Introduction

The aim of this project is to create a model trained on data collected from the house with PV installation and weather data from the same location which would be able to predict the PV installation power generation depending on grid and weather conditions. Also feature importance analysis was performed.

The data used in this project were primary collected for "Validation of Photovoltaic (PV) Connection Assessment Tool" project carried out by UK Power Networks. Objectives of the primary project were as follows:

- "Validate UK Power Networks' guidelines for assessing PV connection requests and develop a formal policy.
- Develop a better understanding of the impact (including weather-related behaviour) that PV clusters have on the LV network by monitoring 20 secondary substations and 10 PV connection points.
- Understand how information available to PV installers could be used by DNOs.
- Gain a better understanding of the solutions available to address network constraints."

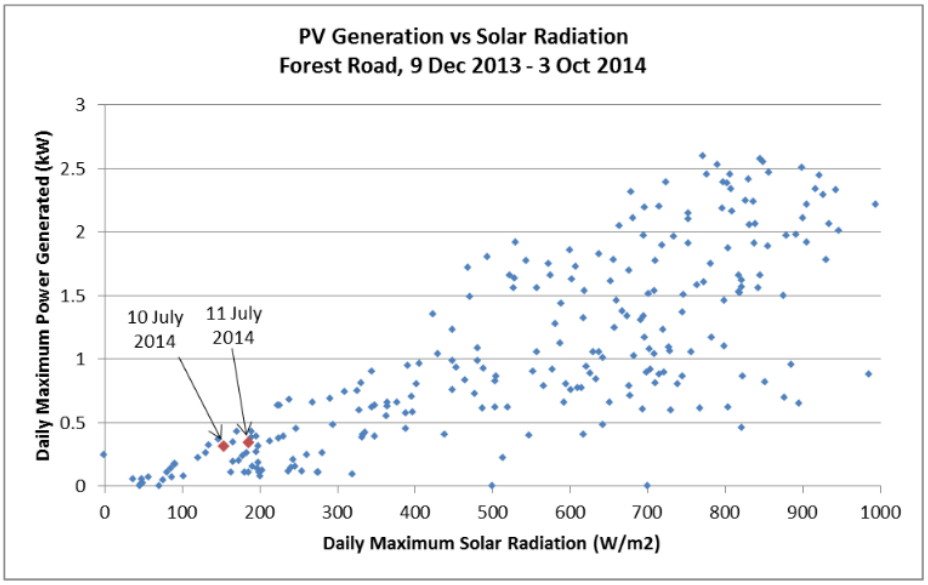
UK Power Networks' business plan predicts a significant and continuous increase in PV installation connected to its network by 2023. According to this calculation 1 in 10 customers will have a PV installation in their neighbour which may have an influence on a voltage rise. The data were collected in years 2013-2014. Measurements were performed in 20 distribution substations (18 including high PV penetration and 2 baseline substations with no PV penetration) and 10 customers PV installations each connected to the one of the 20 trial substations. Due to low respond rate and only 70 % of houses being unsuitable it was possible to install a monitoring equipment only in six houses. In remining four cases the monitoring equipment was installed on the overhead network just outside the customers' premises. The weather data from 5 locations were also collected. Monitoring sites were selected based on the number of PV connected. Less than 1% of EPN/SPN (East Power Networks/South Power Network) substations have more than 50kW PV attached. The six endpoint sites were among the highest PV penetrated locations in studied area but did not caused a voltage rise above the statutory limits. There was observed a measurable voltage rise from PV generation along LV feeders but its impact on endpoint voltages was insignificant since it depends mostly on voltage regulations on HV network and distribution transformer. Harmonics or phase voltage imbalance on LV feeders were not impacted by the PV generation. However in some rural areas only PV customers are at the end of long single-phase feeder and in such cases the voltage rise can be up to six times higher than expected for an equivalent three-phase generator.

From the collected data Forest Road site was chosen for the model because both grid and weather data were available for this location.

According to "Validation of Photovoltaic (PV) Connection Assessment Tool" report PV output depends mostly on time of the year, time of the day, solar radiation and panel orientation which may not create diversity in the output of PV clusters if they include undersized inverters. As it is supposed, the highest PV output peak appeared during summer noon and in general peaks were higher and longer in summer than in winter. As it was observed during the mentioned studies the PV generators are more likely to achieve higher peak output during cloudy days than on sunny days. This is related to the temperature's

effect on efficiency. During cloudy days panels can cool down and later when the sun appear work with full efficiency before they warm up again.

On the picture below (from “Validation of Photovoltaic (PV) Connection Assessment Tool”) it is shown daily maximum power generated vs. daily maximum solar radiation for Forest Road location. It looks like there is some correlation in this data but it’s is not very high.



The question is if there are any other significant features that may have an impact on PV power generation beside those listed in “Validation of Photovoltaic (PV) Connection Assessment Tool” as the picture above would suggest.