Sheffield Solar

Prediction of Capacity Factor for Potential Solar PV System in Fakenham, UK

A case study using third order response surface profiling of real-world PV generation data to predict potential generation per kilowatt-peak per annum.

Jamie Taylor
2014-04-11

Executive Summary

A prediction of the annual generation per kilowatt-peak for a solar PV system located in Fakenham* has been carried out, focusing on a small range of azimuth and elevation angles corresponding to the proposed project specification. The resulting predictions are:

Predictions (kWh/kWp/year)		Elevation								
		20	22.5	25	27.5	30	32.5	35	37.5	40
Azimuth	165	864	866	865	861	854	847	839	832	826
	167.5	867	869	868	864	858	850	842	834	828
	170	869	872	871	867	860	853	844	836	830
	172.5	871	874	873	869	863	855	846	838	831
	175	872	876	875	871	865	857	848	840	833
	177.5	872	877	877	873	867	859	850	841	834
	180	872	877	878	874	868	860	851	842	834
	182.5	872	877	878	875	869	861	852	843	835
	185	870	877	878	875	869	862	853	844	835
	187.5	869	875	877	875	869	862	853	844	835
	190	866	873	876	874	869	861	853	843	835
	192.5	863	871	874	873	868	861	852	843	834
	195	859	868	871	871	866	860	851	842	833

Figure 1; Table of predicted generation per kWp for various azimuth and elevation angles

Introduction

The Microgen Database (MgDb) collects and analyses solar photovoltaic (PV) generation data from over 6000 installations across the UK. Here that data is used to predict the potential annual capacity factor (ACF) for PV systems located in the town of Fakenham (Norfolk) and the surrounding areas for installations consisting of a single array having an orientation between 165° and 195° clockwise from North† and an elevation of between 20° and 40° from horizontal. No account is taken for shading, although it would be feasible to multiply the predicted ACF by a shading factor such as that described in the *Guide to the Installation of Photovoltaic Systems*, published by MCS. Since predictions presented in this report are in the form of generation per kWp, the true prediction of annual generation can be calculate by multiplying by the kWp‡ of the proposed system.

Method

The MgDb ensemble is restricted to the 275 PV installations situated in the region indicated in Figure 2. This region was chosen after careful consideration of the implications of choosing a region that should both maximise sample size and minimise affects which cannot be accounted for in this method, in particular the effect of coastal microclimatic effects as discussed in the *National Meteorological Library and Archive Fact sheet 14* from the Met Office. More standardised regions that were considered can be seen in Figure 7 of the appendix. These regions produced poor fits when profiled and resulted in visibly anomalous predictions which is suspected to be a result of coastal installations being under represented in the sample. The sample has also been restricted to systems with an azimuth between 90° and 270° and an elevation between 20° and 50°, for whom the Microgen Database holds at least one full calendar years' worth of data.

* Latitude 52.829°, longitude 0.852°
† i.e. ±15° from due South
‡ Also referred to as "system size" or "system capacity"

