Predicting Solar Irradiance and Inverter power for PV sites

Area of Expertise:	Data Analytics
Industry:	Solar
Name of Client:	InnovateUK collaborative project
Project End Date:	January 2019
Project Timeline:	1 Year

1. Optimisation Scope

The growing demand for solar PV is having an impact on the low voltage (LV) network. New grid connections can be expensive and can affect a project's viability. In some areas of the country Distribution Network Operators (DNOs) are forced to limit new solar sites. However, new technologies are introducing ways to make smarter use of the abundant free energy provided by the sun and deliver new revenue streams, without the need for costly infrastructure upgrades.

Accurately modelling the commercial benefits of solar PV and battery storage is an important aspect of the project. If predicted solar generation is higher than the export limit of the site, a battery can be charged instead of curtailing generation, discharged to grid during a later period of high demand, and in the meantime the battery can be employed for Demand Side Response (DSR). For a site with no installed storage, generation can be curtailed at times when the network is constrained in response to DSR signals, such as Demand Turn-Up. Accurate predictions allow the DNO or Transmission System Operator (National Grid) to efficiently manage their network

With the UK's solar capacity forecast to rise to 15.7GW by 2020 – from just over 9.3GW at present – using advanced technology to more efficiently integrate and optimise solar PV sites is vital to create a more sustainable energy future. Due for completion in early 2019, this project aims to pave the way for the smarter use of solar PV via peer-to-peer energy markets that benefit local communities, delivering a smarter, more flexible energy system across the UK

Funded by InnovateUK, the UK Government's Innovation Agency, this collaborative research project helps solve this problem by forecasting solar output in near-time with better accuracy. This will allow PV sites to match generation to demand.

2. Project Team and their roles

Project Team comprises four companies:

Meniscus Systems Ltd: Lead Partner providing the data analytics and processing capability to deliver solar intensity predictions. All predictive analytics are delivered using the Meniscus Analytics Platform (MAP).

Open Energi: Providing expertise to deliver accurate, real-time PV-based DSR solutions to DNOs and owner/operators of solar farms to more efficiently manage local networks.

BRE National Solar Centre: responsible for ensuring the system meets the requirements of the PV industry and providing domain expertise and access/advice on technical solar issues.

Cornwall Council: owner/operator of one of the solar farms used to test and demonstrate the system

3. Locations of PV sites used during the project

3.1. Kernow PV site

Cornwall Council provided access to their PV site located at Newquay Airport and operated by Solar Century. This is a 4MW site and was the main location for testing and developing the models. Data provided from the site included both invertor power and pyranometer irradiance at 15-minute periodicity for the period from the 1st January 2018 to the 31st October 2018.

3.2. Frome PV site

Open Energi already monitored, and implemented DSR, at a smaller PV site at Frome operated by Hermitage Solar Farm Limited. It was decided that this would be a better location to run the NRT monitoring as it was cheaper and easier to install a new pyranometer at the site rather than install additional power monitoring equipment at the Newquay site. As such, a RaZON ALL-IN-ONE Solar Monitoring System pyranometer was installed and started collecting data from the end of November 2018.

4. Technical Overview

4.1. Prediction of solar intensity values

Near real time (NRT) satellite imagery from the SEVIRI instrument on the Meteosat Second Generation 0-degree satellite is pre-processed by a third party who deliver a set of $1km^2$ gridded images of the solar reflectance with values from 0 (no cloud) to 254 (100% cloud). The set of images is a cut-out of the South West of England ($112,000km^2$) and includes visible and infra-red images as well as metadata on the estimated cloud height and the likelihood that the image is affected by reflection (see section 5.1). The NRT images are processed every 15 minutes and imported into the Meniscus Analytics Platform (MAP) for subsequent analytics. Historic imagery has been processed for the period from the 1st January 2018 to the 31st October 2018 and this data is used to develop and validate the underlying models.

4.1.1. Cloud cover predictions - Meniscus

The NRT imagery is used to create predictions of the solar reflectance for the next 2 hours at 15-minute increments. These predictions are delivered in under 5 minutes from the time the new NRT imagery is imported into MAP. This involves several key steps.

a) Meniscus uses a Convolutional Long Short-term Memory (LSTM) recurrent neural network (NN) as the basis of the predictive model. This is run using the open source Tensorflow software library. The model is trained using the historical image data from the period and includes over 12,000 cloud images. The output from the LSTM NN is a set of predicted images at 15-minute intervals for the next 2 hours.
b) Meniscus has also implemented a block matching and relaxation approach[1] as a supplement to the LSTM NN model. Reasons for using this approach are detailed in section 5.2. This method creates a speed and direction vector for every 16km2 cell of the images and these vectors are then applied to the last raw image to create a set of predicted images at 15-minute intervals for the next 2 hours.

A PV site ‘template’ has been created in MAP. This includes all the core calculations and site properties used to convert the solar reflectance images into cloud cover and then into irradiance data. It also includes several metrics to allow easy comparison between predicted and actual irradiance data. This ‘template’ is then applied to any new location, making it very quick to add new sites into MAP.

