UNC Workgroup 0754R

30/11/2021

xoserve

Provided by: correla

Contents

- Workgroup Meeting 4 Recap (5th October 2021)
    - Key Discussion Points
- Update on Area 1 Progress:
    - Approach to Analysis
    - Model Development / Verification
    - Test Modelling Error
- Conclusions and Next Steps
- Timeline

Contents

Useful Links

- Uniform Network Code Section H
- Demand Estimation Methodology
- Demand Modelling Approach (2021 version)
- UIG Task Force Findings
- NDM Algorithm Consultation Material
- UNC Request for 0754R Workgroup

Glossary

- For those not familiar with all the industry abbreviations please find full name of those used in this presentation below:
    - ALP: Annual Load Profile
    - AUGE: Allocation of Unidentified Gas Expert
    - CDSP: Central Data Services Provider
    - CWV: Composite Weather Variable
    - DAF: Daily Adjustment Factor
    - DESC: Demand Estimation Sub Committee
    - DM: Daily Metered
    - DOW: Day of Week
    - EUC: End User Category
    - ILF: Indicative Load Factor
    - LDZ: Local Distribution Zone
    - MAPE: Mean Absolute Percentage Error
    - MPE: Mean Percentage Error
    - NDM: Non-Daily Metered
    - PLF: Peak Load Factor
    - SNCWV: Seasonal Normal Composite Weather Variable
    - SND: Seasonal Normal Demand
    - UIG: Unidentified Gas
    - UNC: Uniform Network Code
    - WCF: Weather Correction Factor
    - WSENS: Weather Sensitivity

Meeting 4 Re-cap
(5th October 2021)

Background

- UIG Task Force produced a number of recommendations to help reduce temporary UIG levels/volatility. This included findings associated with the modelling error within the NDM Algorithm
- DESC is responsible for the NDM Algorithm (UNC Section H) and has an obligation to review it every 3 years (UNC H 2.2.2)
- Prior to moving forward with the above a consultation was performed during Q4 of 2020 to assess the levels of support for making improvements to the NDM Algorithm
- A more detailed view of the background to this Workgroup and current state overview is provided in the March meeting papers here

Rationale for Workgroup 0754R:

- Supports DESC's UNC obligation to review the NDM Algorithm
- UIG Task Force findings will be explored and progressed
- Clear industry support for investigating advanced analytical approaches
- A Workgroup maintains focus and increases visibility across the industry
- Improved NDM Allocation will result in a reduction in UIG volatility and subsequent Meter Point reconciliation/UIG volumes (temporary)

Key Discussion Points

The main headlines from Meeting 4 of 754R were……

Area 1: Investigating alternative Machine Learning approaches to develop EUC demand models:

- Data Set Up Considerations
- Exploration of various Machine Learning approaches
- First set of results shared trialing different approaches – 'Gradient Boosting' had provided some promising output
- Talked through the challenges of translating the Machine Learning output into the parameters we require (particularly the DAF)
- Explained that more time was required to develop and understand the various modelling approaches e.g. Neural Networks had not provided good results initially

