Western Power Distribution

EFFS Change Request 01

Final Report

By:	Smarter Grid Solutions Ltd.
Date issued:	29/05/2020
Document number:	200824 06

1. DOCUMENT ISSUE CONTROL

Document name: **Final Report**

Document number: **200824 06**

Version	Issue Date	Author	Reviewed by	Approved by	Description	DUR #
A	**29/05/2020**	**Dimitris Athanasiadis**	**Mark Collins**	**Euan Davison**	**Initial Issue**	**4292**

Contact:

Name: Mark Collins

Job Title: Senior Smart Grid Analyst

Email: mcollins@smartergridsolutions.com

Smarter Grid Solutions Ltd

Optima Building

58 Robertson Street

Glasgow

G2 8DU

CONTENTS

1. DOCUMENT ISSUE CONTROL 2
2. INTRODUCTION 6
3. SITE OVERVIEW AND FEATURE DATA UPDATE 8
4. UPDATING PERFORMANCE METRICS 10
5. UPDATING THE XGBOOST (MACHINE LEARNING) MODELS 11
5.1. Common Test Bench 11
5.2. Bulk Supply Points (BSP) and Primary Substations 12
5.2.1. Base Model (BSP) 12
5.2.2. Update Model with Weather Data (BSP) 12
5.2.3. Summary of Newton Abbot 13
5.2.4. Summary of Sowton BSP 14
5.2.5. BSP/Primary Summary 15
5.3. Photovoltaic Generator Sites 16
5.3.1. Base Model (PV) 16
5.3.2. Update Model with Weather Data (PV) 17
5.4. Summary of PV Updates 18
5.5. Wind Generator Sites 19
5.5.1. Base Model (Wind) 20
5.5.2. Update Model with Weather Data (Wind) 20
5.6. Summary of PV Updates 21
5.7. Machine Learning Model Summary 23
6. CONSTRUCTING ENGINEERING MODELS 24
6.1. Engineering Model PV 24
6.2. PV Engineering Model Behaviour 25
6.3. PV Engineering Model Performance Summary 26
6.4. Engineering Model Wind 28
6.5. Wind Engineering Model Behaviour 30
6.6. Wind Engineering Model Performance Summary 31
6.7. Engineering Models Conclusions 34
7. MACHINE LEARNING MODEL VS ENGINEERING MODELS 34
7.1. Machine Learning Model vs Engineering Models Conclusions 36
8. CONCLUSIONS 37
9. DELIVERABLES 37
9.1. Machine Learning Models 37
9.2. Engineering Models 37

Tables

Table 1: Wind sites 8
Table 2: PV sites 8
Table 3: BSP/Primary Sites 8

Figures

Figure 1: Peak Day error 11
Figure 2: Newton Abbot BSP Current Model RMSE Analysis 12
Figure 3: Newton Abbot BSP TAWF RMSE Analysis 13
Figure 4: Newton Abbot BSP TAWF MAPE Analysis 13
Figure 5: Newton Abbot BSP TAWF Day Of Week (DOW) v no DOW RMSE Analysis 14
Figure 6: Newton Abbot BSP TAWF Day Of Week (DOW) v no DOW Peak to Peak RMSE Analysis 14
Figure 7: Sowton BSP Holiday V No Holiday RMSE Analysis 15
Figure 8: Sowton BSP Holiday V No Holiday Peak to Peak RMSE Analysis 15
Figure 9: RMSE Performance of BSP ML Models 16
Figure 10: Peak to Peak RMSE Performance of BSP ML Models 16
Figure 11: TO RMSE Analysis 17
Figure 12: TO vs TAWF RMSE Analysis 18
Figure 13: TO vs TAWF RMSE Peak To Peak Analysis 18
Figure 14: RMSE Performance of all PV ML Models 19
Figure 15: Peak to Peak RMSE Performance of all PV ML Models 19
Figure 16: MAPE Performance of all PV ML Models 19
Figure 17: Peak to Peak MAPE Performance of all PV ML Models 19
Figure 18: TO RMSE Analysis 20
Figure 19: TO vs TAWF RMSE Analysis 22
Figure 20: TO vs TAWF RMSE Peak To Peak Analysis 22
Figure 21: RMSE Performance of all Wind ML Models 22
Figure 22: Peak to Peak RMSE Performance of all PV ML Models 22
Figure 23: MAPE Performance of all PV ML Models 22
Figure 24: Peak to Peak MAPE Performance of all PV ML Models 23
Figure 25: PV Material Coefficients from reference study 25
Figure 26: MW Output Vs MW Prediction 26
Figure 27: MW Output Vs GHI Input Data 26
Figure 28: All PV Engineering Model RMSE Performance 27
Figure 29: All PV Engineering Model Peak to Peak RMSE Performance 27
Figure 30: All PV Engineering Model MAPE Performance 27
Figure 31: All PV Engineering Model MAPE Peak to Peak Performance 28
Figure 32: Wind Turbine Coeffcients 30
Figure 33: MW Output Vs MW Prediction 31
Figure 34: MW Output Vs GHI Input Data 31
Figure 35: All Wind Engineering Model RMSE Performance 32
Figure 36: All Wind Engineering Model RMSE Peak to Peak Performance 32
Figure 37: All Wind Engineering Model MAPE Performance 33
Figure 38: All Wind Engineering Model MAPE Peak to Peak Performance 33
Figure 39: Average RMSE between PV models, across months compared 35
Figure 40: Average Peak to Peak RMSE between PV models, across months compared 35
Figure 41: Average RMSE between Wind models, across months compared 35

Figure 42: Average Peak to Peak RMSE between Wind models, across months compared 36

2. INTRODUCTION

Smarter Grid Solutions (SGS) has previously completed work relating to the forecast methods and models as part of Western Power Distribution's Electricity Flexibility and Forecasting System (EFFS). This change request covers the proposed additional work to satisfy Western Power Distribution's (WPD) additional requirements.

At the end of the initial work, SGS concluded that:

- Incorporation of weather forecast data as a feature of the machine learning (ML) models could improve MW and MVAr forecast;
- Using weather forecast and applying them to engineering models (EM) may perform better than machine learning models under certain conditions.

For this reason, in this work, this report explores:

- The performance of the XGBoost model with new feature data, i.e. weather forecast data; and
- The performance of engineering models of WPD generator sites driven by weather forecast data.

For each of these model types, this report compares and contrasts model performance, presenting the outcomes to enable WPD to select, from the newly developed models, the best performing model for the wider EFFS project.

The XGBoost models, previously constructed in the earlier phases of EFFS, which currently includes historical weather data only, were updated to have two optional additional inputs:

- Historical weather data only – to be used for forecasts up to 1 week in advance where no useful weather forecast data is expected to be available; and
- Historical weather data and short term weather forecast data – the short-term forecast weather data to be used for forecasts up to and including a week in advance, where weather forecasts are expected to be available and more useful than seasonal average values.

New engineering models were constructed:

- These models use mathematical expressions how a generator behaves, expressing how external weather stimuli interact with the physical construction and limitations of PV and Wind sites to produce MW and MVar export.

The test approach uses similar error analysis as the previous report, however this is extended to analyse error of diurnal peaks in order to provide an improve understanding of procurement risk, where general error performance of a model may be relatively good but consistently misses peak identification, leaving any procurement exercise short or long. Enabling a more considered procurement of service to minimise these occurrences.

The structure of the report is as follows:

- Site Overview and Feature Data update;
- Updating Performance Metrics;
- Updating the XGBoost (Machine learning) Models;
- Testing the XGBoost (Machine learning) Models;
- Constructing the Engineering Models;

- Testing the Engineering Models;
- Comparing Machine Learning and Engineering Model Performance; and
- Conclusions.

3. SITE OVERVIEW AND FEATURE DATA UPDATE

As part of the change request, WPD identified the following sites to create ML models for wind, photovoltaic (PV) and Bulk Supply Point (BSP) and primaries: Table 1, Table 2 and Table 3 respectively:

Table 1: Wind sites

Wind	Weather Station (Location)		
WPD name	Long	Lat	Local Name
DARRACOTT MOOR WINDFARM	50.972178	-4.1186563	Great Torrington
BEARS DOWN WIND FARM	50.47061	-4.9501503	St Columb Major
GOONHILLY WIND FARM	50.041474	-5.2028339	Cross Lanes
ROCKHEAD WIND FARM	50.626964	-4.7127465	Camelford
CARLAND CROSS WIND FARM	50.352263	-5.0361588	Carland

Table 2: PV sites

PV	Weather Station (Location)		
WPD name	Long	Lat	Local Name
HATCHLANDS FARM 33kV SOLAR PARK	50.407229	-4.4508885	Pensipple
KNOCKWORTHY 33kV SOLAR PARK	50.984422	-4.1212784	Great Torrington
REXON CROSS FARM 33kV SOLAR PARK	50.68582	-4.2325046	Broadwoodwidger
WILLSLAND 33kV SOLAR PARK	50.815937	-4.1201772	Hatherleigh
HOPE 33kV SOLAR PARK	50.210221	-5.36875	Trevarnon

Table 3: BSP/Primary Sites

BSP/Primary	Weather Station (Location)		
WPD name	Long	Lat	Local Name
Exeter City BSP	50.725562	-3.5269108	Exeter
Newton Abbott BSP	50.530647	-3.6094889	Newton Abbot
Prince Rock	50.369696	-4.1177251	Prince Rock
Sowton	50.724301	-3.4507478	Sowton
Bridge Mills	50.855202	-3.3927625	Bridge Mills (Cullompton)

This data used in this work is updated under the following methodologies:

- Each model is provided with the data defined in the original report for each site. The data are collected and checked against obvious data quality errors.
- No new electrical or weather parameters are introduced;
- Some sites may share weather data where coincidence relationships hold based on a predefined distance;
- The test range use historical data from **2014-2016** only; all data provided are in this range. This enables the use of widely available public resources;
- Electrical parameters for sites and locations required are provided by WPD;
- The weather data are provided by the NASA Modern-Era Retrospective analysis for Research and Applications (MERRA) dataset[1] ;

[1] Rienecker MM, Suarez MJ, Gelaro R, Todling R, et al. (2011). MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. Journal of Climate, 24(14): 3624-3648. doi: 10.1175/JCLI-D-11-00015.1

- Solar irradiance data were taken from both the MERRA and the European Commission's Photovoltaic Geographical Information System[2] and from the Surface Solar Radiation Data Set - Heliosat (SARAH)[3] [4];
- Historical Weather Forecasts: Since historical weather forecast data is not available, historical weather data with noise added is used; this noise is reflective of the error present in the real world weather forecast data. The MetOffice has identified an accuracy range[5] for its weather forecasts; a random noise band of 10% is added to the historical weather data to create short term weather forecast data for dates in the past.