The cloud cover relevant to the location of interest is stored as a calculated Item in MAP and uses a specific algorithm to convert the solar reflectance data to cloud cover. This is detailed in section

[1] Technical assistance provided from Professor Adrian Evans at the University of Bath

5.3. This calculation considers both the time of day and date (to calculate the solar azimuth and elevation) and the estimated height of the clouds to calculate the cloud cover at the relevant cell that will impact the solar irradiance value at the specific location.

4.1.2. *Solar intensity predictions -- Meniscus*

The solar intensity predictions are based on a Clear Sky solar irradiance model. The Clear Sky values are then modified to account for the cloud cover at the specific location.

To avoid setting up the core calculations for this MAP calls the PVLIB Python[2] open source tool for the Clear Sky results (this includes the direct {poa_direct}, diffuse {poa_diffuse} and total {sum of direct and diffuse} in-plane solar {poa_global}) intensity values. PVLIB provides a range of models and variables to help improve the model accuracy for the specific location.

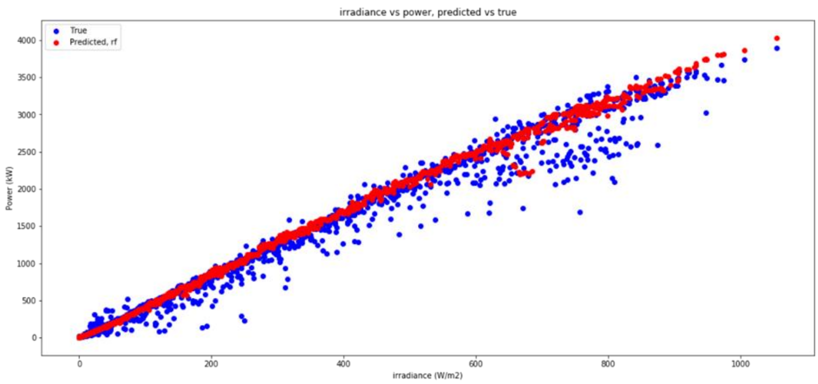
The final output is a calculated metric for the total in-plane cloud solar irradiance at the location. This is calculated every 15 minutes and is predicted for the next 2 hours. This metric considers the actual rainfall data at the location (available from radar rainfall data) as this has a very significant effect on the solar irradiance values – detailed more in section 5.4.

MAP also retains the predicted solar intensity values made at 30 minutes and 1-hour time intervals to allow for easy comparison against the actual solar intensity values.

4.2. Prediction of PV site power data and modelling of battery optimisation – Open Energi

For both the Kernow and the Frome sites the relationship between actual solar irradiance and actual inverter output power is very linear. A Random Forest Regression was used to derive an estimate of the inverter output power using 3 variables: In-plane solar irradiance, temperature and the invertor output power 30 minutes previously. Comparing the actual inverter output power to the estimated inverter power returns a regression coefficient (r^2) of +0.98. A multiple regression works nearly as well as the Random Forest Regression and is easier to apply and this is the model that is used in MAP to calculate the predicted inverter power output.

Figure 1- Irradiance vs power data



5. Lessons Learnt

Any innovative project encounters challenges during its development and this section details the main issues faced during the project, the ways which they were overcome, and the lessons learnt.

5.1. Reflection in imagery

The greatest technical issue in the project related to imagery issues, mainly 'sun glint' and snow and low-lying mist being interpreted as cloud.

Glare in the very early morning and late afternoon generates very high reflectance values or 'sun glint' due to the shallow angle between the sun and the satellite. This results in the software interpreting this as 100% cloud cover.

Snow and low-lying mist are also interpreted as clouds.

The image processing methodology has been continually updated to eliminate the impact of these effects. The metadata associated to each image now includes a value to identify sun-glint, but this can lead to

[2] https://pvlib-python.readthedocs.io/en/latest/

false negative results where the software ignores high reflectance values, due to cloud, and assumes there is no cloud. As a result, the earlier iterations of the models overestimated solar intensity values

5.2. Use of Block Matching

The original proposal for the project included the use of the block matching and relaxation technique to deliver the cloud reflectance predictions. The use of the LSTM NN model delivered improved operational performance and results, but the block matching technique is still used for processing images early in the morning. The LSTM NN model requires 4 images before it can deliver any predictions and so the block matching approach is used instead of the LSTM NN approach during this period.

MAP fuses the block matching and LSTM NN datasets together to deliver a seamless set of predictions from the time the first image arrives till the last.

5.3. Reflectance to Cloud Cover

The solar reflectance data from the satellite imagery is graduated in values from 0 to 254 with 0 being no reflectance and 254 being 100% reflectance.

Initial versions of the algorithm used to convert reflectance to cloud cover used a linear relationship to scale the reflectance so cloud cover (0-100%). This underestimated cloud cover leading to an overestimate of the solar intensity.