Fig. 1

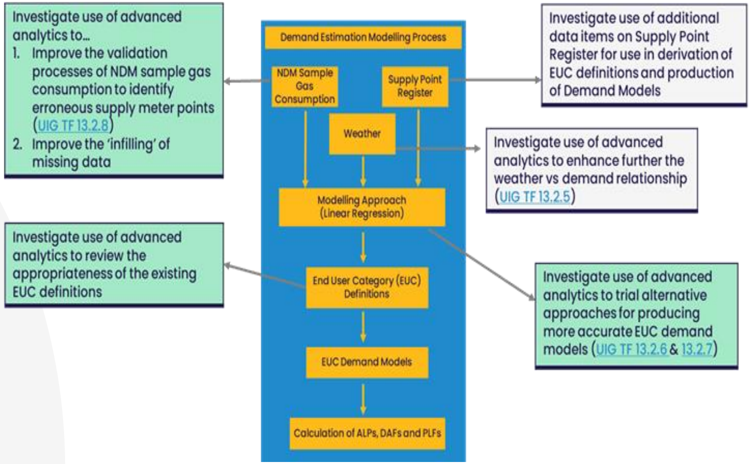
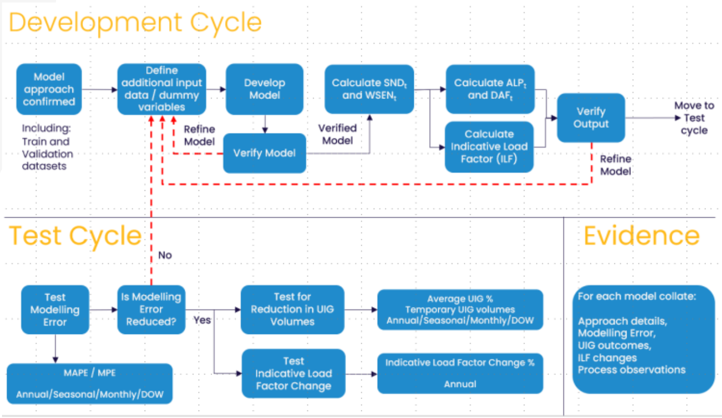
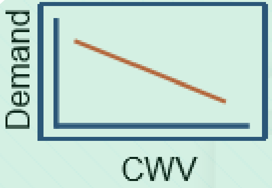


Fig. 2



Update on Area 1 Progress

Area 1: Trial alternative approaches to deriving SND_t



Linear Regression between NDM Sample Demand and Daily Composite Weather Variable

Current Approach

EUC Demand Models (39 per LDZ)

Model Smoothing, the process of averaging [3] individual models, is applied to create the final form of the EUC demand model.

$SND_t = P_t * (C1 + C2 * SNCWV_t)$,
$WSENS_t = P_t * C2$

P_t is a multiplicative factor which is set to 1 for a non-holiday Monday to Thursday and potentially a different value for weekend or holiday day.

C_1 is the constant from the (non-holiday) Monday to Thursday regression of daily demand against daily CWV.

C_2 is the slope of this regression line which represents the weather sensitivity of demand for (non-holiday) Mondays to Thursdays.

Annual Load Profiles

The ALP_t shall be determined as:

$$ALP_t = \frac{SNDE_t}{\frac{[\sum_{t=1}^{N} SNDE_t]}{N}}$$

where

t denotes the value for a particular day

E denotes the EUC

N is the number of days in the Gas Year;

Daily Adjustment Factor

The DAF_t shall be determined as:

$$DAF_t = \frac{WVCE_t}{SNDE_t}$$

where

E denotes the EUC

t denotes the value for a particular day

SND – Seasonal Normal Demand
WVC - Weather Variable Coefficient;

