installation versus the measured peak capacity and (ii) examination of the diversity of the Solar PV panels using the After Diversity Maximum Demand (ADMD) methodology.

# 2 A Comparison of Measured and Rated Capacity

## 2.1 Introduction

In studying solar PV generation, a more useful measure than the absolute power or energy export amounts can be the specific power or energy,

$$p = \frac{P}{\hat{P}}, \qquad e = \frac{E}{\hat{P}}$$

where

$$P = \text{power}$$
$$E = \text{energy generated in a given time period}$$
$$\hat{P} = \text{ the maximum power producible by the solar array}$$

Specific values of power and energy allow arrays of different sizes to be compared and grouped. Absolute values are useful when studying whole populations, or groups of arrays which represent accurately a larger population. For a particular solar array, the value $\hat{P}$ can be found either by measurement, i.e. by finding the maximum power value produced by the array over a given time period, or by using the rated array peak value. This latter value is provided by the PV array installer and is the peak export power under optimal conditions. Although $\hat{P}$ can be defined over any time period, the maximum annual value would be normal and this is used for all calculations in the analysis of the CLNR PV trials.

Because a DNO or supplier will typically only have knowledge of the rated capacity, it is important to understand when analysing solar PV generation profiles whether this value is a good predictor of the actual capacity as measured from data. This study thus investigates the CLNR test cell 5 solar array population and compares the measured and rated values of $\hat{P}$, i.e. $\hat{P}_m$ and $\hat{P}_r$ respectively, for the population.

## 2.2 Method

Rated peak solar PV array generation capacities have been provided by British Gas for a subset of the test cell 5 participants. Solar PV generation data for test cell 5 customers was inspected to give the measured value of $\hat{P}$, i.e. $\hat{P}_m$, for each customer. Where both values were available for customers, these two values (rated and measured) were compared. Histograms are also plotted to show the distribution of peak capacities for both sets of peak estimation. A detailed analysis was conducted to quantify the differences between the measured and rated peak values.

## 2.3 Data Set Observations

Data for a total of 157 customers in test cell 5 were available. Figure 1 shows a histogram of the measured solar peak generation $\hat{P}_m$ for all customers. Bin sizes of 0.1 kW were used.

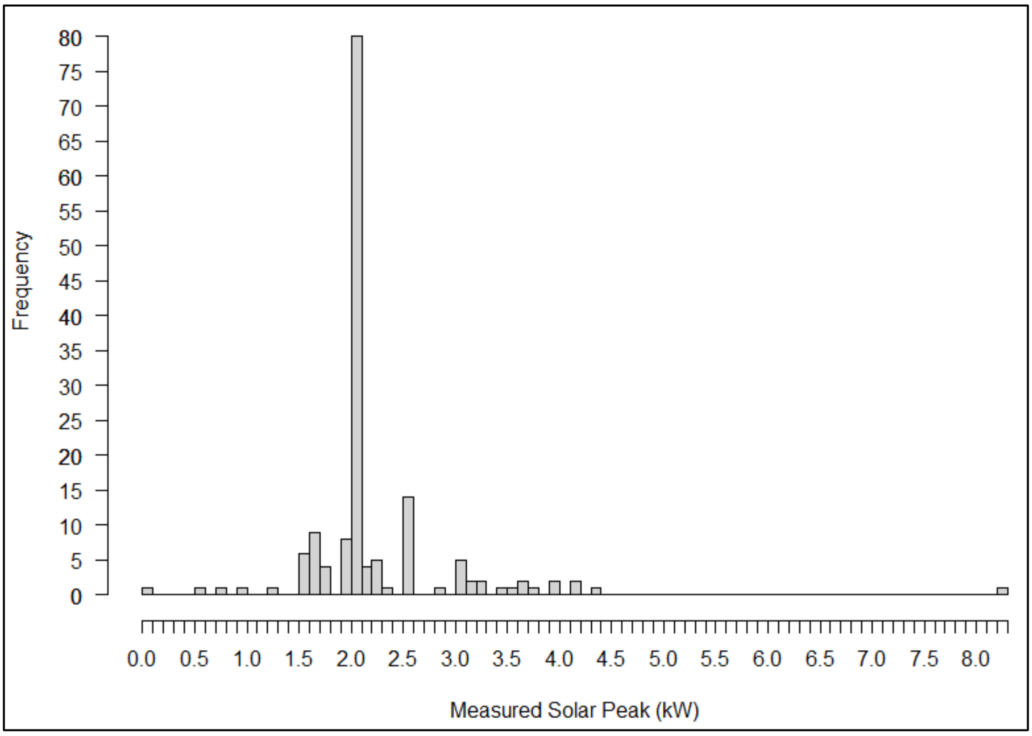


**Figure 1: Histogram of measured solar peak**

Measured solar peaks were mainly observed between 1.5 and 3.5 kW. A mean and median measured peak $\hat{P}_m$ of 2.22 and 2.03 kW respectively were observed for the sample. There seem to be two outliers; between 0.0 - 0.1 kW, and 8.2 - 8.3 kW.

Data for the rated peak was only available for 102 test cell 5 customers and a second histogram

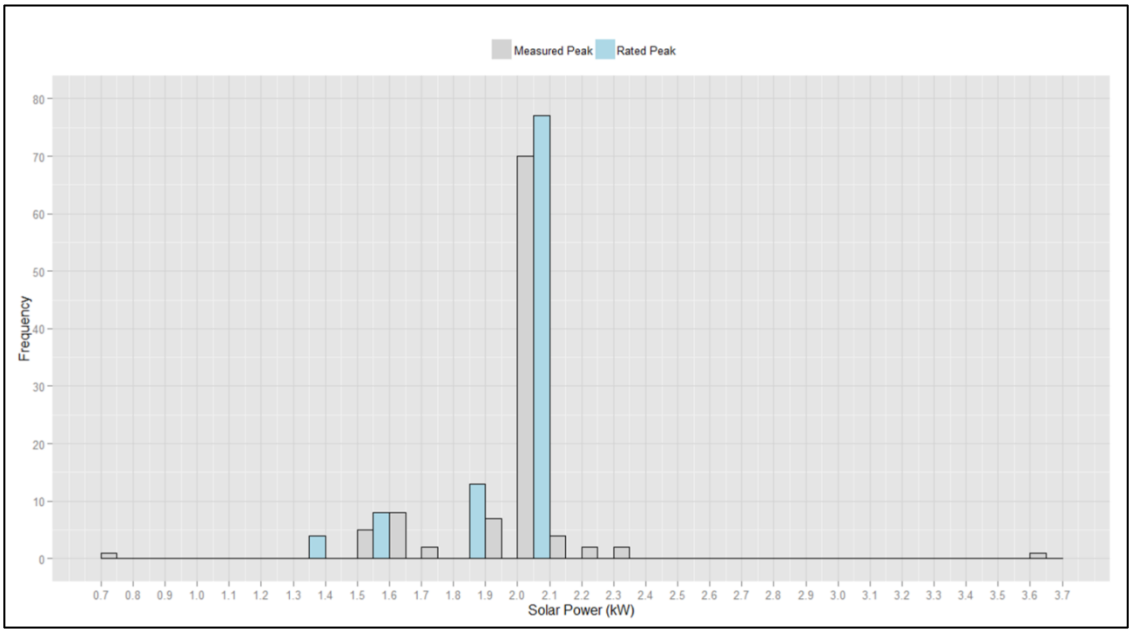


**Figure 2: Side-by-Side Histogram of Customer Measured and Rated Peaks**

(Figure 2) shows the measured solar peaks and rated peaks for these customers. Bin sizes of 0.1 kW were used.

Four distinct rated peaks were observed across the customer sample; 1.35 kW, 1.57 kW, 1.8 kW and 2.02 kW. Rated peaks might be expected to conform to discrete values, as arrays are comprised of a discrete number of panels which have a fixed output rating. In practice, arrays will likely be comprised of between 6 and 14 panels and will more often than not be an even number, giving about five possible different array sizes that would occur most frequently. Measured solar peaks fell within a range from 1.5 to 2.4 kW, with 2 outliers (between 0.5 to 1 kW and 3.5 to 4 kW respectively). Mean and median peaks of 1.98 kW and 2.02 kW, respectively, were observed.

Figure 3 shows a chart of the measured solar peak and rated capacity for each customer.

