UIG Task Force

13.2.5: - Accuracy of NDM Algorithm - use of Weather Data – (basic machine learning)

Summary of Findings

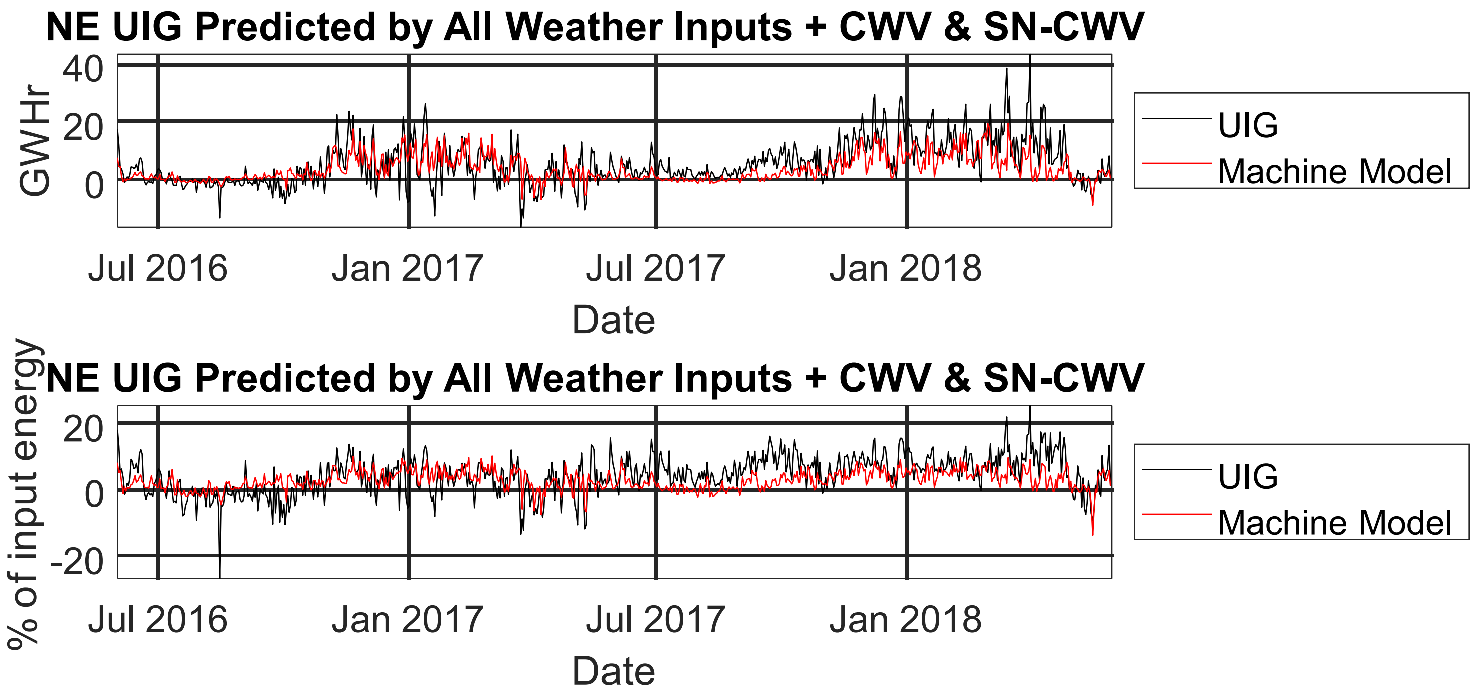
Findings Status	**Closed**
UIG Impact Peak Volatility %	**Reduced by 23%**
UIG Impact Annual Average %	**N/A**
Confidence in Percentages	**M**

Area & Ref #	Accuracy of NDM Algorithm - use of Weather Data – Ref#13.2 (machine learning Ref#13.2.5)
UIG Hypothesis	Can current raw weather data (temperature, possibly wind) and time data (including holidays) be used within a non-linear model to model historical UIG? If machine learning algorithm isn't capable of modelling UIG: can we add additional data (precipitation, irradiance, humidity, wind speed, wind direction). If model does improve UIG, move towards a solution within NDM demand estimation.
Data Tree References	WCF, CWV, ALP, DAF, AQ

