**Summary of the data collected**

Key

Data collected as part of the project run by UK Power Networks:

Validation of Photovoltaic Connection Assessment Tool

<https://www.ofgem.gov.uk/ofgem-publications/93938/pvtoolcdrfinal-pdf>

The project collected a rich dataset at domestic sites with Solar Panels. The data set comprises of 25,775 days of data, and over 171 million individual measurements.

Key stats about the dataset:

* 20 substations and 10 domestic premises
* 480 days of measurement - 27 July 2013 to 19 November 2014
* 10 minute intervals over all time recorded, 1 minute intervals in summer 2014
* 10-minute measurements prior to 10 June 2014, aggregated to hourly minima and maxima

Several important findings are presented in the final project report (link above). For any further queries please email [innovation@ukpowernetworks.co.uk](mailto:innovation@ukpowernetworks.co.uk)

**PV Tool Dataset – Explanatory Notes**

Datasets have been cleansed and processed as follows:

* Aggregated the three datasets provided by Dan
  + Dataset #1 covering 27/07/13 to 27/07/14 (this is the one that was corrupted)
  + Dataset #2 extracted 30/09/13 to 30/09/14
  + Dataset #3 extracted 30/09/14 to 19/11/14
* Deleted the corrupted rows from Dataset #1
  + From 27/07/13 to 31/10/13 there were 696 corrupted rows and 143,795 good rows
  + The magic regexp to delete the corrupted rows is ^[^,T]\*(,[^,]\*){89,100}$
* Deleted duplicate rows where datasets have overlapped
* Deleted blank rows
* Split the timestamp into separate columns for year, month, day, day of week, hour, minute – to make it easier to group data by hour, day of week, month etc.
* Filtered out any data that pre-dates the device’s installation date (e.g. data recorded during bench testing)
* For the monitors that were moved during the trial, assigned the data to the correct location on the correct dates. These rows have retained their original serial number (with an ‘X’ appended), so make sure you group by substation and feeder name, not by serial number.
* Filtered out devices/cards that didn’t have any CT’s connected
* For customer endpoints, transformed A/B/C channels into Generation/Import channels according to how the CTs are connected for that device, and transformed the polarity so that Generation is always +ve, Import is always +ve, and Export is always –ve.
* Filtered out erroneous voltage, current, and thdV measurements:
  + All voltage measurements <200V
  + All current measurements >500A
  + All thdV measurements whenever the corresponding voltage measurement is <200V. (This brings the maximum thd across the entire dataset from 900% down to about 4%)
* For endpoints, provided the corresponding substation busbar voltages in the same row and calculated the voltage rise
* Fixed the problem where numbers were truncated (not rounded) to 2 decimal places. *(Apparently this is a design “feature” of MS Access, and to overcome it you have to force all the numbers into string format before exporting.)*
* Removed commas from substation/feeder names – this was causing issues when exporting to CSV (comma-separated-values) format
* Corrected for the CT polarity that was changed at Alverstone close: Swapped P\_GEN\_MIN with P\_GEN\_MAX and multiplied both by -1 for dates before 2014-01-31 12:00:00
* Deleted erroneous 1min data on 08/05/2014

**Explanation of Field Names**

For all rows:

|  |  |
| --- | --- |
| SerialNo | Serial Number of the LvaDevice |
| Substation | Substation Name |
| Type2 | "Feeder" (At the substation), or "Endpoint" |
| To | Feeder name, or for Endpoints this will say "Customer" or "Network". (Network means the ones we were previously calling service cables, I'm now calling them Network Endpoints because service cable implies they are only looking at one customer, and some of them are looking at multiple phases and multiple customers.) |
| datetime | Date & Time |
| t\_date | Date |
| t\_time | Time |
| d\_y | Year |
| d\_m | Month |
| d\_d | Day |
| d\_w | Day of Week |
| t\_h | Hour |
| t\_m | Minute |
| VA,VB,VC | A/B/C Phase Voltage (V) |
| IA,IB,IC,IN | A/B/C Phase Current (A) |
| PA,PB,PC | A/B/C Phase Real Power (kW) |
| QA,QB,QC | A/B/C Phase Reactive Power (kvar) |
| SA,SB,SC | A/B/C Phase Apparent Power (kVA) |
| thdVA,thdVB,thdVC | A/B/C Phase Voltage Total Harmonic Distortion (%) |
| thdIA,thdIB,thdIC | A/B/C Phase Current Total Harmonic Distortion (%) |
| f | A Phase Frequency (Hz) |
| Ignore\_VoltageNotConnected | Flags records that come from LVAdevices with no voltage connected (you can ignore this field) |
| Ignore\_CurrentNotConnected | Flags records that come from LVAdevices with no current connected (you can ignore this field) |
| \_Filtered | Voltage Columns = Voltages <200V removed  Current Columns = Currents >500A removed  thdV Columns = thdV measurements removed whenever the corresponding voltage is <200V |