Peak Load Factor

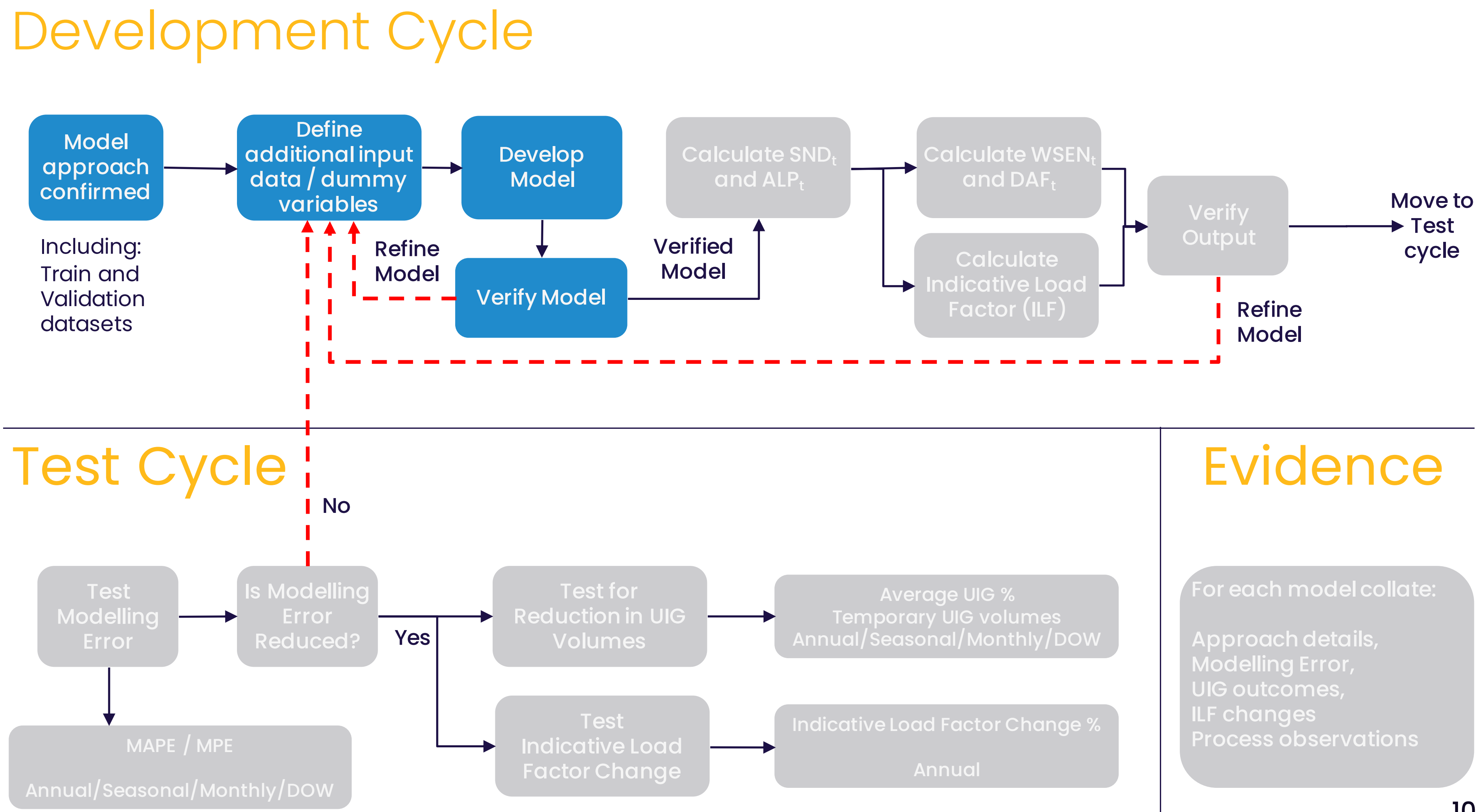
The PLF for the EUC sample is calculated as:

$$\frac{\text{Aggregate AQ from the EUC model}}{\text{1 in 20 peak demand from the EUC model} * 365}$$

An aggregate AQ is derived from the smoothed EUC demand model by setting the composite weather variable to its seasonal normal level in the model and summing the resulting demand values over the 365 days of the forecast year (excluding any 29^{th} February).

For NDM EUCs a 1 in 20 peak day demand estimate is derived from each gas demand EUC model by simulation using the smoothed EUC demand model in conjunction with the database of historic daily composite weather variable values for the appropriate LDZ.

Approach to Analysis


Development Cycle
Model approach confirmed
Including: Train and Validation datasets
Define additional input data / dummy variables
Develop Model
Verify Model
Refine Model
Verified Model
Calculate SND_t and ALP_t
Calculate $WSEN_t$ and DAF_t
Calculate Indicative Load Factor (ILF)
Verify Output
Refine Model
Move to Test cycle
Test Cycle
No
Test Modelling Error
MAPE / MPE
Annual/Seasonal/Monthly/DOW
Is Modelling Error Reduced?
Yes
Test for Reduction in UIG Volumes
Average UIG %
Temporary UIG volumes
Annual/Seasonal/Monthly/DOW
Test Indicative Load Factor Change
Indicative Load Factor Change %
Annual
Evidence
For each model collate:
Approach details,
Modelling Error,
UIG outcomes,
ILF changes
Process observations

Development Cycle

- We have been looking at Machine Learning Techniques in SAS Enterprise Miner
- Mainly focussing on NW 01BND to test methods and understanding, then expanding to the agreed core 6 EUCs of
  - 01BND, 02BNI and 05B in
  - LDZs NW and SE
- We are training using sample data from April 2017 to March 2020, excluding COVID affected days where possible
- Testing is against October 2019 to September 2020 at present

Gradient Boosted and Neural Networks have been a focus as these were the favoured models from the UIG Taskforce. We have tried a number of other methods as well, see meeting 4 slides for more information.