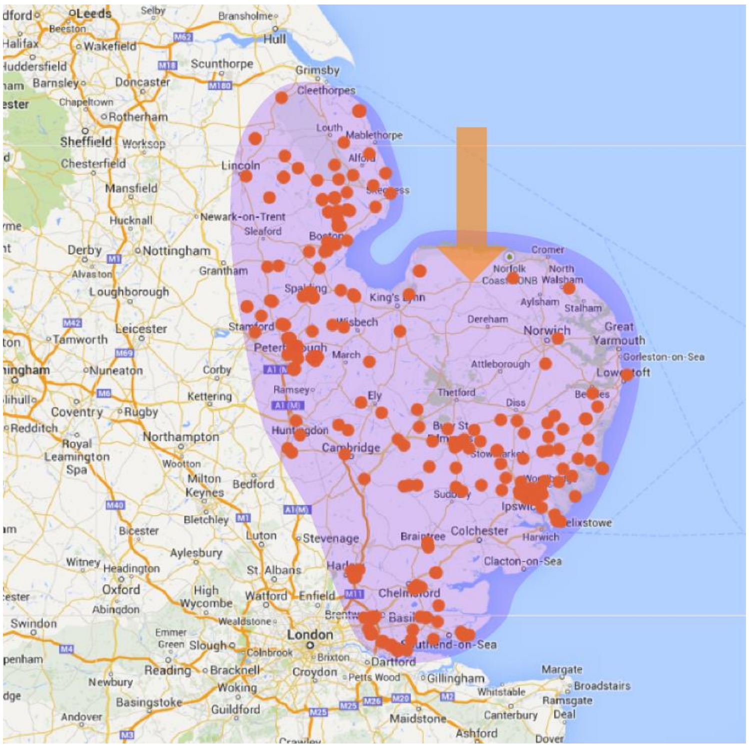


Figure 2; Map of PV installations whose generation data is used in this analysis. Orange arrow marks location of Fakenham.

The generation data for each installation has been used to calculate the annual generation normalised to the system rated power (a.k.a. "generation per kWp" or "capacity factor") by calendar year. The extent of data available for each system varies from 1 year to 5 years, so the capacity factor is averaged over all available years for each installation in order to minimise the effect of varying irradiance levels between years.

A third order response surface analysis is used to profile the capacity factor for this region as a function of latitude, longitude, azimuth and elevation. In order to optimise the profiler, an initial profile was produced which was used to identify outliers in the dataset for whom the response surface was not a good fit. Outliers are defined here as systems where the residual fell outside the following bounds:

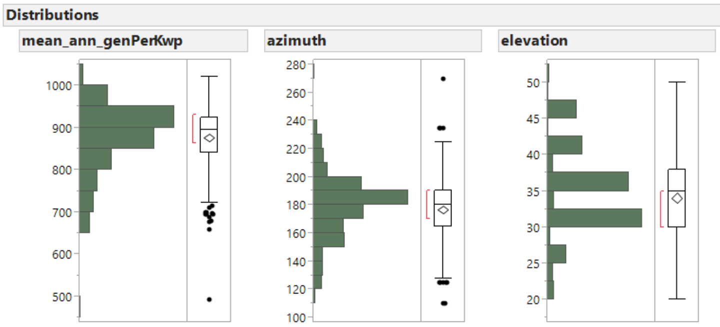
$$Q_1 - 1.5 \times IQR < Residual < Q_3 + 1.5 \times IQR$$

Where Q_1, Q_3 are the first and third quartiles respectively and IQR is the inter-quartile-range.

Since the response surface method gives equal weight to all responses in the sample, it's reasonable to assume that these outliers fit badly due to some unknown variable (e.g. shading) and/or errors in the MgDb data. These outliers were removed and the response surface regenerated, see Figure 9.

Analysis/Results

It can be seen from Figure 2 that the geographical distribution of installations used in this analysis is not ideal in that the majority of systems more than 20 miles away from the desired location of Fakenham (orange arrow). This undesirable geographical distribution is outweighed by the desire to maximise the sample size and distribution of elevations and azimuths (Figure 3) in order to improve the statistical accuracy of the response surface profiler. The primary geographical dependence across this region, namely the varying levels of global horizontal irradiance, ought to be accounted for through the inclusion of latitude and longitude terms in the profile.



Quantiles (mean_ann_genPerKwp)		
100.0%	maximum	1020.15
99.5%		1019.94
97.5%		986.01
90.0%		953.965
75.0%	quartile	924.078
50.0%	median	895.515
25.0%	quartile	840.85
10.0%		761.265
2.5%		694.797
0.5%		556.461
0.0%	minimum	493.702

Summary Statistics (mean_ann_genPerKwp)	
Mean	875.98611
Std Dev	76.086814
Std Err Mean	4.5882075
Upper 95% Mean	885.01873
Lower 95% Mean	866.95349
N	275

Quantiles (azimuth)		
100.0%	maximum	270
99.5%		256.7
97.5%		225
90.0%		200
75.0%	quartile	190
50.0%	median	180
25.0%	quartile	165
10.0%		150
2.5%		125
0.5%		110
0.0%	minimum	110

Summary Statistics (azimuth)	
Mean	177.09818
Std Dev	22.403919
Std Err Mean	1.3510071
Upper 95% Mean	179.75786
Lower 95% Mean	174.43851
N	275