As also seen in and, the most common solar peak occurs between 2.0 to 2.5 kW. The deviation between measured and rated peak for most customers is small (less than 0.5 kW) with the exception of two customers (Customers 12 and 26). Customer 12 has a measured solar peak of 3.6 kW (approximately 1.6 kW higher than its installed capacity) whereas Customer 26 has a measured peak of approximately 0.7 kW.

An overall summary of the panel data is shown in Figure 3. This describes the number of customers

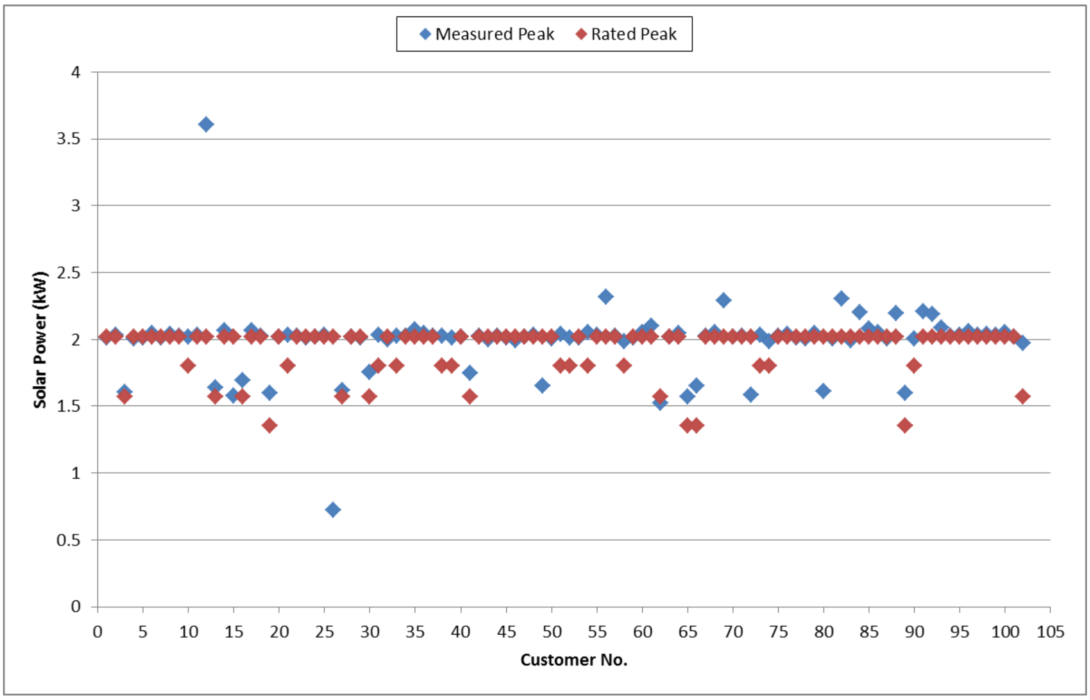


**Figure 3: Measured Solar Peak vs Rated Peak for Each Customer**

for which the measured peak is greater than, less than or equal to the rated peak. There is a notable discrepancy between the measured and rated peak values, with over two-thirds of the panels having a larger measured peak than rated. This is explored in more detail in the Analysis section below.

| | Peak Measured > Rated | Peak Measured < Rated | Peak Measured = Rated |
|---|---|---|---|
| **Number of customers** | 72 | 29 | 1 |
| **Mean difference (Measured - Rated) (kW)** | 0.12 | -0.11 | - |
| **Variance (kW)** | 0.04 | 0.07 | - |
| **Standard Deviation (kW)** | 0.2 | 0.27 | - |

**Table 1: Summary of Differences between Measured and Rated Peak Generation**

## 2.4 Analysis

In order to investigate how well the rated peak agrees with the measured peak, the entire data set was plotted. Figure 4 is a variation of Figure 3 and shows the rated peak vs the measured peak for the 102 solar panel arrays. The rated values fall into four categories 1.35, 1.57, 1.80 and 2.02 kW.

The validity of the classification of measured peak is questionable for several observations, notably the minimum and maximum extreme values observed for a rated value of 2.02 kW (denoted by the blue arrows in Figure 5). It is also possible that the values indicated by the magenta circle should have been classified with a rated peak value of 1.57 kW instead of the recorded 2.02kW. The observation indicated by the green circle in Figure 5 (a rated peak of 1.57 kW) perhaps is more likely to be a rated peak value of 1.80 kW; however this is speculation.

As it is not possible verify data entry for rated values, as these data are held by the PV installers, those observations relating to a rated peak value of 2.02 kW as mentioned above are assumed to be anomalous, and are removed from the data set. However the dubious value relating to a rated peak value of 1.57 kW is left and attention is drawn to it in the following analysis where appropriate.

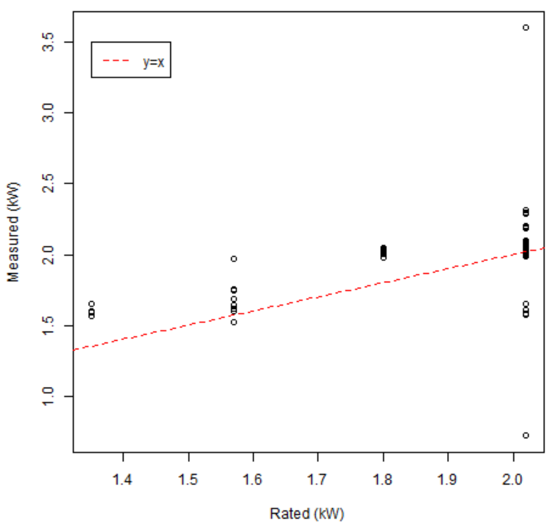


**Figure 4: Measured peak vs Rated peak for 102 solar panel arrays**

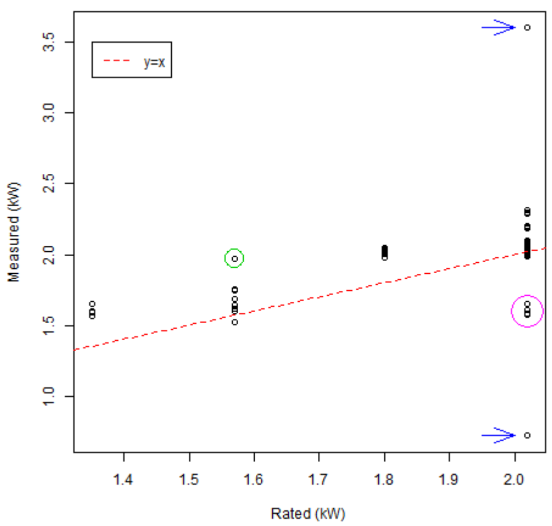


**Figure 5: Measured peak vs Rated peak for 102 solar panel arrays, annotated plot**

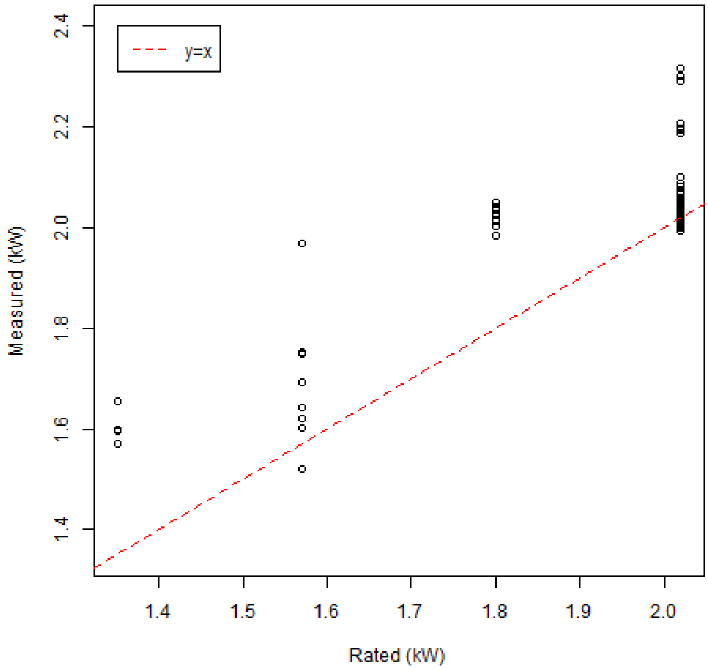


**Figure 6: Measured peak vs Rated peak for 96 solar panel arrays, modified data set**

Figure 6 shows the measured peak vs the rated peak for this modified data set, which now consists of 96 observations. Figure 6 shows that the majority of measured peak values appear to be greater than the rated peak values. It would therefore appear that the solar panel arrays on average are generating more power than their rating.