[2] https://re.jrc.ec.europa.eu/pvg_tools/en/tools.html#MR
[3] Müller, R., Pfeifroth, U., Träger-Chatterjee, C., Trentmann, J., Cremer, R. (2015). Digging the METEOSAT Treasure—3 Decades of Solar Surface Radiation. Remote Sensing 7, 8067–8101. doi: 10.3390/rs70608067
[4] SARAH dataset. doi: 10.5676/EUM_SAF_CM/SARAH/V001
[5] https://www.metoffice.gov.uk/about-us/what/accuracy-and-trust/how-accurate-are-our-public-forecasts

4. UPDATING PERFORMANCE METRICS

The current performance metrics that are considered are:

Mean squared error (MSE)

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2$$

Root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2}$$

Mean absolute error (MAE)

$$MAE = \frac{1}{n}\sum_{t=1}^{n} |e_t|$$

Mean absolute percentage error (MAPE)

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n} \left|\frac{e_t}{y_t}\right|$$

e=Predicted-Actual

This is applied to each data point to give an overall metric for the model's performance. Moreover, it can obfuscate performance around diurnal peak prediction, which will influence procurement.

Error calculation is adapted to consider the predicted peak and actual peak for a 24 hours period between midnight and midnight.

$$e_{PeakDay} = max[Predicted]^{Day} - max[Actual]^{Day}$$

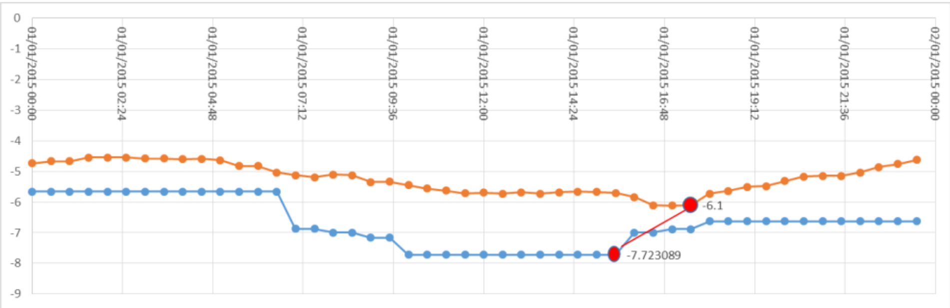

Figure 1: Peak Day error

$$Peak\ to\ Peak\ MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{e_{PeakDay}}{y_t}\right|$$

$$Peak\ to\ Peak\ RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_{PeakDayt}^{2}}$$

5. UPDATING THE XGBOOST (MACHINE LEARNING) MODELS

In the previous report, the XGBoost models used the following features to learn the underlying relationships between inputs to predict future outputs:

- 'date‘
- 'hour'
- 'quarter‘
- 'month’
- 'year‘
- 'dayofyear‘
- 'dayofmonth'
- 'weekofyear‘

These features are all historical temporal features. In this section, the featured data is expanded and further physical historical weather features are added. Furthermore forecasts of weather data are also added.

To aid integration of the models into wider EFFS programme, other considerations with regards to the models are presented:

- The developed models are constructed using the same data structure and scripting approach as in the models of the initial work ;
- The individual performance of the models against historic data are compared using similar metrics as in the original report, as well as a cross comparison between the new and previously developed models.

5.1. Common Test Bench

Each model is tested in the same manner. A year of historical data is used to train the model in order to predict the following week time horizon, for example producing a week ahead prediction for 01/01/2015-07/01/2015 a year of data between 01/01/2014-31/12/2014 is used to train the model. This common test bench is used as a control for this analyses to ensure performance can be suitably compared.

For all models the performance metrics extract are

- RMSE
- MAE
- Peak to Peak MAE
- Peak to Peak RMSE

RMSE provides a good general performance indication. MAPE provides a good observation of prediction underestimating (negative), and overestimating (positive), and for both peak to peak for diurnal performance.

5.2. Bulk Supply Points (BSP) and Primary Substations

The Bulk Supply Points and Primary substations previous models are analysed and the new models are cross compared.

5.2.1. Base Model (BSP)

The base model features for BSP/Primary sites are stated at the start of this section, these features are all temporal only (Temporal Only). An example of Newton Abbot BSP is presented in regards to performance.

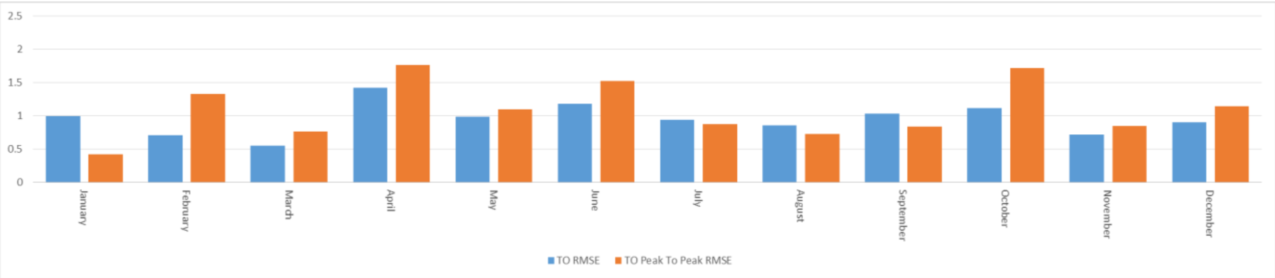

Figure 2: Newton Abbot BSP Base Model RMSE Analysis

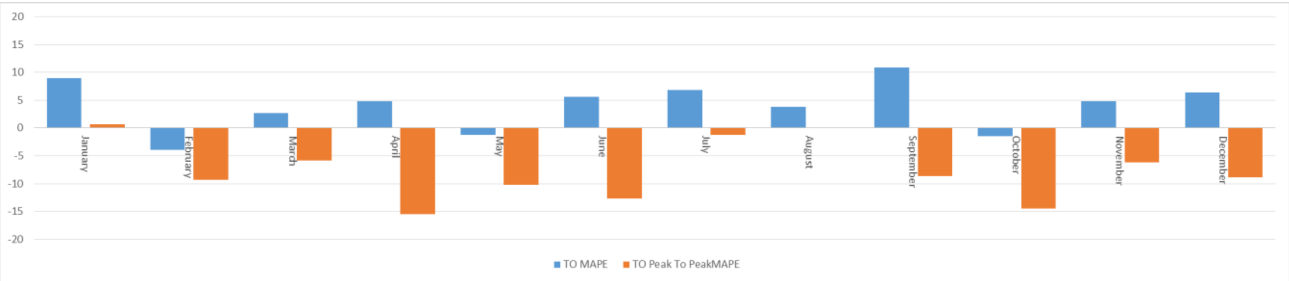

Figure 3: Newton Abbot BSP Base Model MAPE Analysis

5.2.2. Update Model with Weather Data (BSP)

Additional weather feature were added to the existing temporal features creating temporal and weather feature (TAWF) XG Boost Models

These features are:

Weather:

- GHI (MERRA)
- Wind Speed (MERRA)
- Pressure (MERRA)
- Temperature (MERRA)

Weather Forecast:

- GHI (MERRA)
- Wind Speed (MERRA)
- Pressure (MERRA)
- Temperature (MERRA)

This results in the following performance:

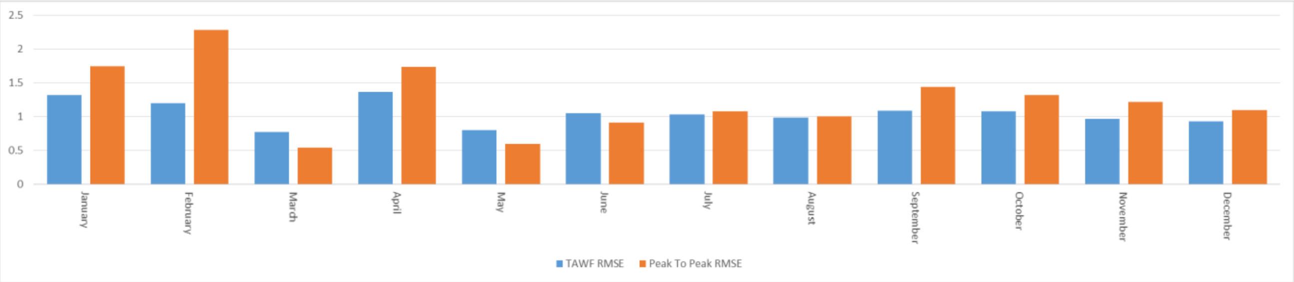

Figure 3: Newton Abbot BSP TAWF RMSE Analysis

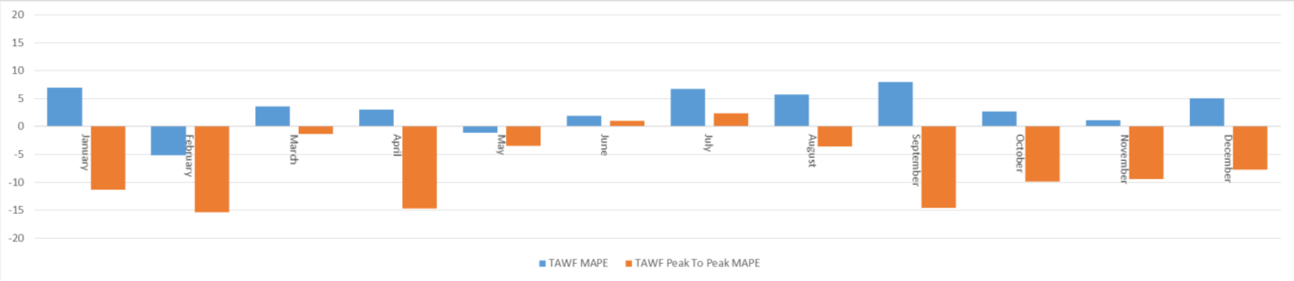

Figure 4: Newton Abbot BSP TAWF MAPE Analysis

5.2.3.Summary of Newton Abbot

Introduction of weather data and forecasts offers a slight improvement in some cases, indicating that weather features are not strongly linked, but do offer a minor relative performance improvements. Furthermore, weekend predictions poor compared to weekday. Therefore a day of the week feature was added to identify this behavioural impact.