Further analysis identified that the relationship was non-linear and changed throughout the year. A monthly lookup is now used that applies a different non-linear relationship to the reflectance data for each month of the year. This increased the cloud cover leading to a reduction in the solar intensity values and a much-improved matching of actual against predicted values.

5.4. Use of rainfall data

During the analysis of the reflectance to cloud cover data the localised rainfall was tested as a parameter in the analysis. Although not used in the final algorithm for the reflectance to cloud cover calculation it was noted that rainfall data matched unexplained reductions in the actual solar intensity data.

Therefore, the rainfall data for the location, taken from radar rainfall data updated every 5 minutes, is used within the overall calculation of the solar intensity values. It is felt that the rainfall data may act as a surrogate for the depth of colour of the clouds. I.e. heavy rain relates to very dark overcast clouds. This data is not apparent when using satellite reflectance values.

The project is continuing to investigate two further areas of analysis to improve the relationship further.

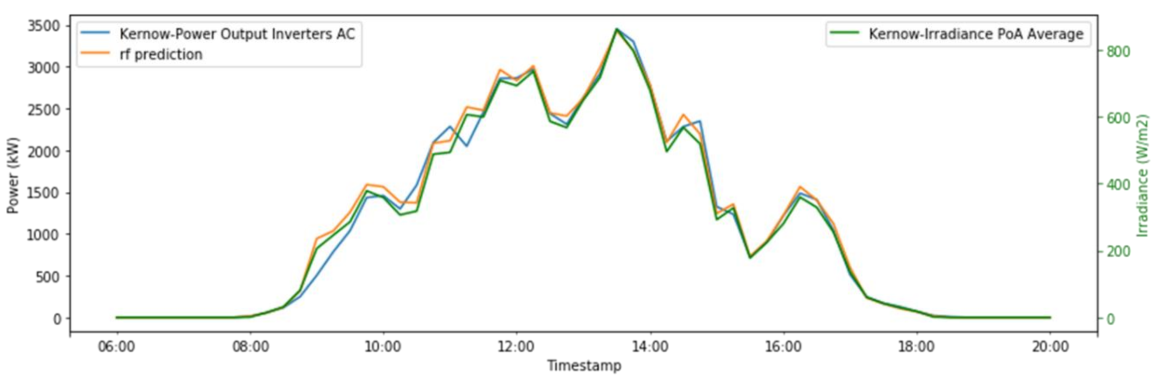
- The use of the infra-red satellite imagery to provide more information on cloud temperatures
- The use of rainfall data but aggregated over a longer time range

6. Benefits realised

6.1. Delivering short term irradiance predictions and site power

- Getting very good relationship between solar irradiance data and site power

Figure 2- Kernow actual vs predicted power



- with a 98% linear regression correlation coefficient. Some noise is present.

6.2. Additional revenue from DSR

Revenue from energy arbitrage is highly dependent on the following:

- Understanding specifics of commercial power purchase agreements on site between different parties
- Forecasting solar generation so charge battery from solar rather than imports from grid; pre-programmed profile to ensure battery has enough energy stored to discharge over peak, rather than trickle charging
  - This is where the forecast will help; days when we don't make as much revenue are due to poor solar generation compared to forecast
  - If solar generation is low, it is cheaper to charge up overnight to ensure reach full charge for 4pm
- Forecasting site demand accurately so as to not push site into export
- Electricity is most expensive between 4-7pm. British Summer Time (BST) months have solar overlap with peak so it may not be best idea to discharge battery as this can to push site into export which can lose money so need to change tactic.

6.3. Commercial analysis – size of PV site for break even

Open Energi carried out a range of scenarios that applied a 'standard' Power Purchase Agreement (PPA) agreement to a range of types and size PV sites and associated battery to quantify the revenue opportunities that a short-term solar forecast can deliver.

Types of PV sites ranged from sites with on-site demand from a battery or internal consumption (behind the meter sites) and those with no on-site demand (In front of meter) and sizes from 500kW up to 5 MW.

This analysis identified that:

1. DSR benefit is driven by shoulder months when solar generation overlaps with peak demand.
2. Enhanced forecasting results in the following uplift in energy markets & embedded benefits
   a. Behind the meter:2-10% uplift
   b. In front of the meter:1-5% uplift
3. Breakeven analysis
   a. Behind the meter: Business case is better so modelling indicates sites **2MW** and above would benefit, dependent on-site demand
   b. In front of the meter: Modelling indicates **>2MW** sites would benefit, depends on ratio of battery to solar inverter rating
4. Enhanced forecasting could lead to additional markets opening up, e.g. Balancing Mechanism access & DNO services, which aren't modelled.

7. Exploitation

7.1. Solar Irradiance - Prediction as a Service

Meniscus will deliver a web-based service to deliver predicted solar intensity values for any location. Initially the service will focus on using the area covered by the existing satellite image cut-out and this will be extended as new customers are found. An extension of the area requires re-training of the LSTM NN model and this takes both time and resource.

The service will be priced in a similar way to the company's existing historic and forecast rainfall location analytics with data available via both a dashboard and an API call.