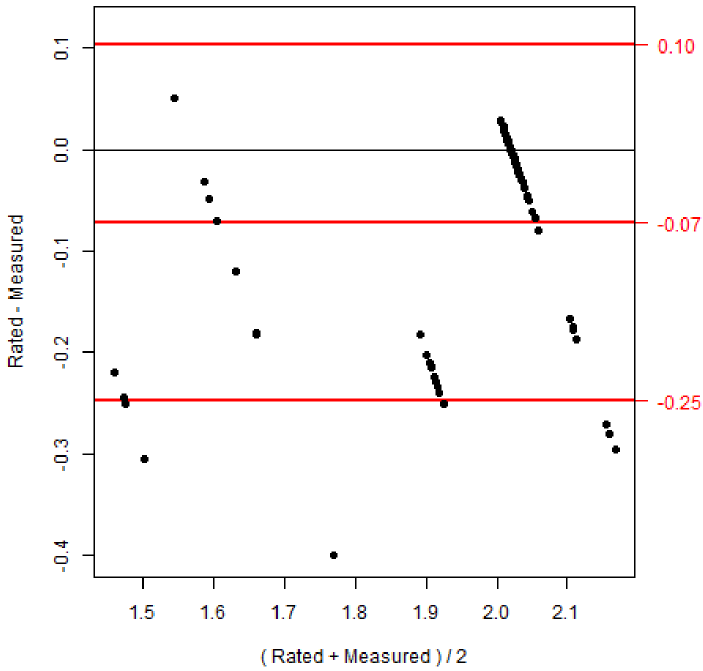

Figure 7 shows a Bland-Altman plot for the modified measured peak vs rated peak data set with 90% confidence limits. A Bland-Altman plot allows us to compare the difference between two methods which are measuring the same thing; the x-axis shows the average of the two methods whilst the y-axis shows the difference between the two methods. If both methods are in agreement with each other, then the majority of points should lie along the zero line. Looking at Figure 7 it can be seen that this is not the case for this data set, with the majority of points lying below zero on the y-axis; this indicates that the measured peak value is greater than the rated peak value. The solar panel arrays with the four rated values (1.35, 1.57, 1.80 and 2.02 kW) are clearly defined in Figure 7 by the four diagonal lines. It can be seen that for all four categories, measured values are greater than the rated values. The mean difference between the measured and rated peak values is -0.07, meaning the measured value is 0.07kW greater than the rated value, and a 90% confidence interval is given by (-0.25, 0.10). The one obvious outlier on the plot (with a Rated-Measured value of -0.4) belongs to the dubious observation with a rated peak value of 1.57kW identified above. Due to the fact that this observation may have been misclassified to begin with, the large difference between the two methods should be regarded as an anomaly. On the whole however it appears that the measured peak value is greater than the rated peak value and a suitable upper bound for the difference may be that the measured peak is 0.25kW greater than the rated peak.

It should be noted that the confidence interval contains the value zero, and as such there is no statistical significant difference between the measured and the rated peak. It may be that the measured peak is 0.25kW greater than the rated peak, although it could also be the case that the measured peak is 0.1kW less than the rated peak. In order to find a value such that there is a 90% chance that the measured peak would be less than the rated peak, 0.25kW could be added to the rated value; although this has the caveat that some measured values could be much less than this new rated value.

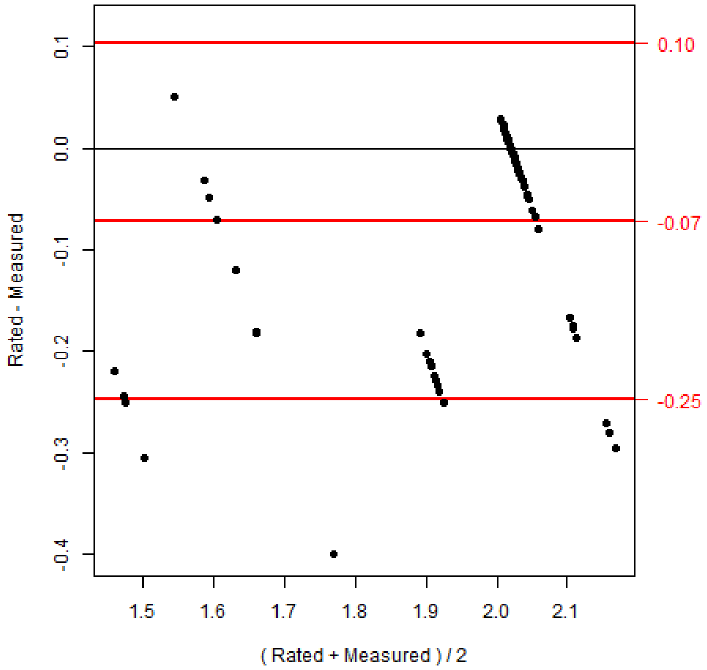


**Figure 7: Bland-Altman plot for the Rated peak vs Measured peak with a 90% confidence interval (outer red lines) and the mean difference (centre red line) for the modified data set**

| Contingency table | | | |
|---|---|---|---|
| **Rated value** | **Measured ≤ Rated** | **Measured > Rated** | **Total** |
| 1.35kW | 0 | 4 | 4 |
| 1.57kW | 1 | 7 | 8 |
| 1.80kW | 0 | 13 | 13 |
| 2.02kW | 24 | 47 | 71 |
| **Total** | 25 | 71 | 96 |

**Table 2: Contingency table for Measured peak vs Rated peak for the modified data set**

Table 2 shows a contingency table that elaborates on the information presented earlier in Table 1. Overall, 76% of measured peak values were greater than the rated peak value. For two rated values (1.35 and 1.80kW) it can be observed that all the measured values were greater than the rated value, although the sample size is small for these rated values and this result should not be taken too seriously. The rated value with the smallest percentage of measured peak values greater than the rated peak value is 2.02kW, although this is still 66% of the total.

# 3 Conclusions

Data from the CLNR solar trials show that the measured peak capacity of domestic solar panels does not match the installer-declared peak capacity. This study examines this discrepancy and attempts to quantify the extent to which these two values (measured and declared) differ. It might be expected that the measured peak output from a PV array might be lower than the rated peak because as solar arrays age, their peak output is expected to reduce. However this study finds that for the population of PV arrays for which rated peak values were available, more arrays produce peak power that is higher than their rated value than produce lower power, in a ratio of about 3:1.

The study examined the rated peak value for a series of solar PV observations. A modified data set consisting of 96 observations after removals of anomalous data was analysed. Using this modified data set it was observed that the measured peak value is often greater than the rated peak value, with 76% of measured peak values being greater than the corresponding rated peak value of the PV installation. A 90% confidence interval for the difference between the two peak values, i.e. rated peak minus measured peak, is (-0.25, 0.10), suggesting that a network designer may wish to consider that a solar panel array may output a peak value of 0.25kW greater than its rated value.

Analysis of CLNR PV trials used the peak power output as a normalisation parameter to allow differently-sized PV arrays to be compared and grouped. Using the rated peak as a normalisation parameter will thus result in computed specific energy, $e$, and power, $p$, values greater than 1.0. The CLNR PV trial analysis thus used measured peak power as a normalising factor. The rated peak for a specific array could then be used an estimator of the measured peak when converting normalised to absolute generation, however further work will be needed to formalise a method of estimating the actual peak and the error in the estimation from the rated peak and this is not included in this study.

When designing LV networks it would not be unreasonable to assume that there will be at least one time during the year when the peak power generation for a group of PV panels is reached simultaneously. In this case, the median rated peak for the group would seem to be a reasonable estimator of the true median, and this could be used to estimate group generation power. However this does not take into account other factors such as PV installed in a batch using a common supplier (where the true peak for the group may be consistently higher or lower than the rated peak), and variability in house orientation, which would reduce the chance of a simultaneous peak.

Using the data available in this study, in order to find a value such that there is a 90% chance that the measured peak would be less than the rated peak, 0.25kW could be added to the rated value; although this has the caveat that some measured values could be much less than this new rated value. Given the limited range of known installed capacities in this study however, this value may also have limited applicability outside of this sample. A broader study using the same method would be recommended to determine a robust working “rule of thumb” for the difference between installed and actual peak domestic solar generation capacity.