Findings	**Approach to analysis**
The use of machine learning showed that it wasn't possible to deliver a 'step change' improvement in results based on the current weather inputs. This indicates that as a next step we need to focus on the inputs to the model, rather than the mechanics of the model itself. Using Machine Learning approaches to predict UIG from data inputs can reduce the absolute volatility by around 23% on average (e.g. a range of volatility of +-10% would be reduced to +-7.7%). The following slides provide the supporting evidence for the results using the existing raw weather data inputs. Confidence in percentages is medium as this analysis is exploratory to identify the potential benefits of using machine learning to inform the NDM Algorithm. A reduction here suggests there could be benefits in using this technology but does not guarantee the same level of improvement when used on supply point level data.	A number of machine learning algorithms were tried against a variety of 'input' data in an attempt to model UIG for that data. This provides an indication of how good the NDM demand estimation algorithm could be with the current data set.

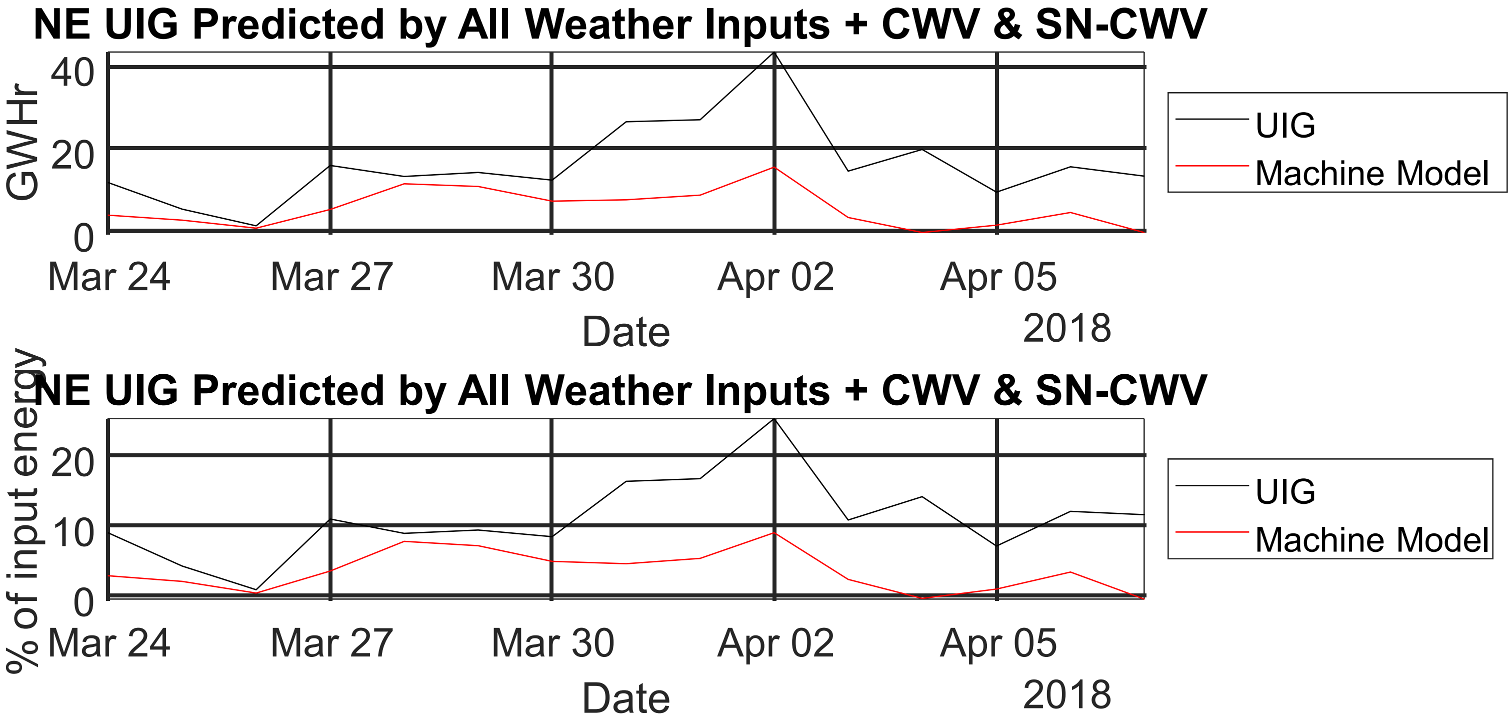
Supporting Evidence (1/3)

Graph illustrating modelled historical vs. observed historical UIG.
Example for the ‘Boosted Tree’ machine learning algorithm shown below



Supporting Evidence (2/3)

Graph illustrating modelled historical vs. observed historical UIG for North Eastern LDZ.
Example result shown for 'Boosted Tree' algorithm below.