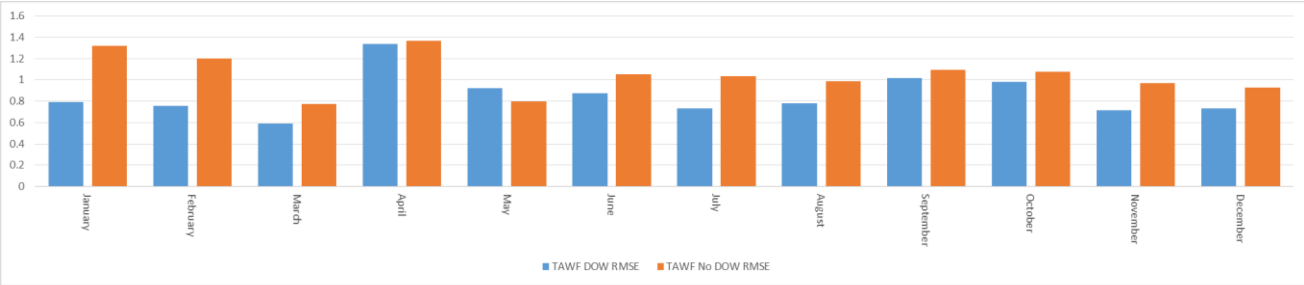

Figure 5: Newton Abbot BSP TAWF Day Of Week (DOW) v no DOW RMSE Analysis

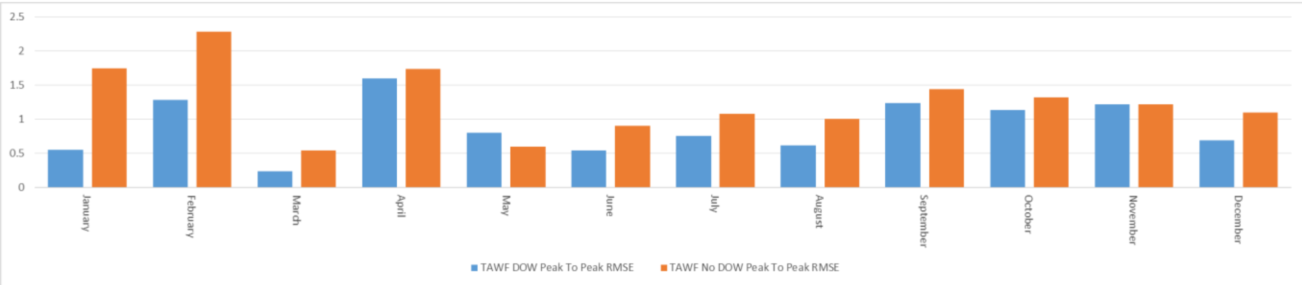

Figure 6: Newton Abbot BSP TAWF Day Of Week (DOW) v no DOW Peak to Peak RMSE Analysis

The improvement by introducing day of week was carried forward as a feature into all the models.

5.2.4. Summary of Sowton BSP

During the analysis of Sowton BSP, it was found further temporal feature additional could improve the performance of the BSP/Primary type ML models. Introduction of identifying specific days as public holidays improves performance of predictions, around holiday periods, Christmas (Jan/Feb), Easter (March/April).

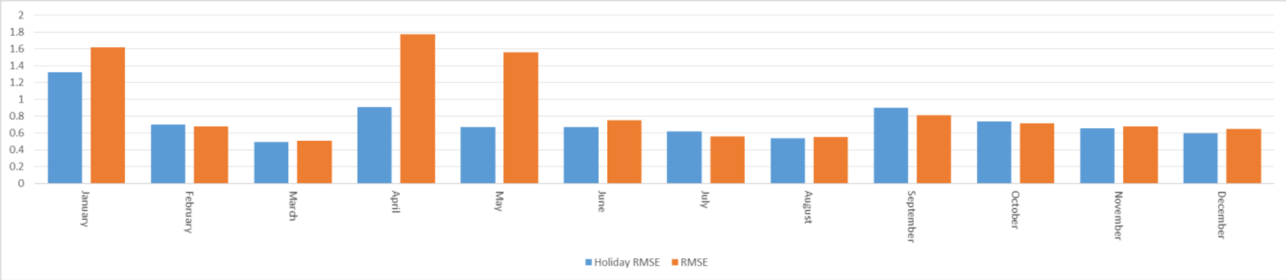

Figure 7: Sowton BSP Holiday V No Holiday RMSE Analysis

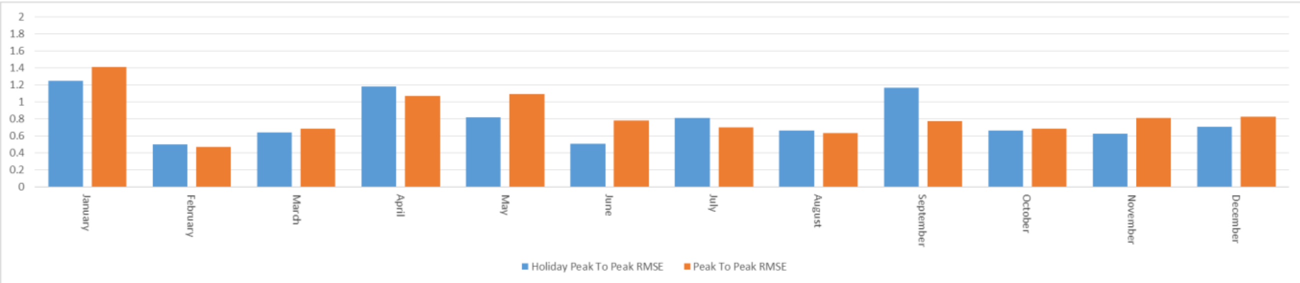

Figure 8: Sowton BSP Holiday V No Holiday Peak to Peak RMSE Analysis

5.2.5. BSP/Primary Summary

The best performance from the machine learning models were achieve with the follow temporal and weather features:

Temporal:

- Date
- Hour
- dayofweek
- quarter
- month
- year
- dayofyear
- dayofmonth
- weekofyear
- holiday

Weather:

- GHI (MERRA)
- Wind Speed (MERRA)
- Pressure (MERRA)
- Temperature (MERRA)

Weather Forecast:

- GHI (MERRA)

- Wind Speed (MERRA)
- Pressure (MERRA)
- Temperature (MERRA)

These features are incorporated in the model and carried forward into all further models.

All sites were analysed and summarised.

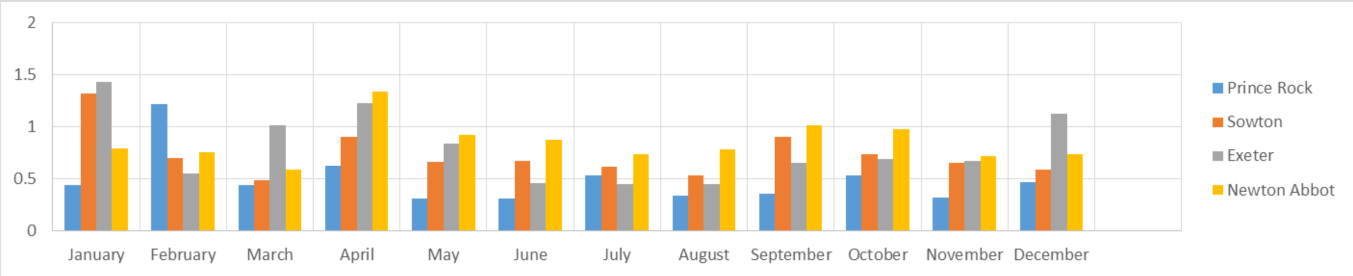

Figure 9: RMSE Performance of BSP ML Models

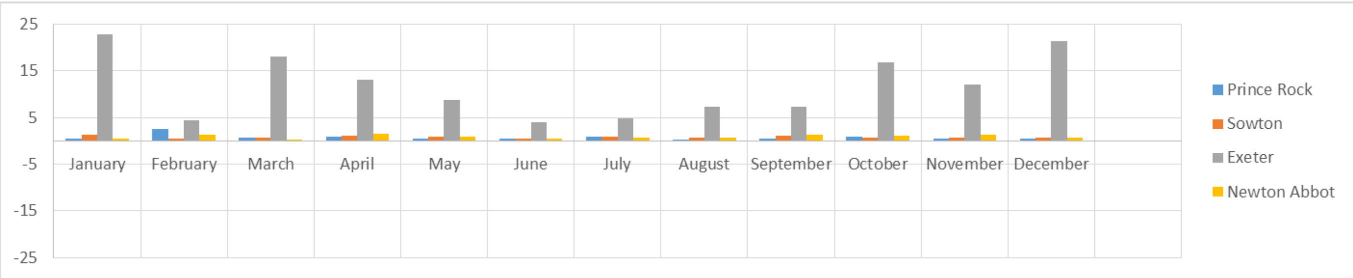

Figure 10: Peak to Peak RMSE Performance of BSP ML Models

Overall, the results can be summarised as follows:

- Weather features are not strongly linked;
- Strong temporal links to diurnal behaviour;
- Strong temporal links to holiday specific behaviour;
- The models will, in general, over estimate output but under estimate peaks; and
- Improved model performances in summer versus winter- showing load behaviour more variable in winter.