Quantiles (elevation)		
100.0%	maximum	50
99.5%		50
97.5%		45
90.0%		45
75.0%	quartile	38
50.0%	median	35
25.0%	quartile	30
10.0%		25
2.5%		22.9
0.5%		20
0.0%	minimum	20

Summary Statistics (elevation)	
Mean	34.054545
Std Dev	6.1282467
Std Err Mean	0.3695472
Upper 95% Mean	34.782058
Lower 95% Mean	33.327033
N	275

Figure 3; Distribution of the real ACF, azimuth and elevation for the systems used in this profile

In Figure 3 we can also see the distribution of the real ACF for the 275 systems profiled. The mean ACF is 876.0 kWh/kWp/year with a standard deviation of 76.1 kWh/kWp/year. This provides a ball park figure for the predicted ACF in this region, but of course fails to account for the distribution of azimuths and elevations across the sample.

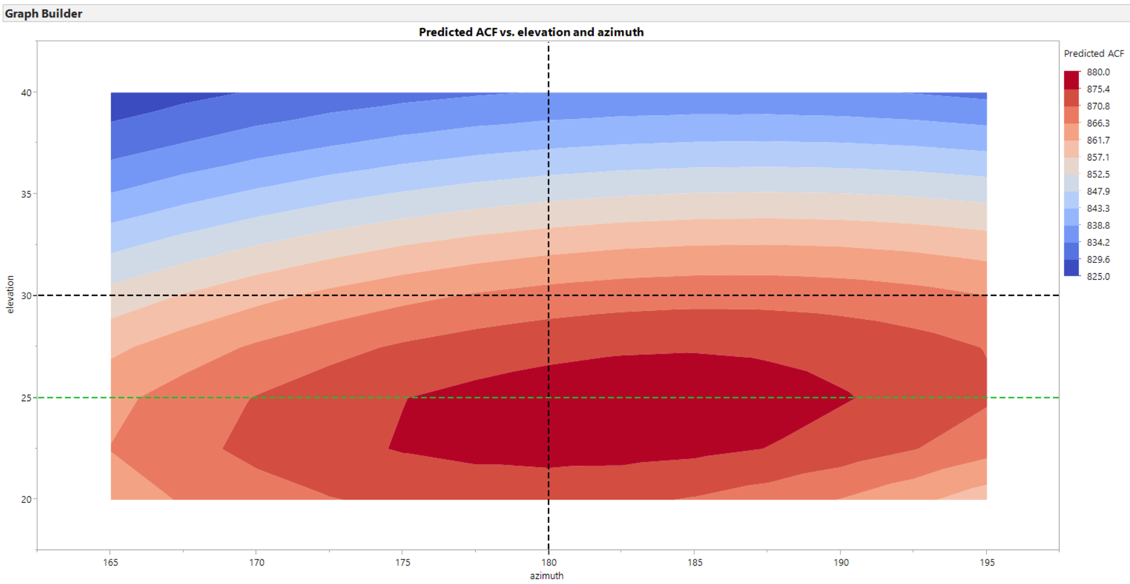


Figure 4; Contour plot of the Predicted ACF as a function of azimuth and elevation for the region analysed

Figure 4 shows the results of the profiler's prediction when applied to azimuths and elevations in the range of interest in intervals of 2.5°. The black dashed lines mark the reference for an azimuth of 180° (due South) and an elevation of 30°, which is generally considered the optimal configuration in the UK. In fact, this sample suggests that an elevation closer to 25°, marked with a green dashed line, achieves the optimal ACF, though the variation from an elevation of 25 to an elevation of 30 is very small relatively speaking. One can see the predicted ACF varies from 875.4-880 kWh/kWp/year for the "optimal" configuration, to 825-829.6 for configurations at the outer edges of the range of interest. The predictions can be seen in more detail in Figure 5.