Supporting Evidence (3/3)

The first numeric column in the following table shows the impact of removing a data set from the predicted UIG model compared against the maximum improvement of 23.1% volatility reduction. The second numeric column shows the potential volatility reduction using *only* that data set to achieve the UIG prediction. 23.1% is the performance when all datasets are used. Green highlight shows where the data item has the biggest impact on the UIG prediction.

Input Data Set	Performance when input is removed (23.1 when all used) (% Gain)	Performance using only this input (% Gain)
4 hourly temperature	22.8	17.7
Morning Temperature	19.5	13.5
Afternoon Temperature	22	16.2
Wind Speed	22.2	2
CWV and SNCWV	20.1	14.3
Day of week	22.9	0
Holidays	23.1	0.1

Legend
Red = minimal affect
Yellow = marginal affect
Green = substantive affect

Machine learning against UIG with additional data to NDM

Summary of Findings

Findings Status	[Closed]
UIG Impact Peak Volatility %	Reduced by 28%
UIG Impact Annual Average %	N/A
Confidence in Percentages	M

Area & Ref #	Accuracy of NDM Algorithm - Use of Weather Data - Weather Data Inputs (machine learning) (Ref #13.2.5)
UIG Hypothesis	As previous analysis 'Machine learning against UIG for raw inputs to NDM', but expanding into additional weather data (precipitation, irradiance, humidity and pressure). Several machine learning algorithms/models of increasing complexity will be used to try to predict UIG by LDZ, based on weather data inputs. Demand modelling behaviour is governed not only by temperature, but potentially by other key weather parameters that are not part of the model. This task will quickly address whether the additional weather data could reduce UIG volatility.
Data Tree References	WCF, CWV, ALP, DAF, AQ

Findings	Approach to analysis
It was found that the additional weather data does improve the prediction of UIG in the highly non-linear models (Gradient Boosted Tree and Gaussian Process). However the increase is incremental from the previous machine learning results with models using all of the weather data capable of reducing root mean square error percentage UIG from ~75% to ~70%. The linear model achieved ~85%, however the performance is more inconsistent across LDZs. The strongest predictor (after temperature) is irradiance, which can reduce UIG by 3.8%, this is followed by humidity (1.7%), wind (1.0%), precipitation (0.4%) and pressure (0.4%). The combined effect of all of these is 6.1% on average. Additional weather built into a non-linear machine learning model may offer improvements to predict UIG but has only a negligible impact on the peak UIG volatility. Incorporating additional coarse (one reading per LDZ) weather data does help, but does not have a dramatic impact on reducing UIG. When peak volatility has been addressed, a non-linear model that incorporates irradiance may improve the annual average UIG. Confidence in percentages is medium as this analysis is exploratory to identify the potential benefits of using machine learning to inform the NDM Algorithm. The result suggest there could be benefits in using machine learning in the Demand Estimation process and we will therefore explore this further under investigation line item 13.2.6.	Machine learning (ML) models were used to predict the UIG using historical weather data on UIG from the range of 02/06/2012 – 31/05/2018. Available weather data for precipitation, irradiance, humidity, and pressure was assigned to LDZs based upon the distance to the LDZ the weather station. The existing 2 hourly temperature data, 4 hourly wind data, and CWV variable along with the new data (4 hourly) was used to train various ML models to predict UIG.

Supporting Evidence (1/2) – RMSE % UIG reduction

	Inputs to Machine Learning model and RMSE % UIG reduction							
LDZ	temperature only	temperature + wind	temperature + humidity	temperature + precipitation	temperature + irradiance	temperature + pressure	all	all, without CWV
EA	23.3	23.5	23.5	23.1	25.8	23.4	26.2	25.3
EM	25.0	25.7	29.4	26.4	30.2	24.8	33.7	32.5
NE	20.6	21.8	24.6	22.7	23.9	22.6	28.7	27.6
NO	20.2	20.3	23.3	20.2	25.5	20.5	28.2	26.8
NT	19.7	20.8	19.6	19.0	23.6	20.2	25.5	23.8
NW	30.2	32.3	30.5	30.2	31.5	30.3	33.7	32.3
SC	18.4	18.4	20.2	18.4	20.7	18.9	22.9	22.2
SE	18.0	19.4	19.1	17.8	21.1	18.8	23.5	22.5
SO	29.0	29.1	29.5	29.0	33.7	28.6	34.3	33.5
SW	21.4	22.2	23.9	22.1	26.7	22.1	28.4	27.2
WM	17.8	19.5	19.4	18.5	23.7	17.7	26.3	24.7
WN	26.3	27.2	26.8	27.2	29.4	26.8	30.4	27.2
WS	15.1	16.4	17.1	17.1	21.4	16.6	23.6	23.6
national	22.3	23.3	24.0	22.7	26.1	22.7	28.4	27.2