Suggestions for further improvements:

- Split models into weekday/weekend specific models;
- Split models into holiday specific models; and
- Split models into summer/winter specific models.

5.3. Photovoltaic Generator Sites

The PV generator site previous models are analysed, the base models are cross compared with the addition of weather data. Hatchlands PV is highlighted in these examples.

5.3.1. Base Model (PV)

The base model uses the following features:

Temporal:

- date
- hour
- quarter
- month
- year
- dayofyear
- dayofmonth
- weekofyear

Resulting in a temporal only (TO) feature machine learning model, whose performance is presented below.

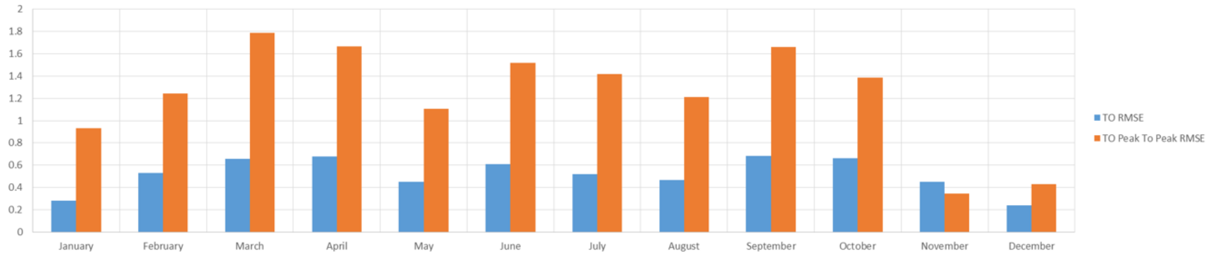

Figure 11: TO RMSE Analysis

5.3.2. Update Model with Weather Data (PV)

The additional weather forecast features were added:

Temporal:

- date
- hour
- quarter
- month
- year
- dayofyear
- dayofmonth
- weekofyear

Weather:

- GHI (MERRA)
- Wind Speed (MERRA)
- Pressure (MERRA)
- Temperature (MERRA)

Weather Forecast:

- GHI (MERRA)
- Wind Speed (MERRA)
- Pressure (MERRA)
- Temperature (MERRA)

5.4. Summary of PV Updates

A comparison of Hatchlands PV for the updated model (TAWF) with the base model (TO) is presented below:

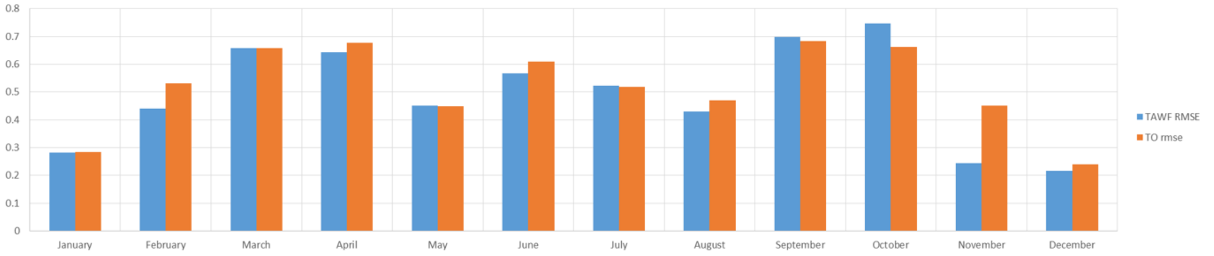

Figure 12: TO vs TAWF RMSE Analysis

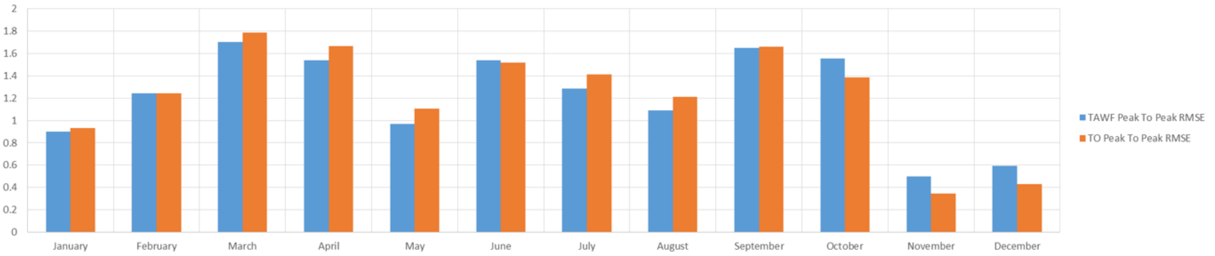

Figure 13: TO vs TAWF RMSE Peak To Peak Analysis

Introduction of Weather Data and Forecasts offers an improvement in some cases. Indicating that weather features are linked, but still a strong temporal trend due to daylight hours and seasonal changes.

Weather ML (TAWF) performance consistently better during summer for both RMSE and Peak RMSE, where predictions matter most.

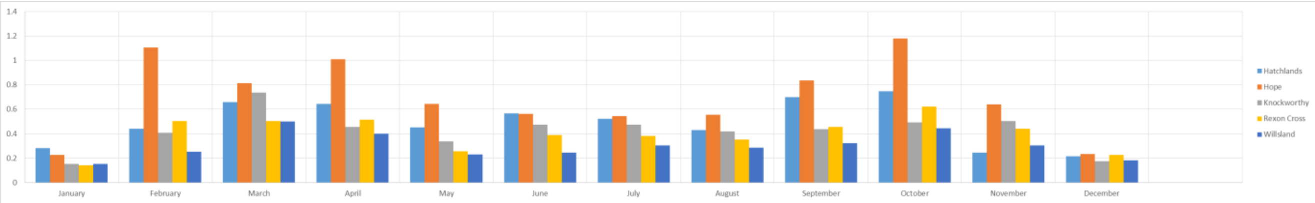

Figure 14: RMSE Performance of all PV ML Models

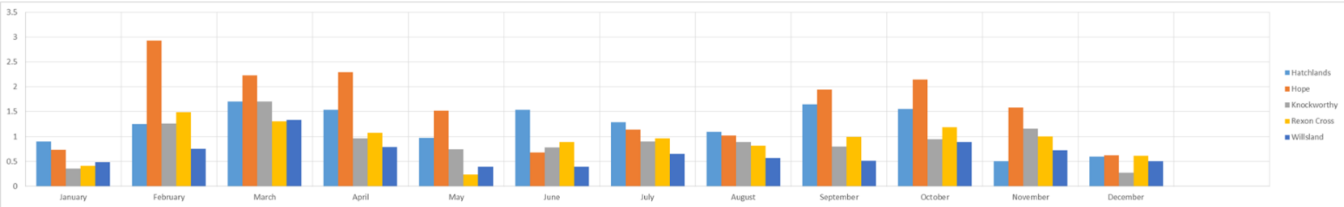

Figure 15: Peak to Peak RMSE Performance of all PV ML Models

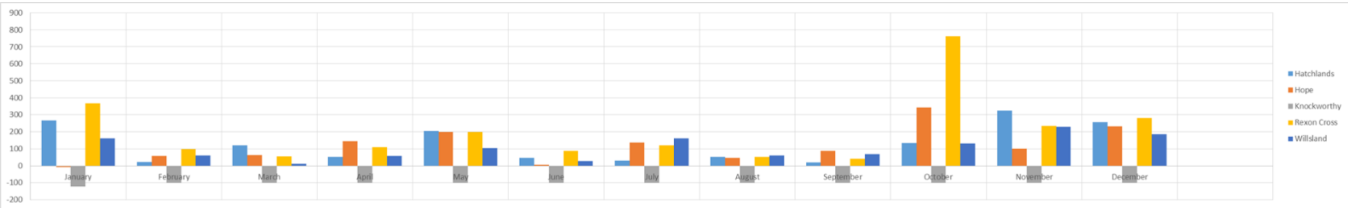

Figure 16: MAPE Performance of all PV ML Models

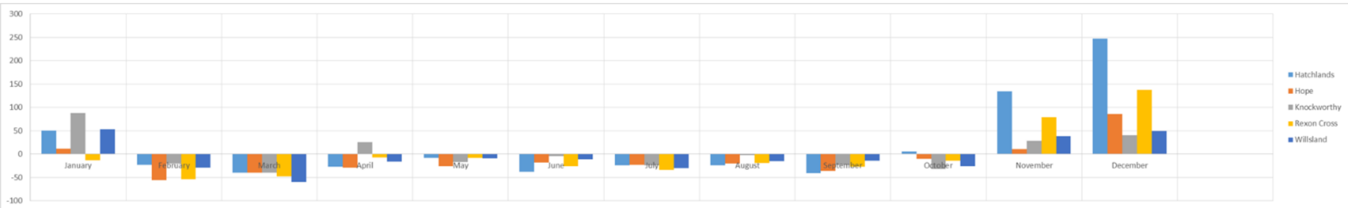

Figure 17: Peak to Peak MAPE Performance of all PV ML Models

Overall, this suggest that:

- Weather features are linked;
- Weather data improves peak prediction in summer, when it matters most;
- Strong temporal links to diurnal behaviour;
- Weak links to holiday/weekends (work concepts);
- Improved model performances in summer versus winter - showing site behaviour is more variable in winter.

Suggestions for further improvements:

- Split Models into Summer/Winter Specific Model –Daylight Range Configurable

5.5. Wind Generator Sites

The Wind generator site previous models are analysed, the base model are cross compared with the addition of weather data. Goonhilly Wind Site is highlighted in these examples.