# 4 Diversity Study: Applying the After Diversity Maximum Demand (ADMD) Methodology to Highly-Distributed PV Generation

## 4.1 Introduction

Test Cells 5 and 20 of the CLNR field trials provided 1-minute resolution average power output data for PV units installed at domestic sites across the UK. This report considers the effect of the physical diversity of the panels on total aggregated power output during high generating conditions. Diversity in nominal panel output arises due to variations in orientation, shading, degradation, PV performance characteristics and capacity rating inaccuracies. The concept of group diversity is well known from the demand perspective, where it is called After Diversity Maximum Demand (ADMD), but is less well understood in the context of group micro-generation such as residential PV. The acronym 'ADMD' is used here even when referring to group generation since the context is clear. This work uses empirical network data to quantify PV group diversity.

## 4.2 The Dataset

Domestic PV output data was collected for 310 sites across the UK, each minute, over the course of approximately 731 consecutive days. The data for each customer therefore consisted of one 731 x 1440 array of power output values, in watts. For consistency, this array often included several days of zero data inserted as padding at the beginning and end of the 731 day period, whenever real data was missing. This does not affect the results of this work. Site locations around the UK were not known.

Output capacity ratings were known for a total of 165 sites. The vast majority of these sites were rated 1.5kW, 1.8kW or 2.0kW, although other ratings observed were 1.3kW, 2.7kW and 3.6kW. Data from sites with unknown rating was of no value in this study and was discarded.

## 4.3 Identifying Peak Days

In order to investigate exceptionally high PV group generation events, days on which the majority of sites generated high power were identified. The data for each site was searched and two metrics of maximum power output were used to identify the following days for each site:

i) the day with the highest mean power output over the entire two year period
ii) the day with the absolute maximum power output over the entire two year period

Histograms showing the frequency of these two high generation days over all sites are shown in Figures 8 and 9. They mostly occur around days 340-420.

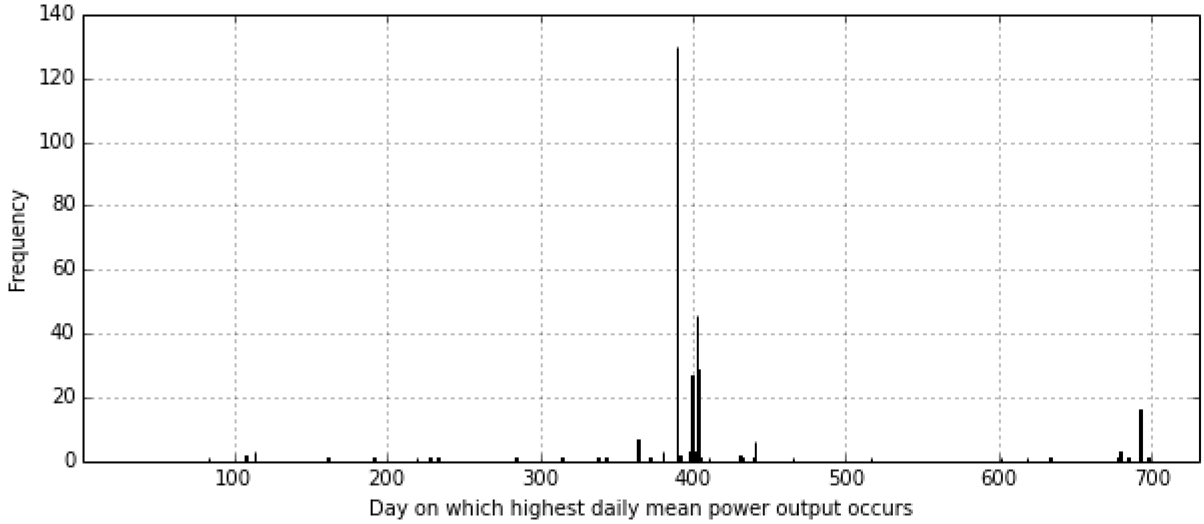


**Figure 8: Histogram of maximum daily mean power output days**

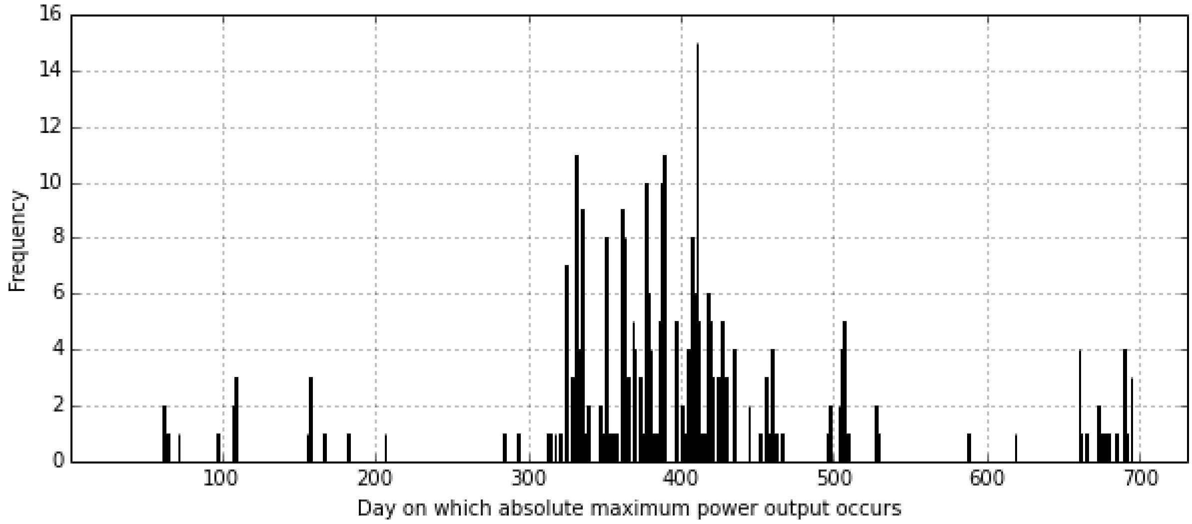


**Figure 9: Histogram of maximum power output days**

Figure 8 shows that the day on which the highest number of sites produced their maximum average output is day 389, with 130 sites in this group. Days 399, 402 and 403 are the next highest.

Figure 9 shows that the day on which the highest number of sites achieved their absolute maximum output is day 410, with 15 sites in this group. Days 331, 377, 387 and 388 are the next highest.

Day 389 probably corresponds to a summer day with uninterrupted sunshine all day at all locations, while day 410 probably corresponds to a summer day on which there were periods of particularly intense sunshine in many locations but not necessarily at the same time, and not lasting all day.

Note that although the data collection period is two years, there is only one cluster of these high output days. This is because the data is generally patchy near the beginning and end of the data collection period, due to logistical and operational reasons (staggered or delayed power meter installations for example).

Figure 10 shows a selection of typical raw output profiles on these two peak days, 389 and 410. The profiles are consistent with the descriptions given.

Output profiles for several other high generation days were inspected for comparison, and were seen to be similar. It is therefore reasonable to consider days 389 and 410 to be fairly representative, from a network perspective, of the two types of peak generation days considered here.

## 4.4 Output Profiles on Peak Days

The output profiles were then normalized by capacity rating in order to analyse the underlying diversity of PV panels regardless of installation size.

On both peak days 389 and 410, a total of 138 PV units produced high quality output profiles, and 27 were discarded due to having zero or clearly faulty output. However, this is not an insignificant issue: if it is the PV equipment itself which is malfunctioning, not just the data recording equipment, then these data points *should* be included in the diversity analysis. This was not known here and the data was simply removed.

Figures 11 and 12 show all valid output profiles for days 389 and 410 respectively. In each case the upper chart shows all the capacity-normalized output profiles plotted together, while the lower chart shows the day’s mean profile.