Table showing relative predictive capability of inputs. Numbers are RMS (root mean squared) percentage reductions per day in the UIG per LDZ as determined using the Gradient Boosted Decision Tree model. All predictions were made using the current day's weather data. Information on holidays and day of the week were not included.

By including the previous day's temperature readings in task 'Machine learning against UIG for raw inputs to NDM', we were able to reduce the UIG by a further 0.6%. The effect should be similar here, taking the total reduction in UIG to 28.4%.

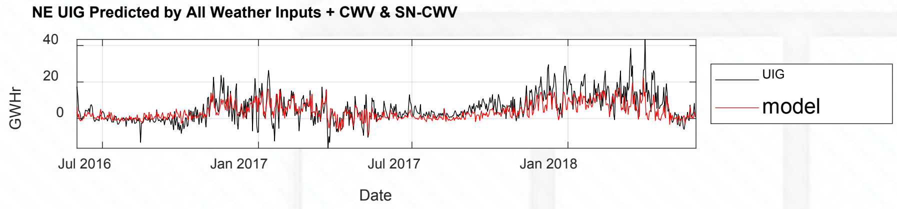
Input Data 01/10/2012 – 31/05/2018

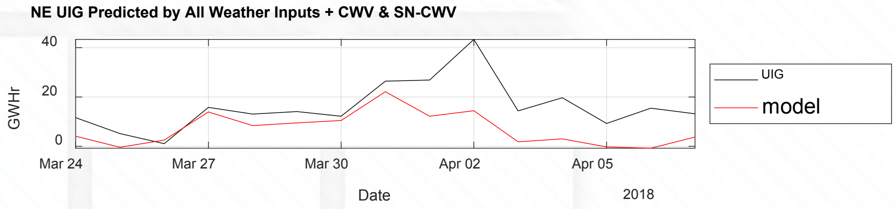
A non-linear model can predict UIG using temperature and wind data which effectively reduces UIG volatility by 23.3% nationally

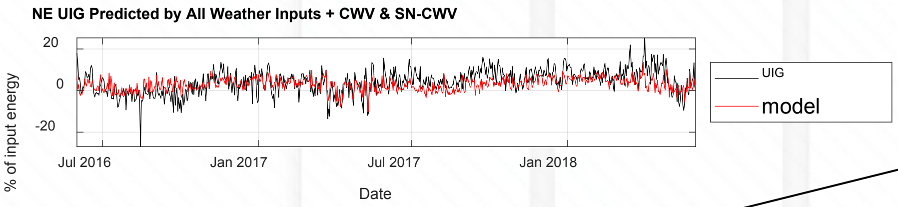
A non-linear model that uses temperature and irradiance can predict UIG using temperature and wind data which effectively reduces UIG volatility by 28.4%

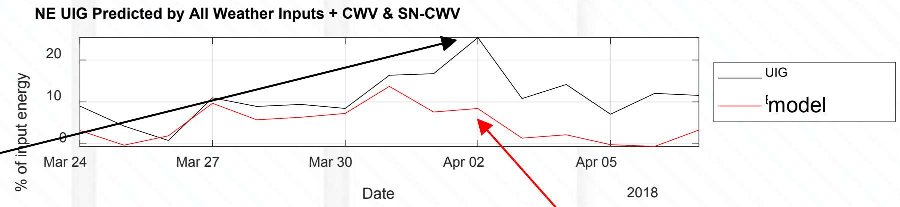
Supporting Evidence (2/2) – Modelled UIG plots

Graphs illustrating modelled historical (i.e. UIG predicted by the machine learnt model) vs. observed historical UIG. This is for the Gradient Boosted Decision Tree model using of the weather inputs, (temperature, humidity, precipitation, irradiance, wind speed). The data on the right illustrates a volatile region in more detail.