5.5.1. Base Model (Wind)

The current model uses the following features:

Temporal:

- date
- hour
- quarter
- month
- year
- dayofyear
- dayofmonth
- weekofyear

Resulting in a temporal only (TO) feature machine learning model, whose performance is presented below:

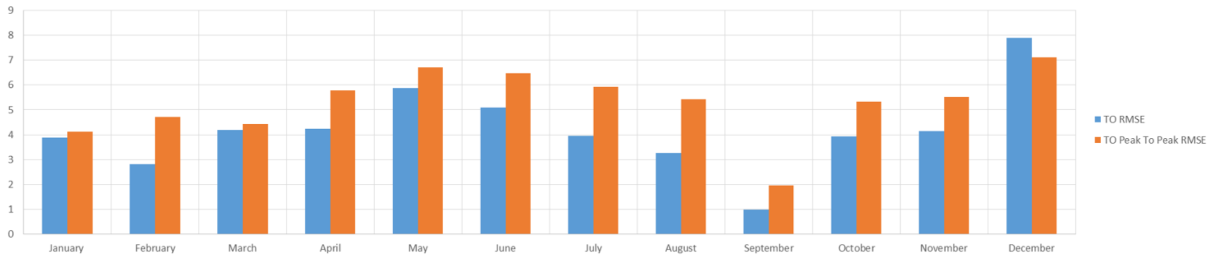

Figure 18: TO RMSE Analysis

5.5.2. Update Model with Weather Data (Wind)

The additional weather forecast features are added:

Temporal:

- date
- hour
- quarter
- month
- year
- dayofyear
- dayofmonth
- weekofyear

Weather

- GHI (MERRA)
- Wind Speed (MERRA)
- Pressure (MERRA)

- Temperature (MERRA)

Weather Forecast

- GHI (MERRA)
- Wind Speed (MERRA)
- Pressure (MERRA)
- Temperature (MERRA)

5.6. Summary of PV Updates

The comparison of Goonhilly Windfarm updates are presented showing the base model (TO) versus the model create using the additional weather data (TAWF):

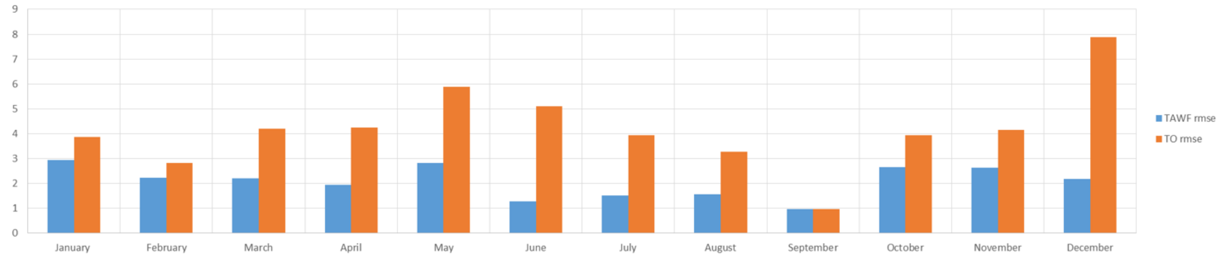

Figure 19: TO vs TAWF RMSE Analysis

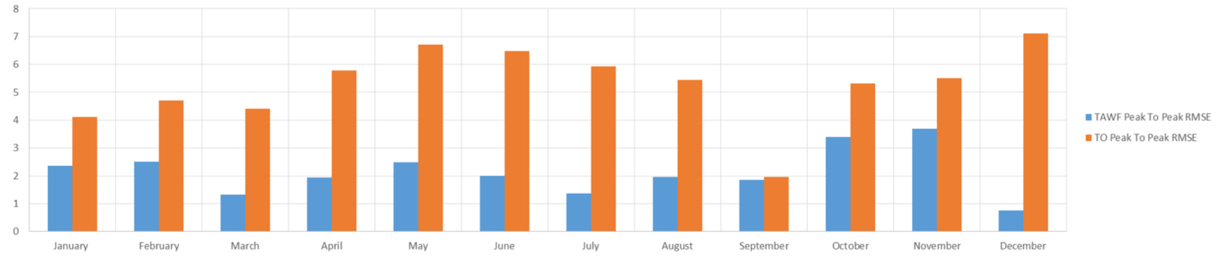

Figure 20: TO vs TAWF RMSE Peak To Peak Analysis

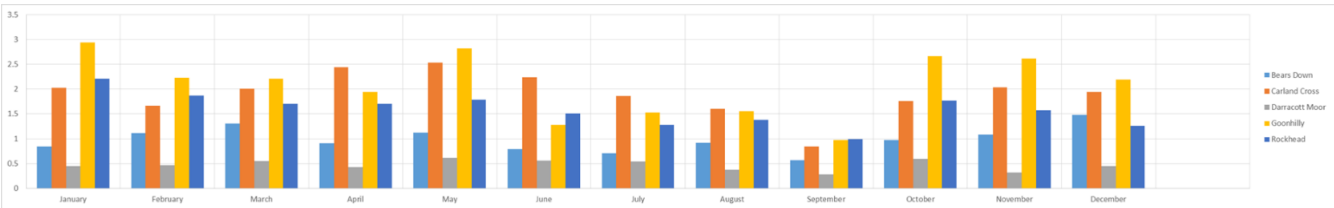

Figure 21: RMSE Performance of all Wind ML Models

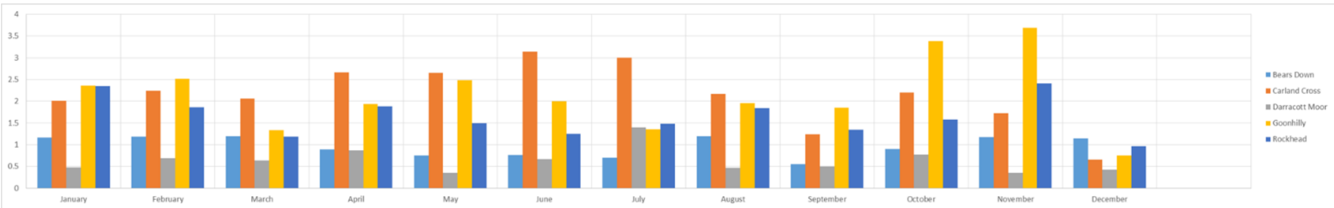

Figure 22: Peak to Peak RMSE Performance of all PV ML Models

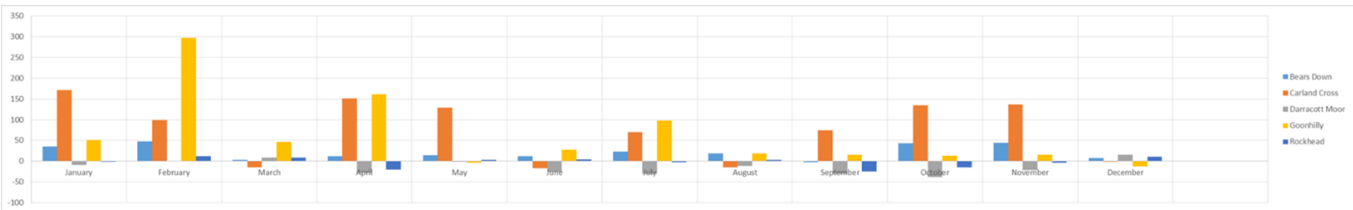

Figure 23: MAPE Performance of all PV ML Models

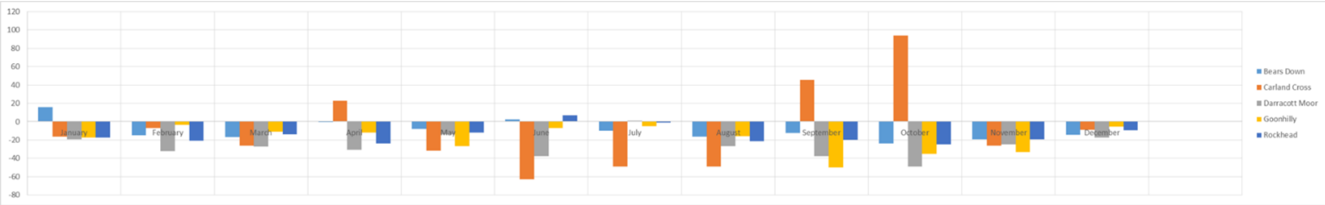

Figure 24: Peak to Peak MAPE Performance of all PV ML Models

Overall the model suggests:

- Weather features are strongly linked;
- Weaker temporal links to historic output data; and
- Weak links to holiday/weekends (work concepts).

Suggestions for further improvements:

- Investigate why certain peak predictions are better than others.

5.7. Machine Learning Model Summary

The overall update can be summarised as:

- Historical weather and weather forecast data introduces improvements across all ML models;
- Each ML type benefits from different features: BSP/Primary- Temporal, Solar/Wind-Weather;
- Peak to Peak data shows that generally good average metrics can hide range of peaks;
- Model in General over estimate outcome, but under estimate peaks; and
- Understanding peaks per model could add to further improvements in the performance of the ML models.

6. CONSTRUCTING ENGINEERING MODELS

The engineering models create a Generator Forecast Model for MW and MVAr via an engineering model using weather data forecasts. WPD provided Renewable Ninja as a model source. This was used for PV, however, the Wind models were poorly developed. Therefore a new wind model source was used, still meeting the criteria that the models should be open source and configurable.

Each model is mathematically explained and all parameters are accessible for future tuning and development.