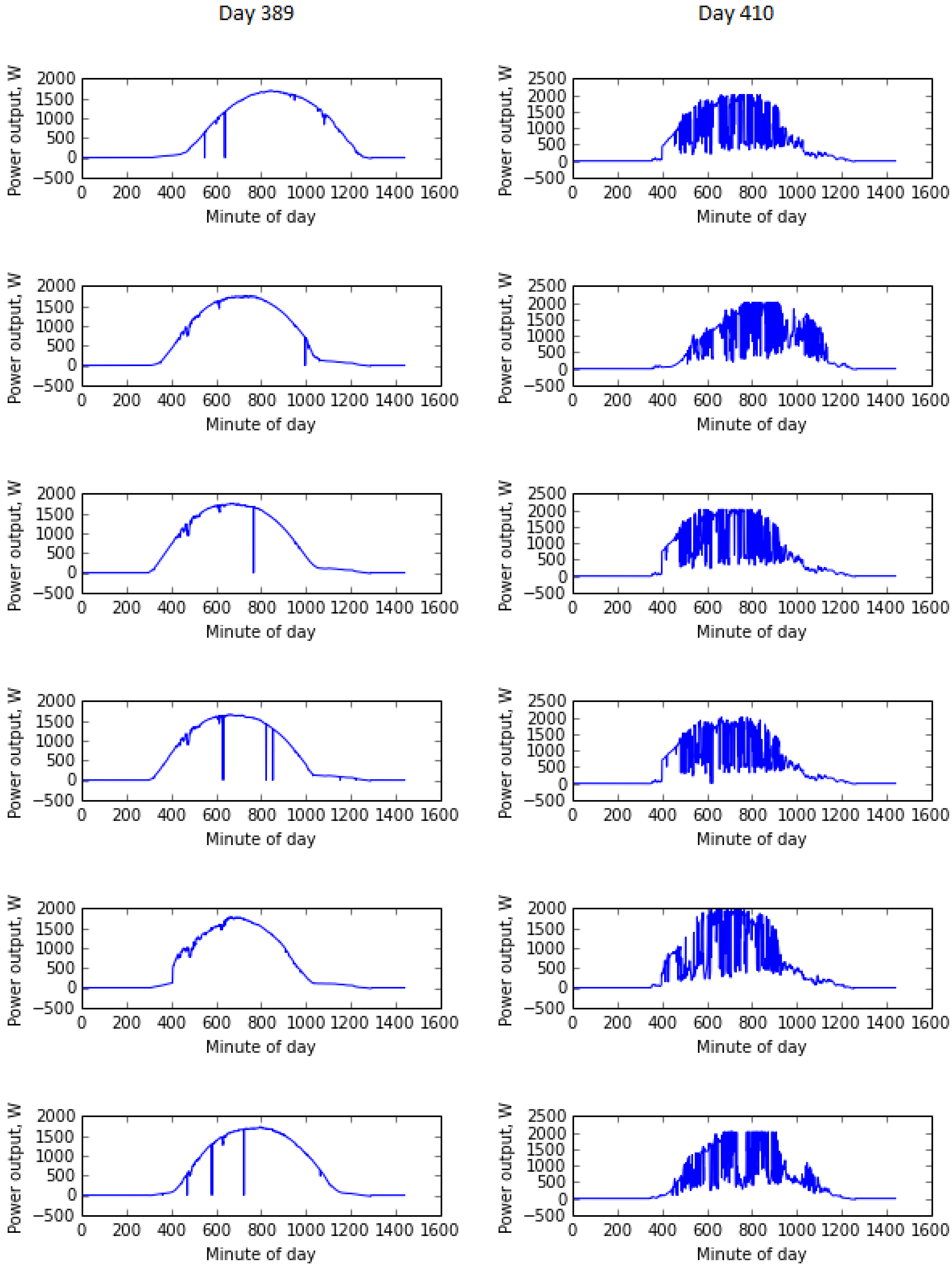


**Figure 10: Selection of output profiles for the two peak days**

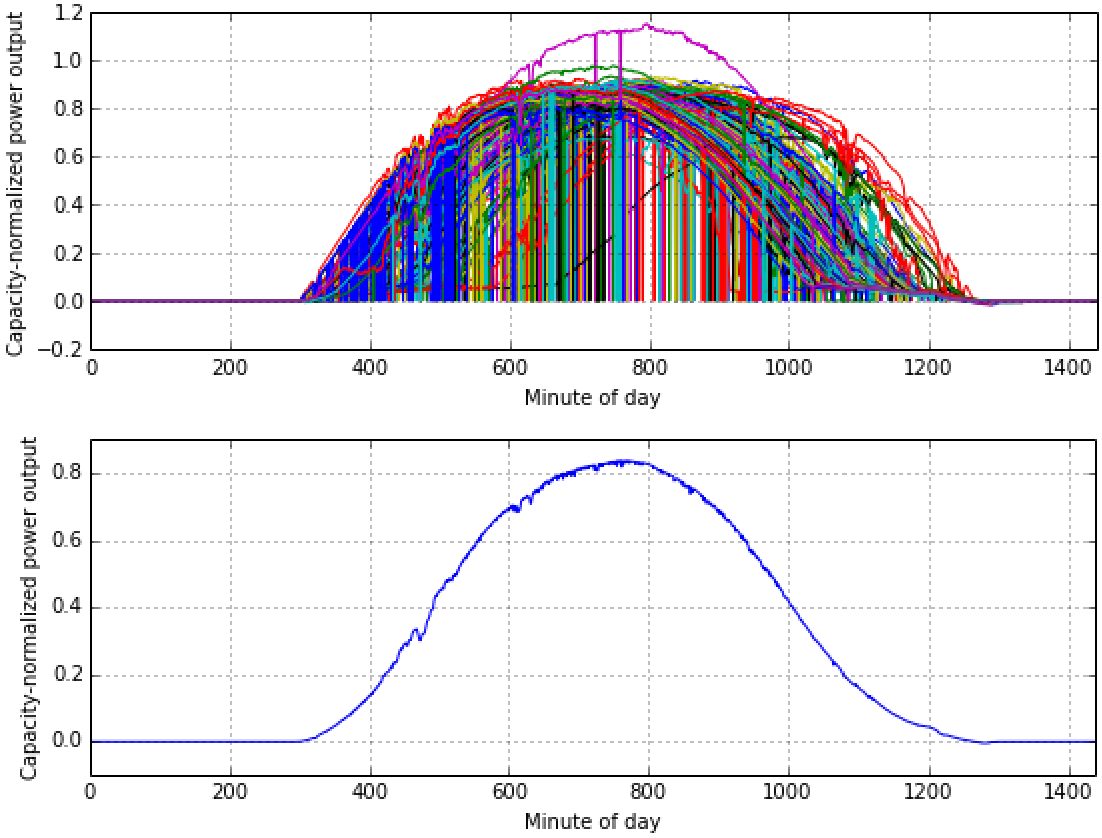


**Figure 11: profiles and mean profile on day 389**

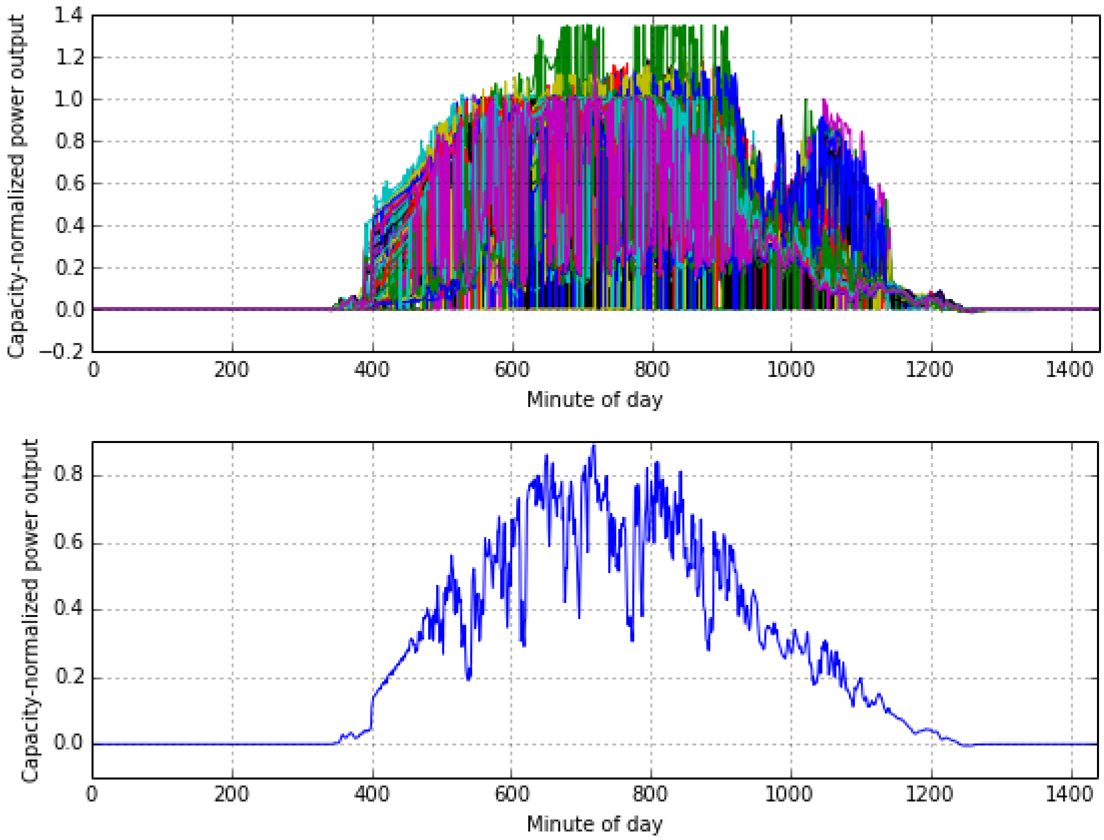


**Figure 12: profiles and mean profile on day 410**

Figure 11 shows that when all sites are producing near their maximum possible output all day, the combined mean profile is smoother, broader and lower than the individual site profiles. The peak of this mean profile occurs at minute 759 of the day and has value 0.838.

Figure 12 shows that for day 410 the picture is less clear, because the individual site profiles are very erratic. However, the mean profile does appear to be somewhat smoother, broader and lower than the individual profiles. The peak of this mean profile occurs at minute 719 and has value 0.891. This is considerably less than the peaks of many of the individual profiles, but higher than the mean profile on day 389.

## 4.5 After Diversity Maximum Demand

The process for generating the ADMD curve for a given time period (in this case just one single day) is as follows:

1) Pick a customer group size N.
2) Repeat the following steps a large number *i* of times:
   a) Randomly select a group of N customers, with replacement,
   b) For these N customers, calculate the mean profile for the day,
   c) Find the peak of this mean profile over the whole day; store this value in an array associated with N.

This builds up an array of *i* values associated with this N. In other words, for each customer group size N, a distribution of *i* points is generated, each of which is the *day maximum* of the *mean profile* of *a randomly selected group of N customers*.