For the peak at 2nd April 2018, UIG was 25%. The non-linear machine learnt model (with all the weather inputs) predicts UIG of 9%. **This suggests that such a model might be able to reduce this peak UIG from 25% to 16%.** There was no improvement on this value from the comparable machine learning results in task ‘Machine learning against UIG for raw inputs to NDM’.

Correlation of UIG against raw inputs on sample set for low EUCs

Summary of Findings

Findings Status	**Closed**
UIG Impact Peak Volatility %	**N/A**
UIG Impact Annual Average %	**N/A**
Confidence in Percentages	**N/A**

Area & Ref #	Accuracy of NDM Algorithm - Use of Weather Data - Weather Data Inputs (machine learning) (Ref # 13.2.5)
UIG Hypothesis Specific Item	The finding for '13.3.2 NDM Sample Data Outliers' illustrated that there was some possibility that low EUCs appeared to track each other very well, and showed some signs of tracking UIG volatility but might be masked by the behaviour seen in the weather insensitive large EUCs.
Data Tree References	Weather, EUC, Day of the week, Market Sector, CWV

Findings	**Approach to analysis**
• UIG for Industrial users in EUC1 has very significant correlations with day-of-week • UIG for Domestic users also have significant day of week correlations in some LDZs, but not in others (SC, EA, NE, SO & SW) • There are significant correlations with CWV in EUC1 industrial when the EUC derived from the sample measurement is different to the UK Link EUC, occasionally also in higher EUCs • Day-of-week correlations are more prominent in LDZs where the sample set UIG is large relative to the total gas usage. This suggests where this problem exists, it is very significant. • **Observation 1:** By looking at correlations on low EUCs (and further breaking down by market sector code), the model error volatility (i.e. sample UIG peak to peak) may be reduced to 30% of its value by a simple linear correlation. This supports the hypothesis that effects from higher EUCs are masking effects in the lower EUCs. • **Observation 2*:*** The clear differences across LDZs suggest some LDZs model the sample set better than others. As the sample set is changing this may be due to the new sample set being different in some significant way to the sample set the model was developed on. We suspect this is due to different AQ representations. • **Observation 3*:*** There is a difference between the EUC1 Industrial/Commercial and Domestic behaviours. This suggests that the current segmentation into EUCs based solely on AQs should be extended to include the I/C & D split. • This means that there may be benefit in producing Demand models from the sample set using techniques to compensate for the AQ distribution in the sample set being different from the user population.	For EUCs 1-5, the demand estimation error (sample UIG) in the sample set in each EUC was split to highlight the different contributions from sites with correct AQs (AQs for which the stated AQ matches the measured AQ) and incorrect AQs, and EUC 1 was further split into Domestic and Industrial users according to the market sector code. The resultant sample set UIG was correlated against day-of-week, CWV, and change in CWV. Residual plots and R-squared statistics are presented

Supporting Evidence (1/4) – R-squared values for day-of-week correlation

These tables show the R-squared values (i.e. strength of correlation described as a range 0 to 1 where higher is better) correlating the sample data demand error (sample UIG) with a weekend / weekday flag.

- **There are significant correlations with day of the week in EUC1, occasionally also in EUC2**
  - **Industrial users in EUC1 (EUC1_I) have very significant correlations with day-of-week**
  - **Domestic users in EUC1 (EUC1_D) also have significant correlations in some LDZs, but not in others (SC, EA, NE, SO & SW)**
- **NOTE – The green cells show strong correlation, the red show low/no correlation**

R-squared values for sample data allocated into EUCs band that match the measured consumption (**'matching' EUC**)

LDZ	EUC1_D	EUC1_I	EUC2	EUC3	EUC4	EUC5
EA	0.02	0.45	0.09	0.06	0.01	0.1
EM	0.24	0.59	0.25	0.01	0.01	0
NE	0.06	0.28	0	0.02	0.01	0.01
NO	0.48	0.68	0.43	0.37	0.19	0.12
NT	0.38	0.57	0.01	0.01	0	0.15
NW	0.31	0.62	0.33	0.08	0	0.03
SC	0.09	0.53	0.04	0.03	0.08	0
SE	0.23	0.67	0	0.07	0.04	0
SO	0.08	0.59	0	0	0	0.16
SW	0	0.41	0.03	0.06	0.03	0.04
WM	0.42	0.7	0.21	0.06	0	0.11
WN	0.19	0.64	0.29	0.24	0.08	0.07
WS	0.45	0.56	0.41	0.26	0.02	0.1