6.1. Engineering Model PV

The method used in this analysis provides the site output power [6], expressed as a function of incident irradiance and module temperature- bounded by inverter efficiency and maximum site output.

$$
\begin{aligned}
P(G', T') &= G' P_{\mathrm{STC}} \left(1 + k_1 \ln(G') + k_2 (\ln(G'))^2 + k_3 T' \right. \\
&+ \left. k_4 T' \ln(G') + k_5 T' (\ln(G'))^2 + k_6 T'^2 \right) \\
G' &\equiv G / 1000\mathrm{W} \cdot \mathrm{m}^{-2} \\
T' &\equiv T_{\mathrm{mod}} - 25^{\circ}\mathrm{C}
\end{aligned}
$$

$$W_{\mathrm{mod}} = \left(\frac{d_{\mathrm{mod}}}{d_{\mathrm{ane}}}\right)^{0.2} \cdot W_{\mathrm{ane}}$$

$$T_{\mathrm{mod}} = T_{\mathrm{amb}} + \frac{G}{U_0 + U_1 W_{\mathrm{mod}}}$$

Where:

P_{STC}=Module Peak 1kW

T_{amb} =Air Temperature (C)

W_{ane} =Wind Speed at 80m

d_{mod} =Module Height

d_{ane} =Anemometer Height

The coefficients have been calculated through analysis, these will vary from site to site. Two types of PV material have been assumed for this study, CdTe: Cadmium telluride and C-Si: crystalline silicon Figure 25.

[6] Estimating PV Module Performance over Large Geographical Regions: The Role of Irradiance, Air Temperature, Wind Speed and Solar Spectrum. Thomas Huld * and Ana M. Gracia Amillo ECJRC

Module Type	U_0	U_1
c-Si	26.9	6.20
CdTe	23.4	5.44

Module Type	c-Si	CdTe
k_1	−0.017237	−0.046689
k_2	−0.040465	−0.072844
k_3	−0.004702	−0.002262
k_4	0.000149	0.000276
k_5	0.000170	0.000159
k_6	0.000005	−0.000006

Figure 25: PV Material Coefficients from reference study

The external weather input to the PV models is:

- Global Horizontal Irradiance (W/m2) @ 2m
- Temperature (C) @ 2m -
- Diffuse Fraction (Unit)

The data applied is the weather forecast with a noise band applied, as discussed in section 3. Diffuse Fraction is set to zero due to GHI data already being combination of Direct and Diffuse. Fraction is needed if GHI provided as a single element.

6.2. PV Engineering Model Behaviour

The PV Engineering model behaviour is presented as a time-series analysis, using Hatchlands as an example:

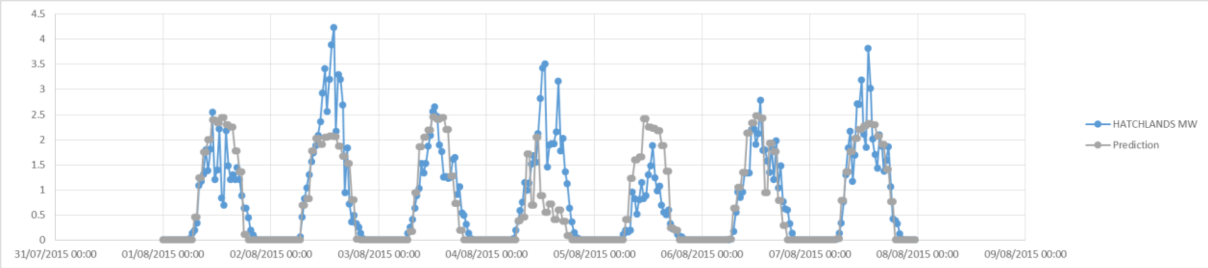

Figure 26: MW Output Vs MW Prediction

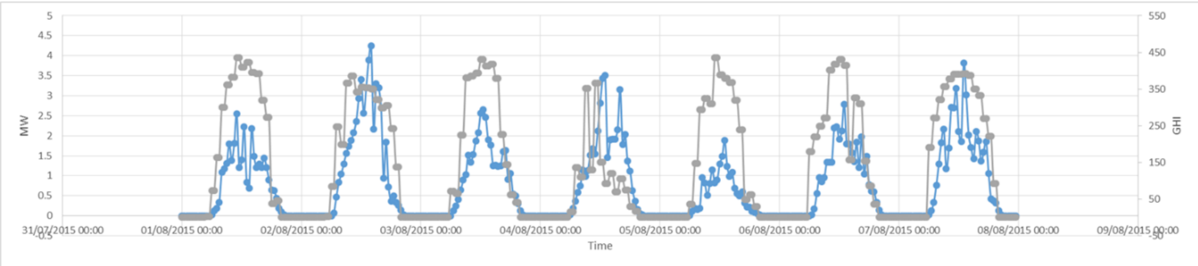

Figure 27: MW Output Vs GHI Input Data

The model follows input data GHI data but site performance is not 100% correlated to GHI data alone. This means that the coefficients could be improved to provide a closer representation of the site behaviour.

6.3. PV Engineering Model Performance Summary

The error performance, similar to the machine learning model analysis was undertaken. The results can be seen below.

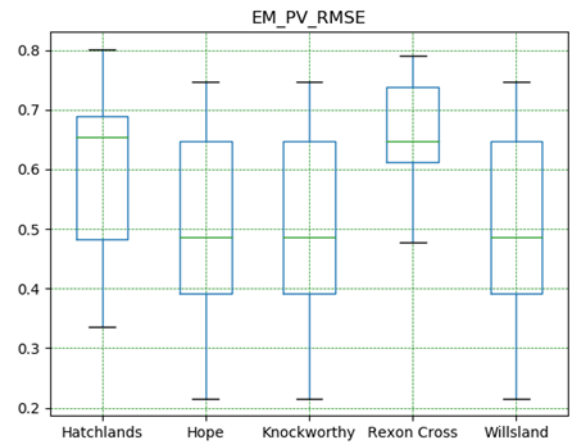

Figure 28: All PV Engineering Model RMSE Performance

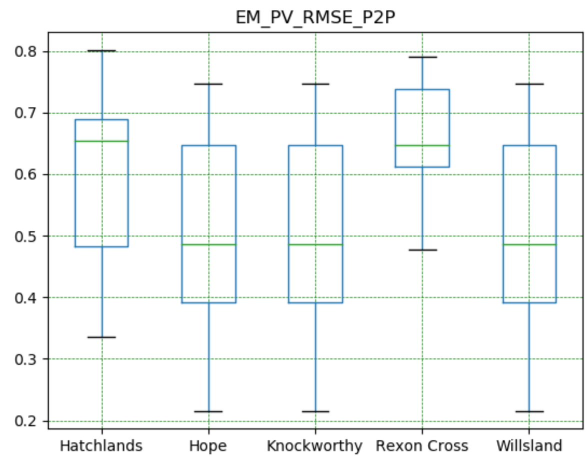

Figure 29: All PV Engineering Model Peak to Peak RMSE Performance

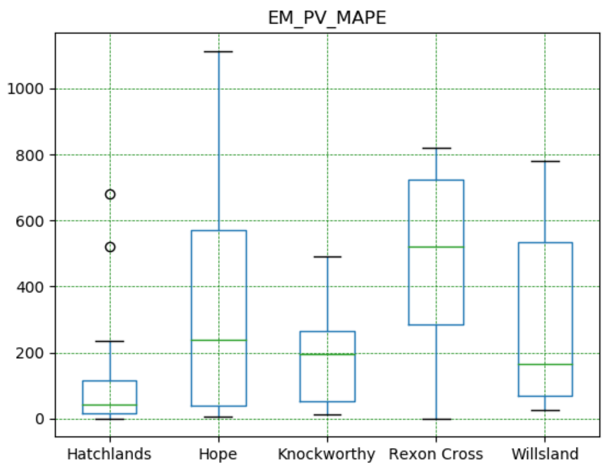

Figure 30: All PV Engineering Model MAPE Performance

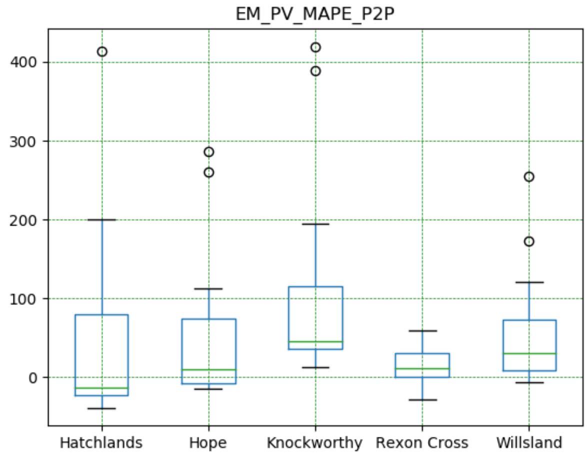

Figure 31: All PV Engineering Model MAPE Peak to Peak Performance

For the results presented in Figure 28, Figure 29, Figure 30 and Figure 31, the general predictions perform similarly across months, compared to one another.

The models provide better peak prediction in summer relative to winter predictions. This is when the performance matters most for PV sites.

Outliers can be quite extreme. The majority are when the site doesn't fully correlate to irradiance available potential. The behaviour shape is not entirely consistent with the site, highlighting some site performance features not embedded in model.

Suggestions for further improvements:

- Adjusting coefficients to improve fit of model behaviour- currently based on test data not associated fully with site;
- More information about the site/sites may allow for improvements to be made to the coefficients (but this illustrates the need to tune the coefficients to obtain better performance and could be required on a site by site basis)
- Wind speed feature not currently live in GSEE module.

6.4. Engineering Model Wind

The method used in this analysis provides the site output power[7] and was provided by the Open Energy Modelling Framework (OEMF). That mode is expressed as a simple model based on the coefficient power (cp)-values of a specific type of a wind turbine. The cp-values are provided by the manufacturer of the wind turbine as a list of cp-values for discrete wind speeds in steps of 0.5 or 1 m/s.

$$P_{wpp} = \frac{1}{8} \cdot \rho_{air,hub} \cdot d_{rotor}^2 \cdot \pi \cdot v_{wind,hub}^3 \cdot cp\left(v_{wind,hub}\right)$$

Where:

- drotor the diameter of the rotor in meters,

[7] **OEMF**Documentation:://readthedocs.org/projects/windpowerlib/downloads/pdf/v0.0.4/

- ρair,hub the density of the air at hub height,
- vwind,hub the wind speed at hub height
- cp the cp-values against the wind speed[8].