For All Endpoints:

|  |  |
| --- | --- |
| Substaton\_ | Indicates measurements taken at the substation |
| Rise fields | Difference between endpoint voltage and substation voltage.  For customer endpoints, the customer’s (single-phase) voltage is compared to all three phases at the substation.  For network endpoints, phase A at the endpoint is compared to phase A at the substation, etc. NB I have not confirmed that this relationship is correct, phase A at the endpoint may in fact correspond to phase B or C at the substation. |

For Customer Endpoints:

|  |  |
| --- | --- |
| V | Single Phase Voltage |
| I\_GEN | Generator Current (always +ve regardless of direction) |
| I\_IMPORT | Import Current (always +ve regardless of direction) |
| P\_GEN, Q\_GEN, S\_GEN | Generator Real Power (kW), Reactive Power (kvar), Apparent Power (kVA).  +ve = Generating  Should never be -ve |
| P\_IMPORT, Q\_IMPORT, S\_IMPORT | Imported Real Power (kW), Reactive Power (kvar), Apparent Power (kVA).  +ve = Import  -ve = Export |

**Customer Endpoint Apparent vs Registered sizes**

|  |  |  |
| --- | --- | --- |
| Substation | Apparent PV Size | Size in G83 Register |
| Bancroft Close | 3.50 | 1.89 |
| Forest Road | 3.00 | 3.29 |
| Suffolk Road | 0.50 | 1.52 |
| Alverston Close | 3.00 | 3.29 |
| YMCA | 0.45 | 0.60 |
| Maple Drive East | 4.00 | 3.83 |

**Site-specific Issues**

|  |  |
| --- | --- |
| Site | Notes |
| Old Mill Nordelph | * This is a 2 phase transformer * Endpoint current measurements appear to be scaled incorrectly (up to 1000A) |
| Upper staplefield common | * This is supposed to be a baseline site with no PV installed |
| Rookery Farm | This is an industrial/commercial estate with lots |

**Site Listing**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Site Name** | **Lic. Area** | **Type** | **Tx Size (kVA)** | **LV Cable Type** | **PV (kW)** | **Customer Monitor** | **Weather Station** |
| ALVERSTON CLOSE | SPN | Ground | 500 | 185H Al | 24 | Premises | - |
| BANCROFT CLOSE | EPN | Ground | 315 |  | 49 | Premises | - |
| BANKFIELD WAY | SPN | Ground | 1000 | 300 Al wv | 86 | - | - |
| CARTERS MEAD | EPN | Ground | 500 |  | 83 | - | - |
| CHAPEL LN GT GLEMHAM | EPN | Pole | 100 | 0.1 Cu | 15 | Pole | - |
| EAST HILLS ROAD COSTESSEY | EPN | Ground | 315 |  | 116 | - | - |
| ELM CRES COLCHESTER | EPN | Ground | 500 |  | 92 | - | Yes |
| FAIRVIEW ROAD | SPN | Ground | 315 | 300 Al wv | 56 | - | Yes |
| FOREST ROAD COLCHESTER | EPN | Ground | 1000 |  | 85 | Premises | Yes |
| MAPLE DRIVE EAST | SPN | Ground | 315 | 120H Al | 63 | Premises | Yes |
| OLD MILL NORDELPH | EPN | Pole | 50 | 0.1 Cu | 14 | Pole | - |
| PRIESTHAWES PMT | SPN | Pole | 100 | 95 ABC | 23 | Pole | - |
| RAMPLING COURT | SPN | Ground | 1000 | 300 Al wv | 0 | - | - |
| ROOKERY FARM BESTHORPE | EPN | Ground |  | 185 Al wv |  | - | - |
| SOUTHCROFT HETHERSETT | EPN | Ground | 300 |  | 115 | - | - |
| SUFFOLK ROAD | SPN | Ground | 1000 | 0.2 Cu | 24 | Premises | - |
| UPPER STAPLEFIELD COMMON | SPN | Pole | 100 | 95 ABC | 0 | - | - |
| VILLAGE BIRCHAM NEWTON | EPN | Pole | 200 | 95 ABC - 50 ABC | 18 | Pole | - |
| WARNINGLID LANE | SPN | Pole | 200 | 0.05 Cu | 23 | - | - |
| YMCA | SPN | Ground | 500 | 185H Al | 25 | Premises | Yes |