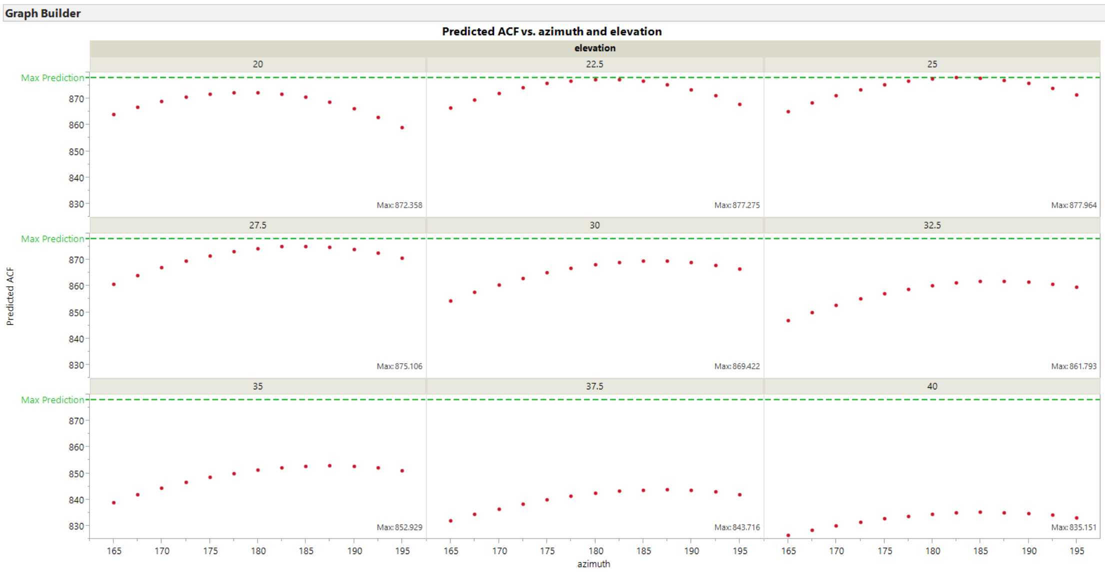


Figure 5; Scatter plot of predicted ACF as a function of azimuth and elevation for the analysed region

In Figure 5 we see the detailed variation of the predicted ACF with azimuth and elevation, with the green dashed line once again referencing the maximum prediction. As expected, the predicted ACF is nearly always symmetrical around an azimuth of 180°. Using the dashed line as a reference point, Figure 5 also shows the effect of elevation on the predicted ACF. Both Figure 4 and Figure 5 show that the dependence of the predicted ACF on the azimuth in this small range is weak, that is, slight deviations from due South result in relatively small decrease in the predicted ACF.

The predicted annual capacity factor for the "optimised" PV array (azimuth:180°, elevation:30°) is 869.0 kWh/kWp/year. The mean response of the profile was 876.0 kWh/kWp/year, as expected[§], with a root mean square error of 47.5, indicating an error of approximately ±5.4% in the predictions. The standard deviation of the residuals for the sample is 45.6 kWh/kWp/year, indicating that 68.2% of predictions will be subject to approximately ±5.2% error or less. Detailed analysis of the errors can be found in Figure 8 of the appendix.

Comparison

There are many prediction tools available to provide estimates of the ACF for a potential solar PV array. Until recently (7th May 2013), UK PV installers were required to present clients with SAP2009 predictions, based on an algorithm developed by the Building Research Establishment (BRE). This algorithm used irradiance measurements based in Sheffield, UK, assuming annual irradiance levels to be constant at all locations across the UK. It thus accounted poorly for the geography of a system, but is included here for reference. SAP2009

§ The response surface method using a least squares fit should always return a mean response approximately equal to the mean of the sample.

was superseded by the MCS *Guide to the Installation of Photovoltaic Systems*, which uses irradiance measurements specific to the regions shown in Figure 10 for the prediction calculation, and also improves upon the method for assessing shading (though no account is made for shading here). Another widely-used tool employed by many installers is the Photovoltaic Geographical Information System (PVGIS) developed by the European Commission's Institute of Energy and Transport (IET). PVGIS "Classic" makes use of ground based irradiance measurements to predict PV potential, whilst the more recent tool employs satellite irradiance data from the Satellite Application Facility on Climate Monitoring (CM SAF). Several other commercial tools are available, such as PV*SOL and PVSyst, none of which are considered here.

In Figure 6 the predictions made by SAP2009, MCS, PVGIS-Classic and PVGIS-CMSAF[**] are shown alongside the prediction made in this report for a system based in Fakenham with an azimuth of 180° and an elevation of 30°.