R-squared values for sample data allocated into EUCs band that do not match the measured consumption (**'mismatch' EUC**)

Row	EUC1_D	EUC1_I	EUC2	EUC3	EUC4	EUC5
EA	0.25	0.01	0.15	0.22	0	0.01
EM	0.01	0.08	0	0.03	0.13	0.03
NE	0.36	0.11	0.07	0	0.03	0.04
NO	0.07	0.13	0.56	0.19	0.02	0.02
NT	0.08	0.01	0.32	0.08	0.02	0.02
NW	0.05	0	0.05	0.01	0.34	0.02
SC	0	0.02	0.01	0.01	0.05	0.12
SE	0	0.27	0.11	0	0	0.02
SO	0.45	0.44	0.03	0	0.09	0.11
SW	0.04	0	0.12	0.06	0.09	0.03
WM	0.01	0.13	0.04	0.02	0.09	0.16
WN	[No Mismatch]	0.15	0.07	0.02	0.01	0.03
WS	0.01	0.3	0.69	0.22	0.02	0.09

Supporting Evidence (2/4) – R-squared values for CWV correlation

These tables show the R-squared (i.e. strength of correlation described as a range 0 to 1 where higher is better) values correlating the sample data UIG with CWV.

- **There are significant correlations with CWV in EUC1 industrial, occasionally also in higher EUCs**
- **NOTE – The green cells show strong correlation, the red show low/no correlation**

R-squared values for sample data allocated into EUCs band that match the measured consumption (i.e. **'matching' EUC**)

Row	EUC1_D	EUC1_I	EUC2	EUC3	EUC4	EUC5
EA	0.18	0.11	0.03	0.09	0.09	0.1
EM	0	0.01	0.03	0	0.07	0.19
NE	0.01	0.01	0.02	0.1	0.04	0.01
NO	0.02	0.01	0.1	0	0	0.03
NT	0.02	0.23	0.02	0.03	0	0.01
NW	0	0	0.11	0.01	0.01	0.01
SC	0.01	0.06	0	0.18	0.15	0
SE	0.01	0.25	0.2	0.06	0	0.08
SO	0.1	0	0.13	0.01	0.01	0.2
SW	0.44	0.04	0	0.04	0	0.07
WM	0.04	0.01	0.08	0.06	0.08	0.25
WN	0.15	0	0.1	0.01	0.03	0.01
WS	0.02	0	0.01	0	0.07	0.01

R-squared values for sample data allocated into EUCs band that do not match the measured consumption (i.e. **'mismatching' EUC**)

Row	EUC1_D	EUC1_I	EUC2	EUC3	EUC4	EUC5
EA	0.11	0.75	0.55	0.27	0.22	0.25
EM	0.21	0.77	0.38	0.02	0.57	0.24
NE	0.01	0.39	0.16	0.57	0.39	0.36
NO	0.04	0.83	0.44	0.47	0.23	0.21
NT	0.19	0.61	0.24	0.06	0.36	0.38
NW	0.33	0.77	0.08	0.42	0.01	0.35
SC	0.14	0.91	0.75	0.38	0.03	0.61
SE	0.21	0.77	0.96	0.58	0.56	0.75
SO	0.2	0.24	0.93	0.08	0.39	0.84
SW	0.03	0.81	0.67	0.16	0.01	0.58
WM	0.27	0.76	0.15	0.38	0.18	0.37
WN	[No Mismatch]	0.07	0.01	0.26	0.73	0.05
WS	0.86	0.71	0.03	0.07	0.2	0

Supporting Evidence (3/4) – WM EUC1 Sample Set UIG correlations with Domestic/Commercial Split

These graphs show the sample demand estimation error (Effectively UIG in the Sample set, blue) and the residual error after correlation with CWV (green) and a combination of day-of-week with CWV (black), for West Midlands (WM LDZ), which has a large UIG percentage on the sample set.

The analysis attempts to remove the UIG predicted by various inputs to leave the remaining UIG which could be driven by other factors. The smaller these residual values, the more UIG can potentially be explained by changes in the tested input variable.