3) Repeat this process for all values of N from 1 up to some chosen limit $N_{max}$, thereby generating the entire ADMD curve for that day. Plot the data.

After coding and testing, an upper limit of $N_{max}$ = 100 was chosen for customer group size, and the number of repeat random samplings for each N was chosen to be *i* = 2000. The process is quite computationally intensive, so there is a trade-off between $N_{max}$, *i* and time duration analysed. In this work, due to constraints on both coding time and program running time, only individual days were analysed. This is not ideal, since the ADMD procedure should really be applied to the whole time duration for which data exists.

Figures 13 and 14 show the ADMD curves for days 389 and 410 respectively. For each customer group size N a box-and-whisker plot is drawn, illustrating the corresponding distribution of ADMD values. The red line is the median, the blue box is the interquartile range, and the 'whisker' end points are the maximum and minimum outlier values of the distribution.

On day 389, for N=100, the ADMD distribution has: mean = 0.839, max = 0.865, min = 0.809.

On day 410, for N=100, the ADMD distribution has: mean = 0.900, max = 0.949, min = 0.848.

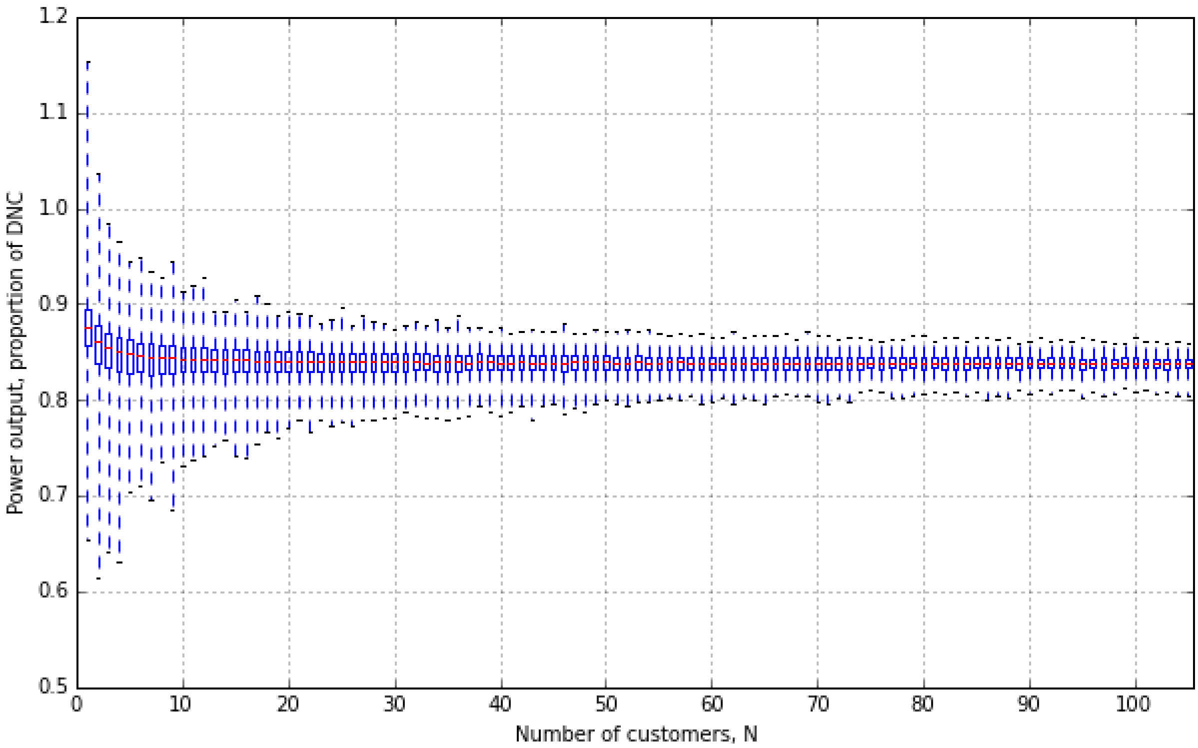


**Figure 13: ADMD curve for day 389. Mean for N=100 is 0.839**

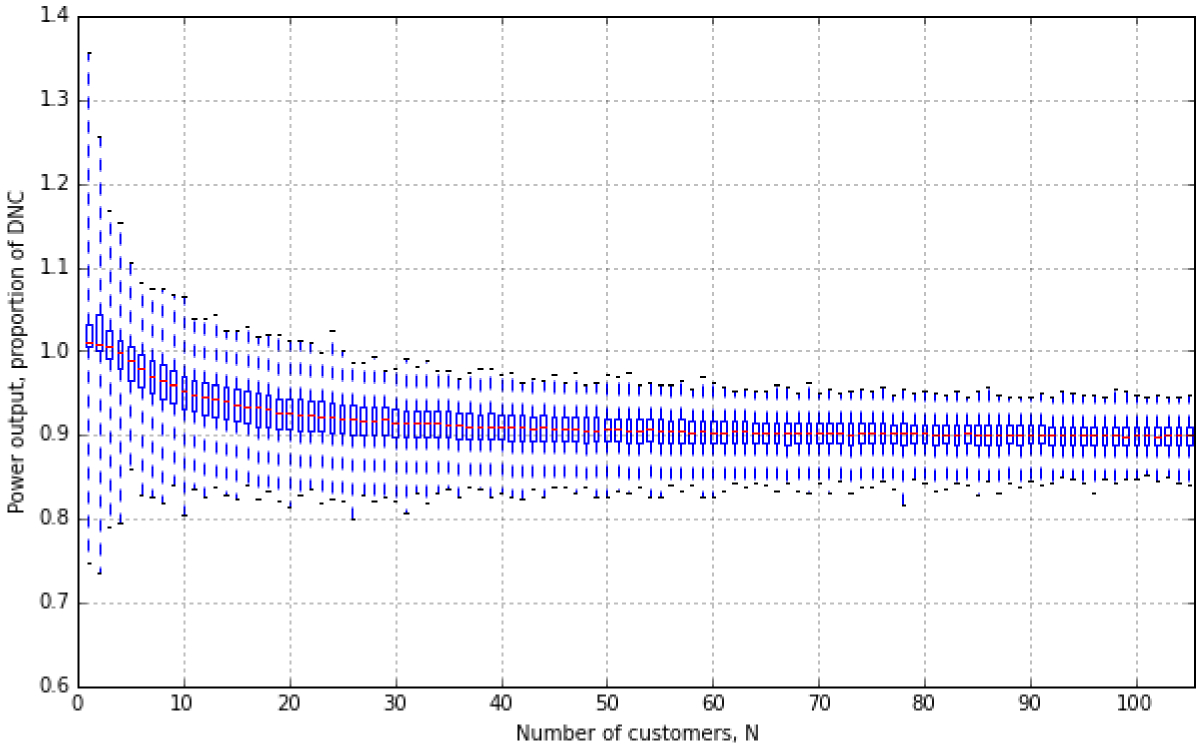


**Figure 14: ADMD curve for day 410. Mean for N=100 is 0.900**

For comparison, the same procedure was implemented for the next most frequent absolute maximum days (387 and 388), and the next most frequent maximum average output days (402 and 403). The results are shown in Figures 15-18, and the corresponding mean profiles on the given days are shown in Figures 19-22.