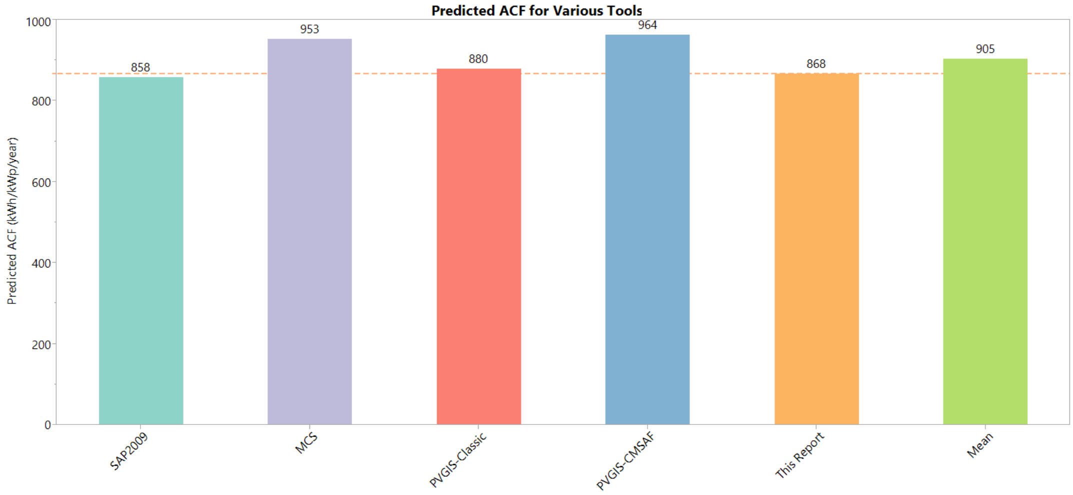


Figure 6; Plot of the predicted ACF for various tools alongside the prediction made in this report

The variation in the predictions shown in Figure 6 demonstrates the huge discrepancies that can occur when employing different calculation methods and irradiance datasets in the derivation of ACF. Once again it is important to note that no account is made for shading here in that the sample used for the response surface profile is assumed to have zero shading for all of the installations - as such in all of the above predictions shading is ignored[††] or the shading factor set to 1.

[**] PVGIS has additional inputs of 'PV technology' and 'Estimated system losses' which for the purposes of this comparison were set to the defaults of "Crystalline Silicon" and 14% respectively
[††] PVGIS claims to take into account horizon shading (i.e. due to landscape features), but does not account for local shading by trees, buildings etc

Appendix

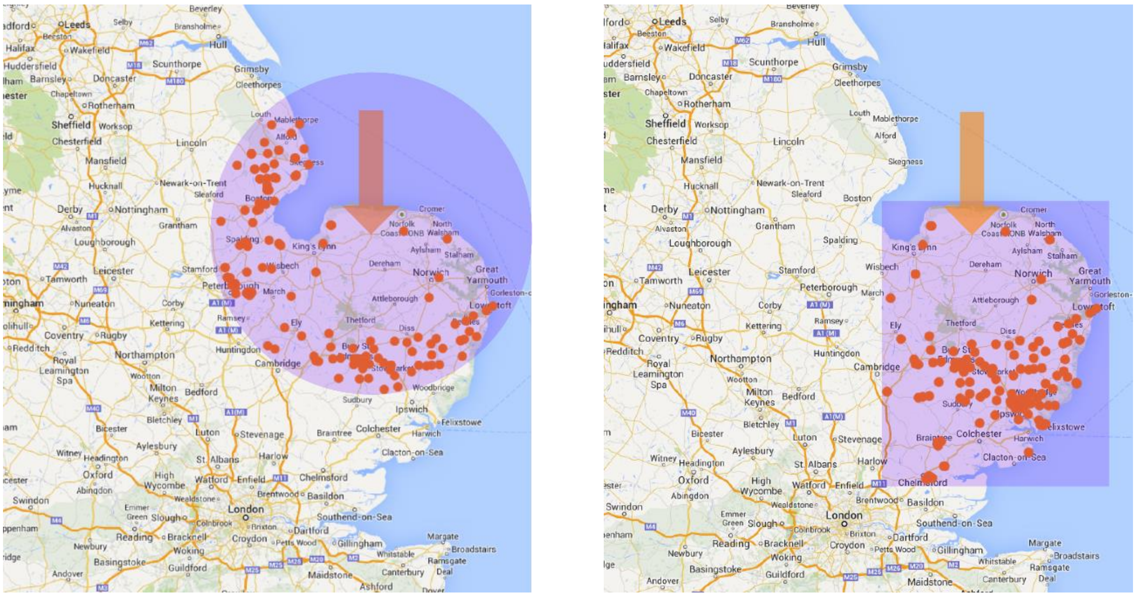
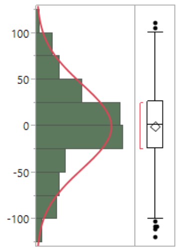