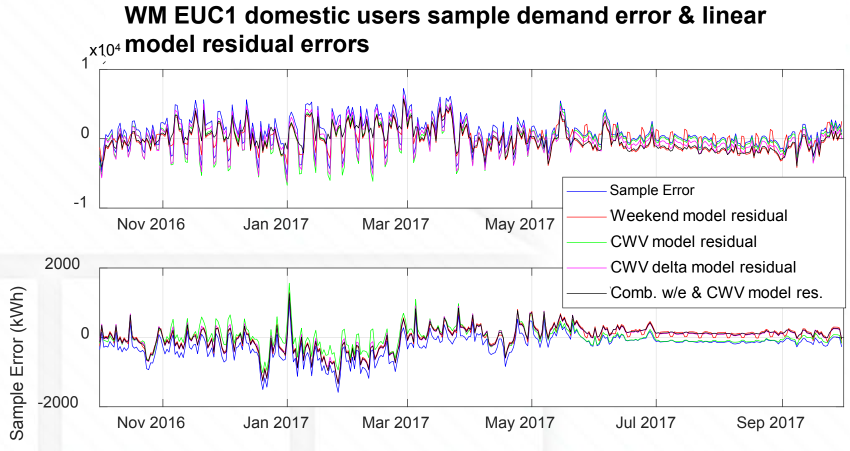
The top graph shows the residuals for Domestic users with UK Link EUCs that match an EUC derived from the NDM sample consumption. There is clear day-of-week effects (R=0.42), significantly mitigated by the correlation (although some residual remains, suggesting the relationship could be more complicated than linear)

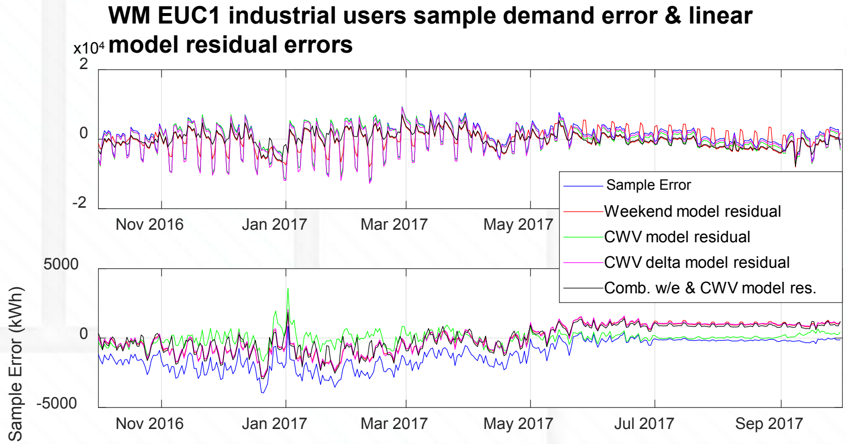
The second graph shows the residuals for Domestic users with UK Link EUCs that do not match an EUC derived from the NDM sample consumption. There is a clear CWV (R=0.27) and day-of-week effect, but only the CWV leads to an improvement. A combination of factors may lead to further reduction.

The third graph shows the residuals for Industrial users with matching EUCs. There is a very large day-of-week effect (R=0.7), which is significantly reduced by the day-of-week linear model.

- **Note the necessity for the combination of day-of-week with CWV here – red is the day-of-week only residual, and its performance is limited because it injects errors in the summer.**

The fourth graph shows the residuals for Industrial users with mismatching EUCs. There is a possible weak day-of-week effect, but this is dominated by a CWV effect(R=0.76).





Supporting Evidence (4/4) – SC EUC1 Sample Set UIG correlations with Domestic/Commercial Split

These graphs show the sample error (blue) and residual after correlations with CWV (green) and a combination of day-of-week with CWV (black). Scotland has a relatively small UIG percentage on the sample set

The top graph shows the residuals for Domestic users with matching EUCs. There is no day-of-week or CWV effects (R<0.09),

The second graph shows the residuals for Domestic users with mismatching EUCs. Again, there is no day-of-week or CWV effects (R<0.14),

The third graph shows the residuals for Industrial users with matching EUCs. There is a large day-of-week effect (R=0.53), which is significantly reduced by the day-of-week linear model.

The fourth graph shows the residuals for Industrial users with mismatching EUCs. There is a possible weak day-of-week effect, but this is dominated by a very large CWV effect(R=0.91), .