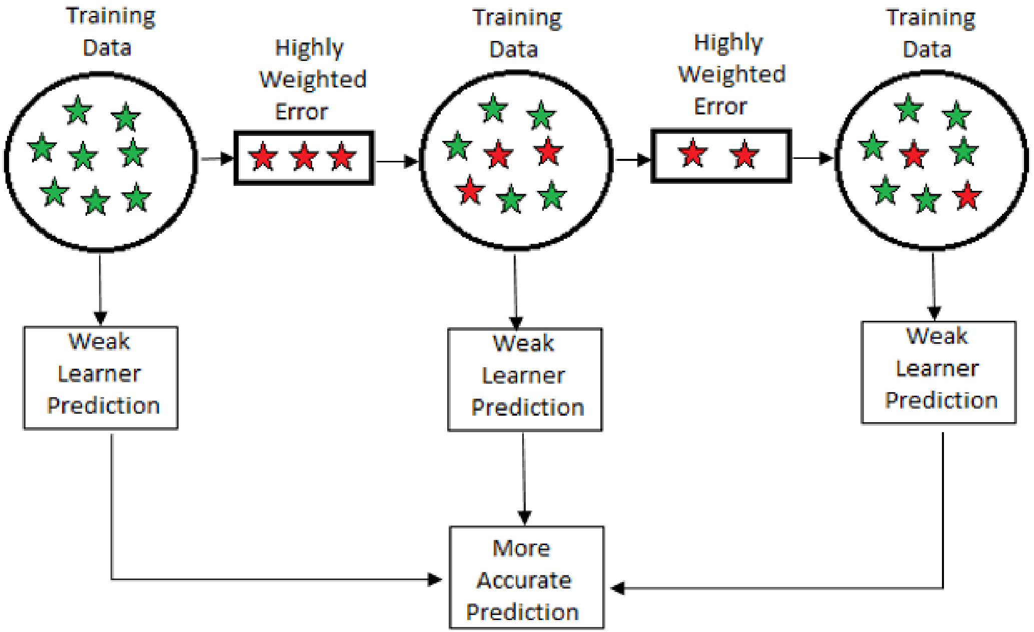
In meeting 4 we presented the progress so far and noted 2 areas of focus which are covered on the following slides

- Improvement to models
  - Including further investigation into previously visited methodology such as Neural Networks
- Delivering DAFs and Load Factors
  - These are fundamental as they present a risk to not being able to deliver a Machine Learning solution

Results shared here today are a progress update and are not final, as further model understanding / refinement is planned

Gradient Boosting

- Boosting is an ensemble learning method that combines a set of weak learners into a strong learner to minimize training errors.
- In boosting, a random sample of data is selected, fitted with a model and then trained sequentially—that is, each model tries to compensate for the weaknesses of its predecessor.
- With each iteration, the weak rules from each individual classifier are combined to form one, strong prediction rule.
- Gradient boosting trains on the residual errors of the previous predictor. The name, gradient boosting, is used since it combines the gradient descent algorithm and boosting method.