Figure 7; Two of the regions considered for the sample, both of which produced a poor fit and anomalous predictions

Distributions

Residual of Predictions



Normal(-2e-13,45.577)

Quantiles

100.0%	maximum	111.258
99.5%		108.817
97.5%		94.4494
90.0%		59.8854
75.0%	quartile	26.7594
50.0%	median	1.86679
25.0%	quartile	-23.89
10.0%		-65.807
2.5%		-91.496
0.5%		-115.63
0.0%	minimum	-118.77

Summary Statistics

Mean	-1.61e-13
Std Dev	45.576959
Std Err Mean	2.748394
Upper 95% Mean	5.4106523
Lower 95% Mean	-5.410652
N	275

Fitted Normal

Parameter Estimates

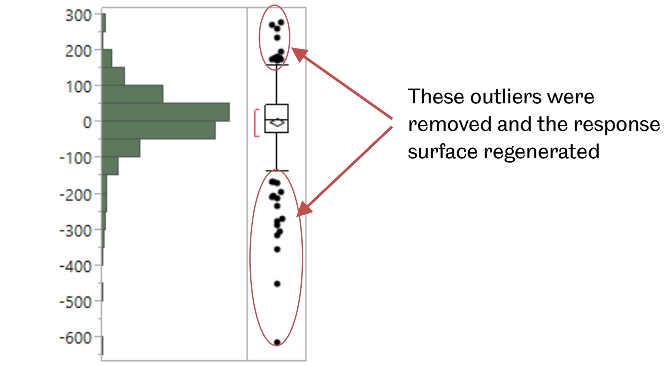
Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	-1.61e-13	-5.410652	5.4106523
Dispersion	σ	45.576959	42.059523	49.741392

-2log(Likelihood) = 2880.08745433777

Figure 8; Details of distribution of residuals for the prediction

Distributions

Residual of Predictions



Quantiles

100.0%	maximum	277.959
99.5%		273.079
97.5%		179.327
90.0%		98.062
75.0%	quartile	48.7593
50.0%	median	4.81001
25.0%	quartile	-29.724
10.0%		-95.938
2.5%		-267.37
0.5%		-514.41
0.0%	minimum	-612.7

Summary Statistics

Mean	6.695e-13
Std Dev	96.542433
Std Err Mean	5.3968861
Upper 95% Mean	10.617987
Lower 95% Mean	-10.61799
N	320

Figure 9; The distribution of residuals following the first attempt at response surface

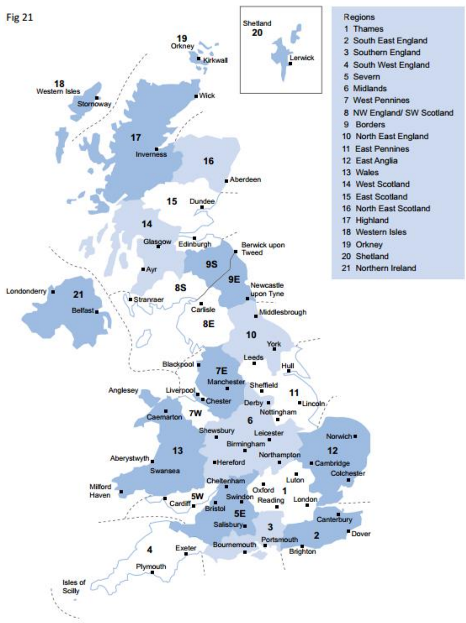


Figure 10; UK regions used by the MCS to determine annual solar irradiance[‡‡]

[‡‡] These regions were originally defined in the SAP2012 guidance which was not adopted for PV predictions.