The wind speed at hub height is determined by the following equation, assuming a logarithmic wind profile.

$$v_{wind,hub} = v_{wind,data} \cdot \frac{\ln\left(\frac{h_{hub}}{z_0}\right)}{\ln\left(\frac{h_{wind,data}}{z_0}\right)}$$

With vwind, hub the wind speed at the height of the weather model or measurement, hhub the height of the hub and hwind, data the height of the wind speed measurement or the height of the wind speed within the weather model.

Z0 relates to roughness length which models the horizontal mean wind speed near the ground.

The density of the air is calculated assuming a temperature gradient of -6.5 K/km and a pressure gradient of -1/8 hPa/m.

$$T_{hub} = T_{air,data} - 0.0065 \cdot (h_{hub} - h_{T,data})$$

with Tair, data the temperature at the height of the weather model or measurement, hhub the height of the hub and hT, Data the height of temperature measurement or the height of the temperature within the weather model.

$$\rho_{air,hub} = \left(p_{data}/100 - (h_{hub} - h_{p,data}) * \frac{1}{8}\right) / (2.8706 \cdot T_{hub})$$

with pdata the pressure at the height of the weather model or measurement, Thub the temperature of the air at hub height, hhub the height of the hub and hp,data the height of pressure measurement or the height of pressure within the weather model.

The coefficients have are provided by turbine type data Figure 32.

[8] Reiner Lemoine Institut http://vernetzen.uni-flensburg.de/~git/cp_values.csv

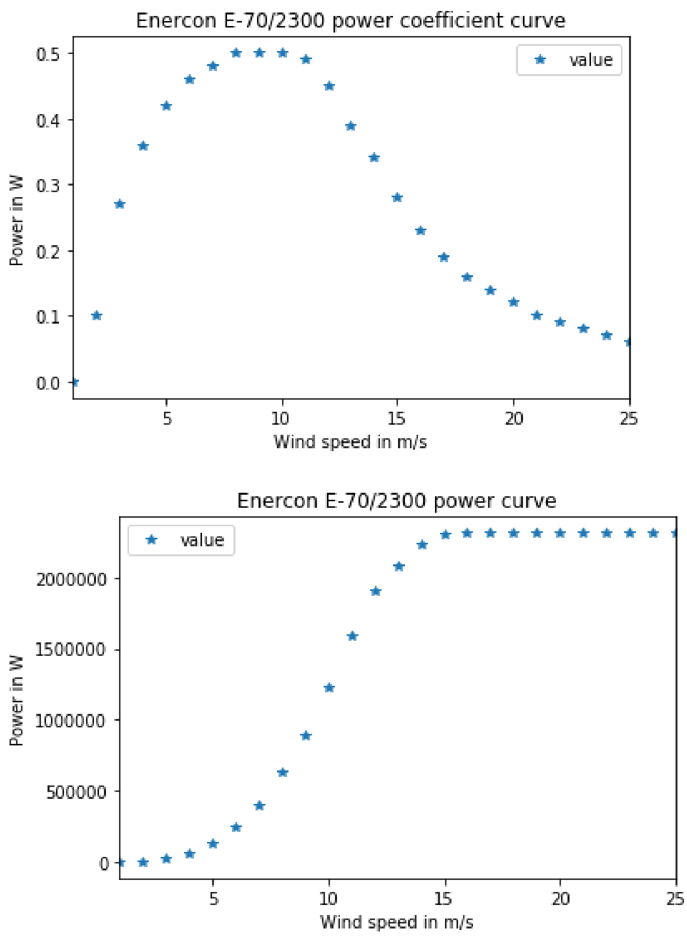

Figure 32: Wind Turbine Coeffcients

The external weather input to the PV models are presented below:

- Wind Speed (m/s) @ 80m
- Temperature (K) @ 2m
- Pressure (kPa) @ 0m (Sea level)

The data applied is the weather forecast with a noise band applied, as discussed in section 3.

6.5. Wind Engineering Model Behaviour

The Wind Engineering model behaviour is presented as a time-series analysis, using Goonhilly as an example:

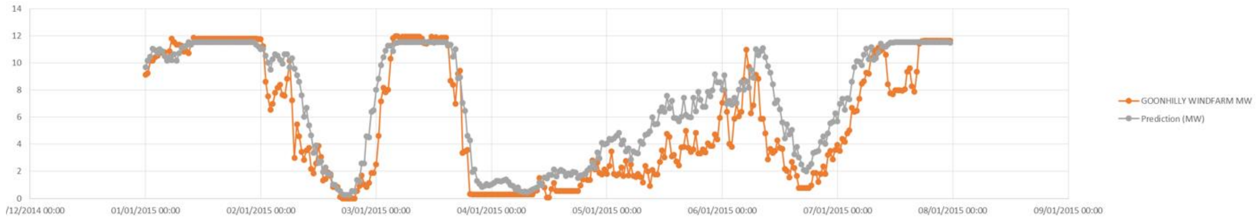

Figure 33: MW Output Vs MW Prediction

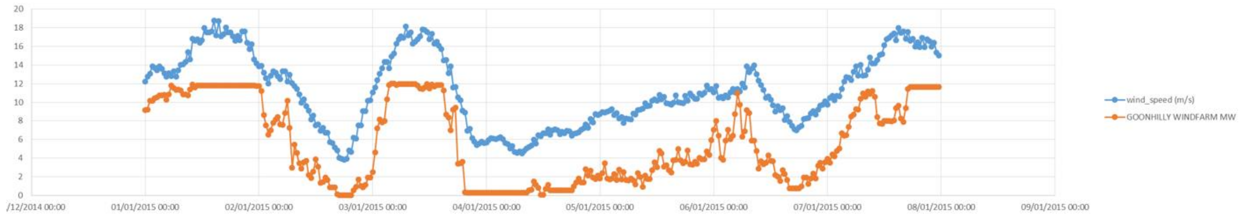

Figure 34: MW Output Vs GHI Input Data

The results show a good correlation between input wind speed and model output.

Due to the output correlation, it shows the majority of the features captured in the input data.

The differences are due to coefficients that vary per site and per turbine.

6.6. Wind Engineering Model Performance Summary

The error performance, similar to the machine learning model analysis was undertaken. The result can seen below:

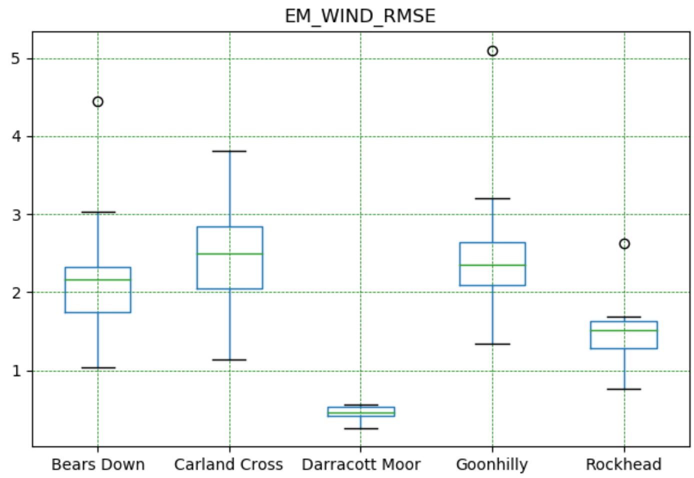

Figure 35: All Wind Engineering Model RMSE Performance

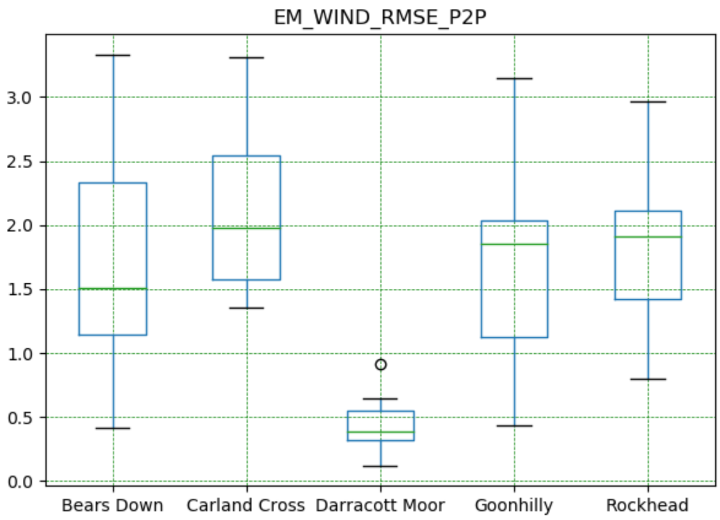

Figure 36: All Wind Engineering Model RMSE Peak to Peak Performance

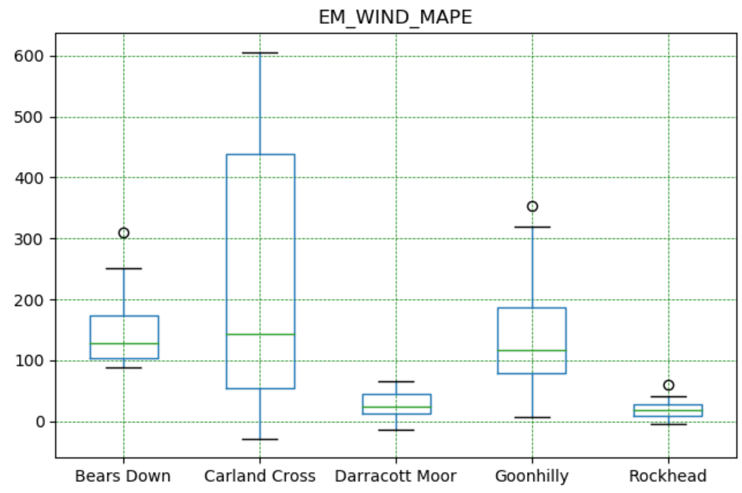

Figure 37: All Wind Engineering Model MAPE Performance

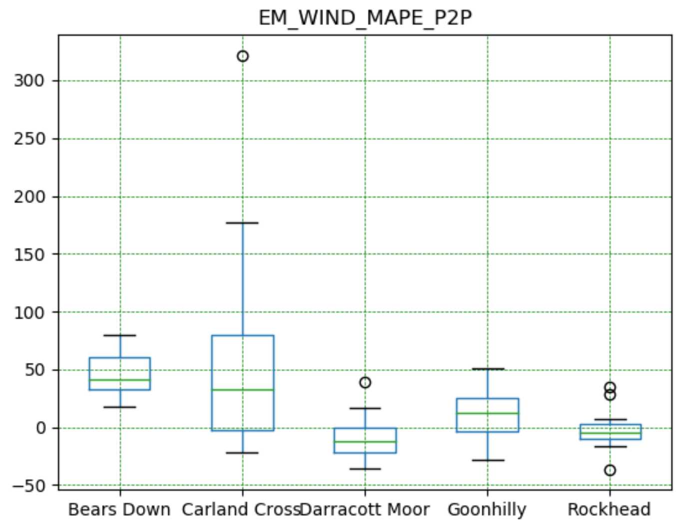

Figure 38: All Wind Engineering Model MAPE Peak to Peak Performance

From the results, Figure 35, Figure 36, Figure 37, and Figure 38 it could be concluded that: wind models with better represented coefficients perform relatively better than others, e.g. Darracott Moor versus Carland Cross. The results show that all the main features captured in the current physical model and that the predictions are consistent across all months.

The following change may lead to a improve performance across the models:

- Adjust coefficients around model to improve fit in line with best performing models. It should be noted this may be required on a site by site basis.

6.7. Engineering Models Conclusions

The following conclusions can be drawn from the analysis of the performance of the engineering models for both wind and PV:

- PV summer models perform well relative to winter;
- Wind models perform well where coefficients are well-tuned;
- Peak prediction was good across well-tuned (coefficient fit) wind models and summer PV models;
- All model will benefit from improved coefficient tuning. Herein lies the issue with operationalising these models. This may be required on a site by site basis.

7. MACHINE LEARNING MODEL VS ENGINEERING MODELS

Comparisons were made for each model against its counterpart. The average performance RMSE and RMSE peak to peak metric for each model across each month was taken.

The difference between these results is presented beow, where a positive represents a better performance from machine learning models, and a negative a better performance from engineering models.

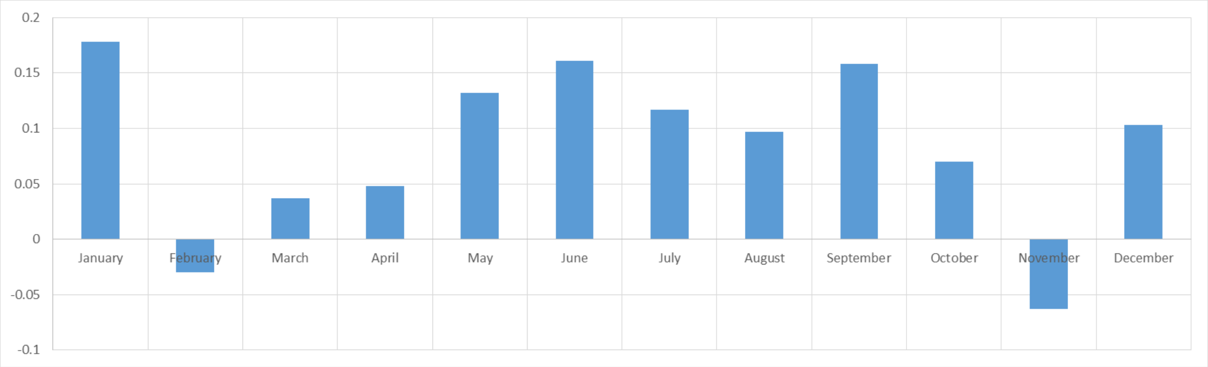

Figure 39: Average RMSE between PV models, across months compared

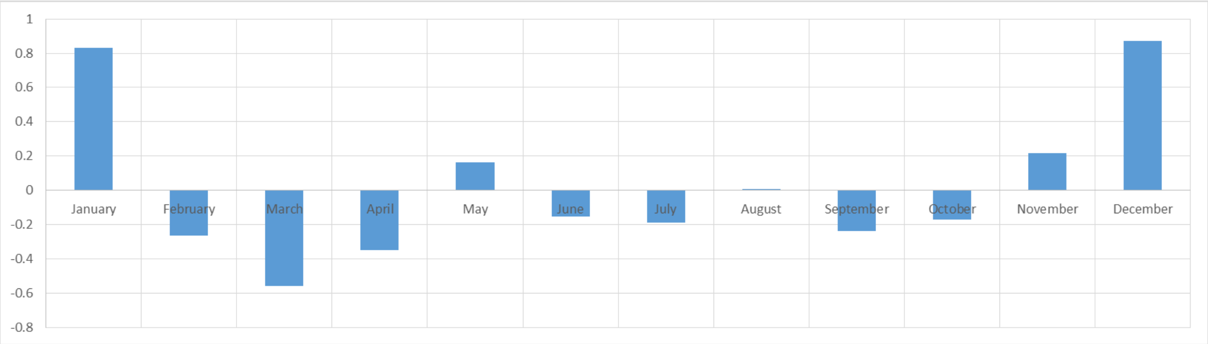

Figure 40: Average Peak to Peak RMSE between PV models, across months compared

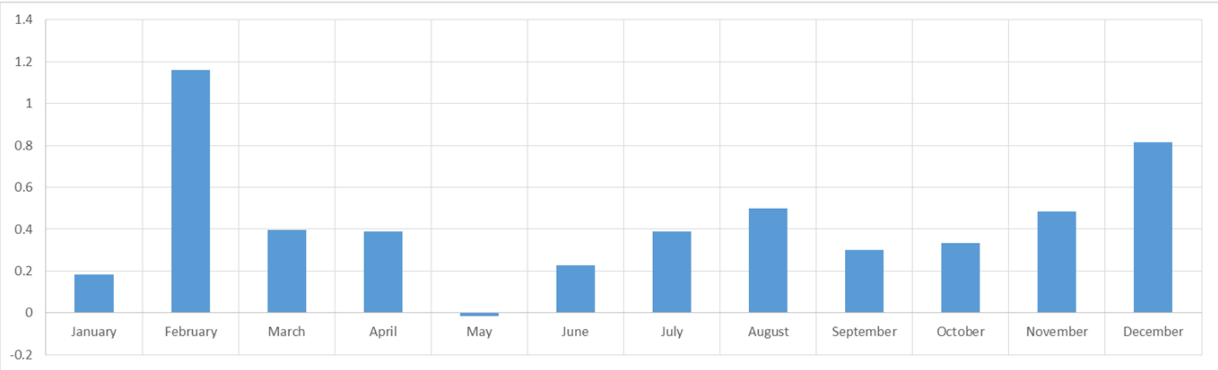

Figure 41: Average RMSE between Wind models, across months compared

Figure 42: Average Peak to Peak RMSE between Wind models, across months compared

7.1. Machine Learning Model vs Engineering Models Conclusions

ML Performs better generally, but EM performs better at peak prediction. The following should be noted:

- ML performs better in general because site operational coefficients are embedded in ML decision trees, whereas EM model a baseline coefficient applied across board. ML tunes itself from the data.
- This is supported by improved behaviour where coefficients more accurate (Darracott Moor Wind)
- EM Peak prediction better than ML. ML is trying to minimise error across the whole time series and not just the peaks and appears to smooth out some peaks.
- Once site engineering models are tuned to site nuances, prior operation is not relevant. Prior erroneous operation, due to site running issues, could get embedded as they do in the ML model if those periods are inadvertently used for training.
- It is easier to tune engineering models to peak, than to site features (embedded in ML decision tree weightings) between zero and peak.

Therefore, the two models provide insight on the general, and peak behaviour of site behaviour. This suggests both model types could be used in together to provide a spread of scenarios bounded by model predictions.

8. CONCLUSIONS

Each element of the change request has been satisfied: improving the machine learning modelling, creating engineering models; and showing the benefits and limitations of each.

The machine learning models were improved by:

- Introducing specific temporal features to machine learning model type, where behavioural changes have an impact e.g. holidays and weekends.
- Introducing weather data and forecast data to enable better pattern identification for machine learning models where rapidly changing weather has a greater impact than diurnal and seasonal effects, e.g. irradiance peaks for solar, and pressure changes for wind.

The engineering models introduced provided:

- An improved view of diurnal peak export with room to improve by more accurate depicting the physical traits of their modally through empirical evaluation (site to site).

The comparing of both has provided insight that a possible ensemble approach to forecasting should be considered in the wider EFFS work to reduce error when procuring services. Or if a single solution is sort, how it highlights the gap in performance that a machine learning model will have to close in order to become as insightful in regards to peak performance.

9. DELIVERABLES

Throughout the project multiple slide decks were presented, discussing the progress and providing further detail as the change request work progressed, these have been provided to the project lead and associated project team and partners, but are available at request.

[1] 200824 WPD EFFS CR01 Forecasting Meeting 1

[2] 200824 WPD EFFS CR01 Forecasting Meeting 2

[3] 200824 WPD EFFS CR01 Forecasting Meeting 3

[4] 200824 WPD EFFS CR01 Forecasting Meeting 4

Further to this the update models were provided to AMT-Sybex and for the wider purposes of model integration, as part of the change request in Jupyter Notebook format.

These have been provided to the project lead and associated project team and partners but are available at request.

9.1. Machine Learning Models

Sowton_Primary_Substation_ML_model.ipynb provided on 29/04/2020

9.2. Engineering Models

PV: PV_Engineering_Model.ipynb provided on 14/05 2020

Wind: Wind_Engineering_Model.ipynb provided on 14/05 2020