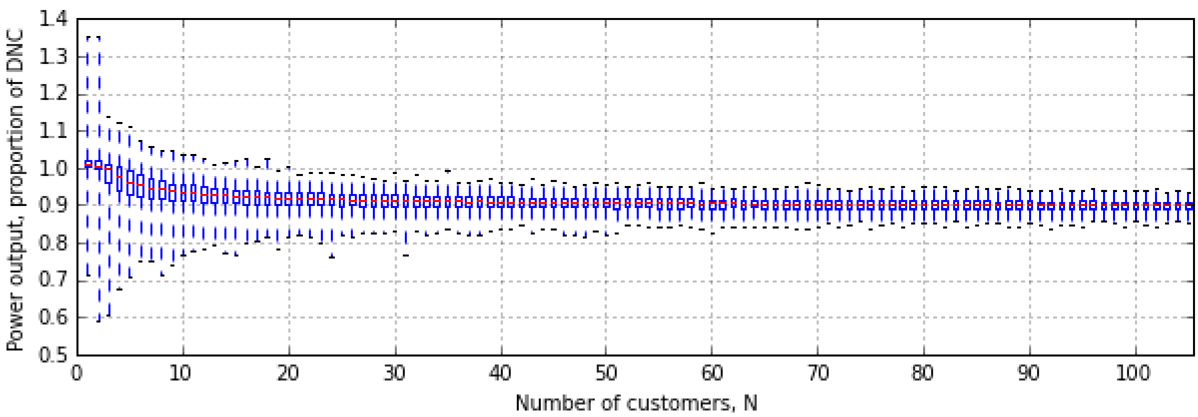


**Figure 15: ADMD curve for Day 387. For N=100, mean = 0.901**

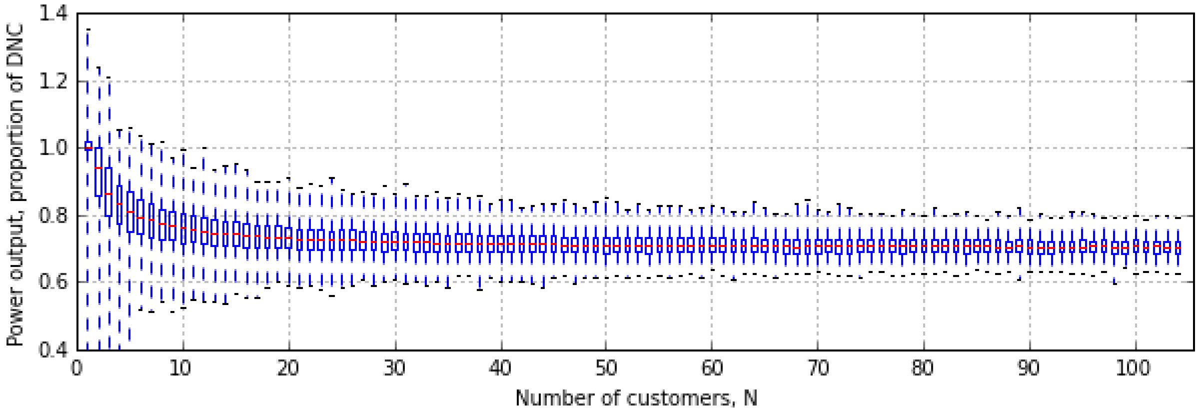


**Figure 16: ADMD curve for Day 388. For N=100, mean = 0.708**

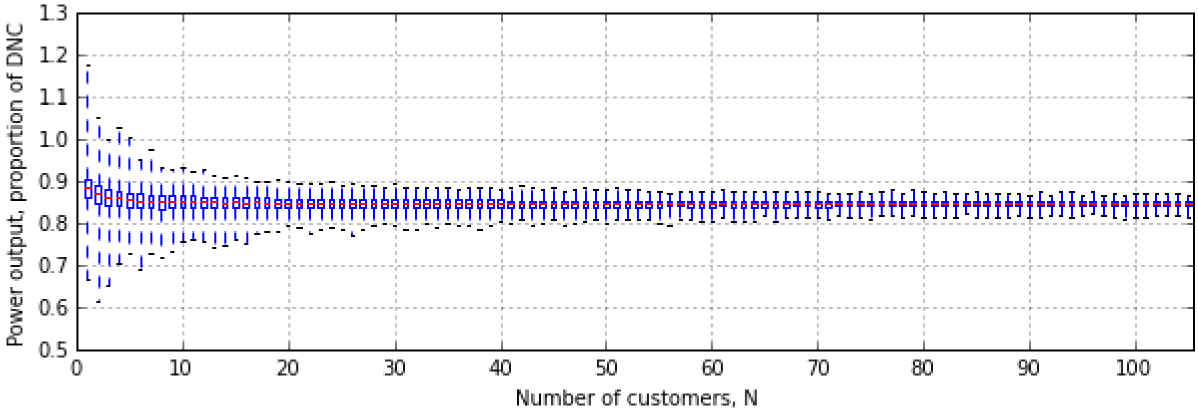


**Figure 17: Day 402. For N=100, mean = 0.846**

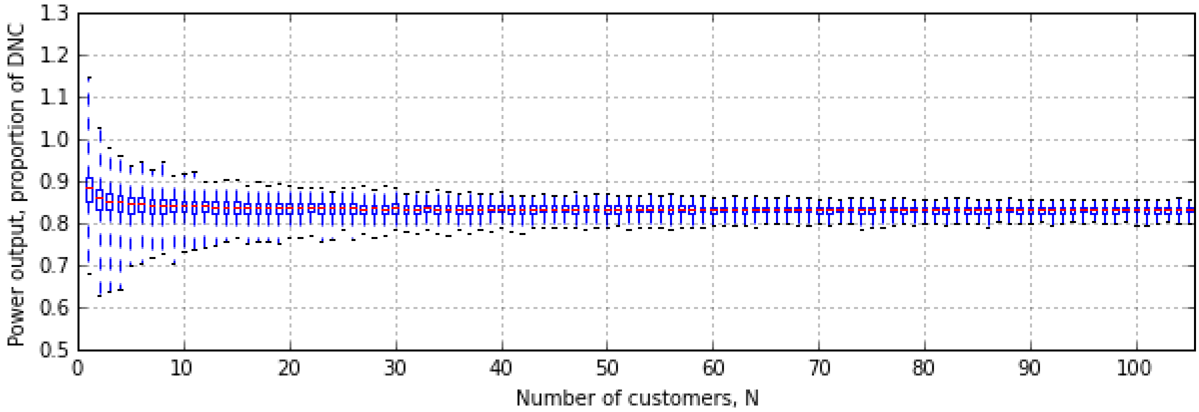


**Figure 18: Day 403. For N=100, mean = 0.832**

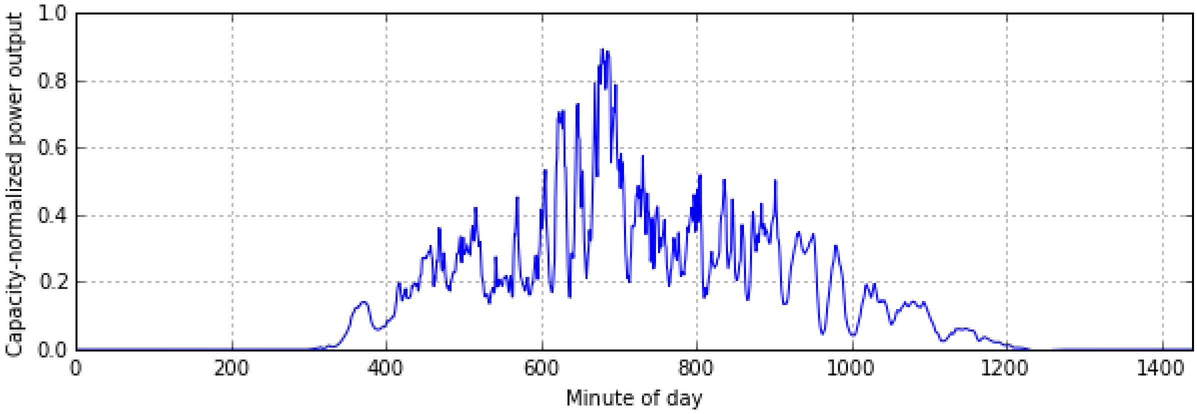


**Figure 19: Day 387 mean profile (N=100 ADMD mean = 0.901)**

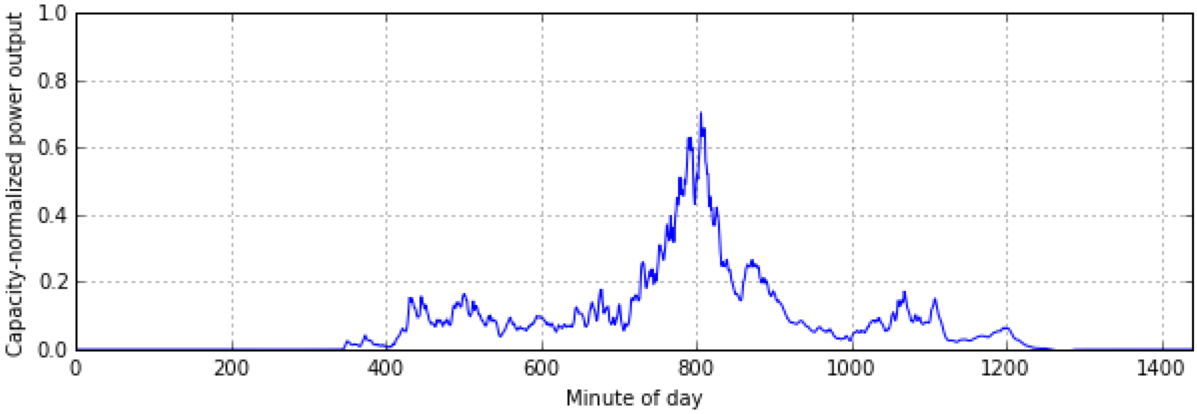


**Figure 20: Day 388 mean profile (N=100 ADMD mean = 0.708)**

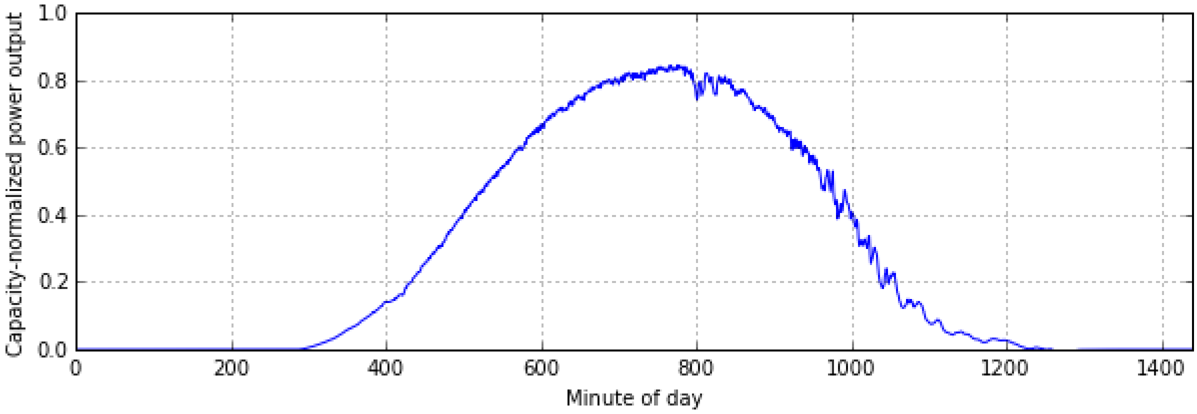


**Figure 21: Day 402 mean profile (N=100 ADMD mean = 0.846)**

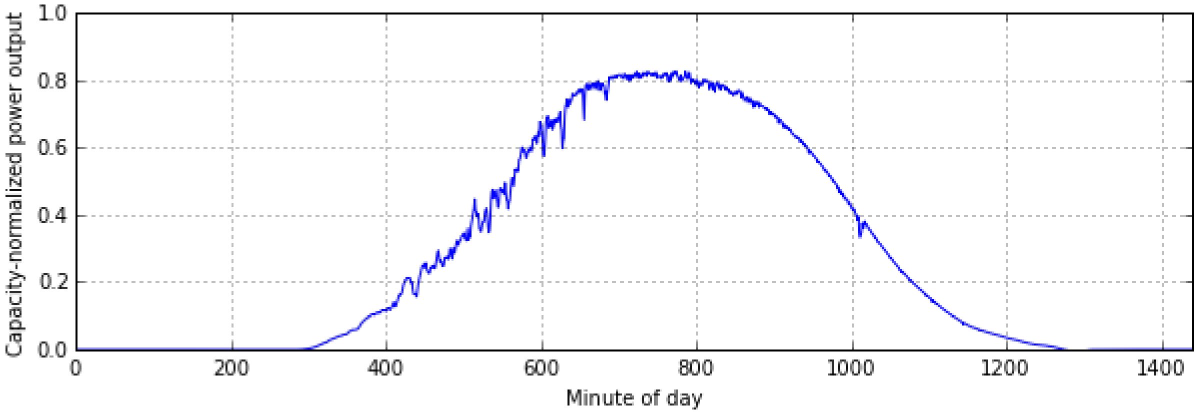


**Figure 22: Day 403 mean profile (N=100 ADMD mean = 0.832)**

For these few days analysed, it is clear that the smooth, broad, symmetrical mean profiles which correspond to high mean power output days also correspond to ADMD mean values for N=100 of around 84% of DNC. On the other hand, the erratic, peaky mean profiles which correspond to days of absolute maximum power output produce a range of ADMD mean values for N=100; 71% of DNC and 90% of DNC were observed, and the mean profiles reflect these differences.

## 4.6 Conclusions

For this dataset, using the methods described here, the highest observed mean After Diversity Maximum Generation power for groups of around 100 PV units is 90% of DNC. This maximum value was observed on day 410 (Figure 14) and day 387 (Figure 15). These are days on which the largest numbers of sites achieved their absolute maximum power output for the entire two year period under consideration. On these days, all sites' output profiles, and indeed the mean profile, are very 'peaky' and erratic (Figures 10, 12, 19 and 20). However, the mean ADMD value for large customer groups varies considerably from day to day.

Due to the specific manner in which the work progressed, the ADMD methodology was not simply applied to the whole two year period, which would give one definitive result, but instead was applied to individual high generation days. Thus the results are not exhaustive and with more time the study should be extended to cover longer time periods (though it is also possible that the full analysis would be prohibitively computationally intensive - this can't be known at this stage). Given all this, two metrics were used to locate 'high generation' days, and it is likely that the analysis of these days undertaken here has indeed produced the most extreme ADMD curves possible. In addition, this approach of applying the ADMD procedure on single days is useful because it allows greater insight in to the specific correlations between daily mean profiles and their associated ADMD curves and mean values.

For days 389, 402 and 403, on which the highest numbers of sites produced their highest mean power, the mean ADMD value for 100 customers was only 83-84% of DNC, significantly less than the 90% figure observed on days 410 and 387. The cause of this difference was not investigated in this work. It could simply be due to differences in characteristic atmospheric conditions on these types of days. Or it could be due to variations in panel output characteristics in different weather conditions, since Figure 12 shows many profiles generating at well above 100%, while Figure 11 shows only one profile over 100%. Further analysis would be required to separate out the precise causes of these differences, but the clear correlations seen here indicate that there is high potential for predicting ADMD values accurately on a day by day basis.