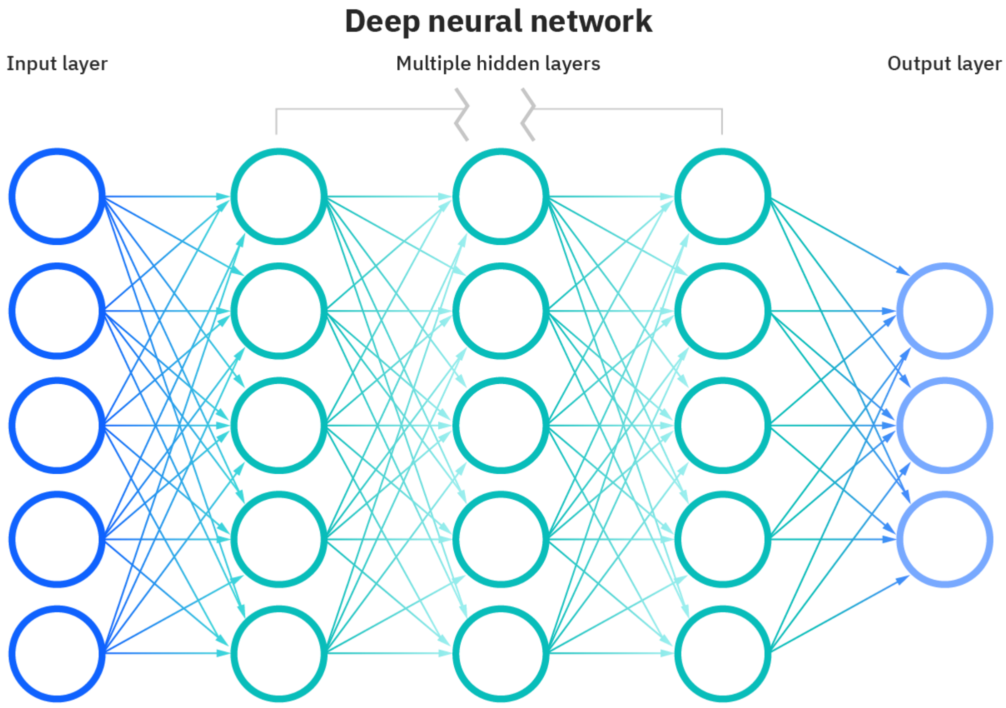


Neural Networks

Neural networks, also known as Artificial Neural Networks (ANNs) or Simulated Neural Networks (SNNs), are a subset of machine learning and are at the heart of deep learning algorithms.

- Their name and structure are inspired by the human brain, mimicking the way that biological neurons signal to one another.
- Artificial Neural Networks (ANNs) are comprised of node layers, containing an input layer, one or more hidden layers, and an output layer.
- Each node, or artificial neuron, connects to another and has an associated weight and threshold.
    - If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network.
    - Otherwise, no data is passed along to the next layer of the network.

We got little success with the full Neural Network approach, however combining a Neural Network with Generalised Linear Modelling has produced relatively good results which are covered on later slides.



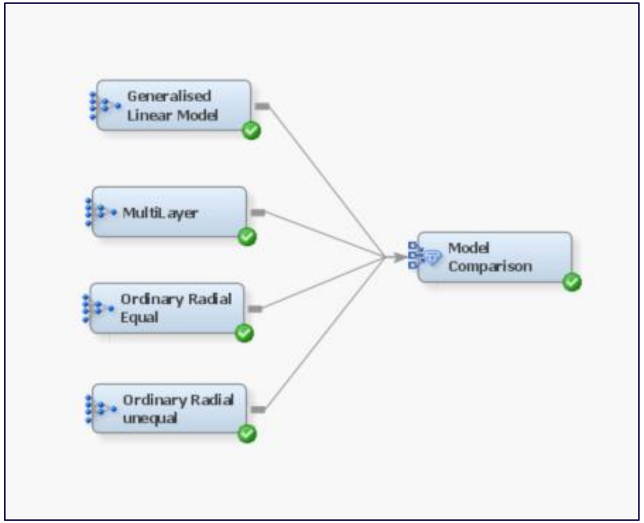
Model Verification Methods

Our initial results were based on a default optimisation approach. After investigation a number of alternatives were identified.

- The diagram to the right shows the set up for investigating different Neural Network approaches
  - Multilayer - default
  - Generalised Linear Model (GLM)
  - Ordinary Radial (Equal width)
  - Ordinary Radial (unequal width)

- Based on Average Square Error measure, the GLM model has produced the lowest value – a measure of how well the training and prediction follows each other

- This is not the only technique we have used as this does not take into account factors such as seasons, month, day of the week , holidays etc.

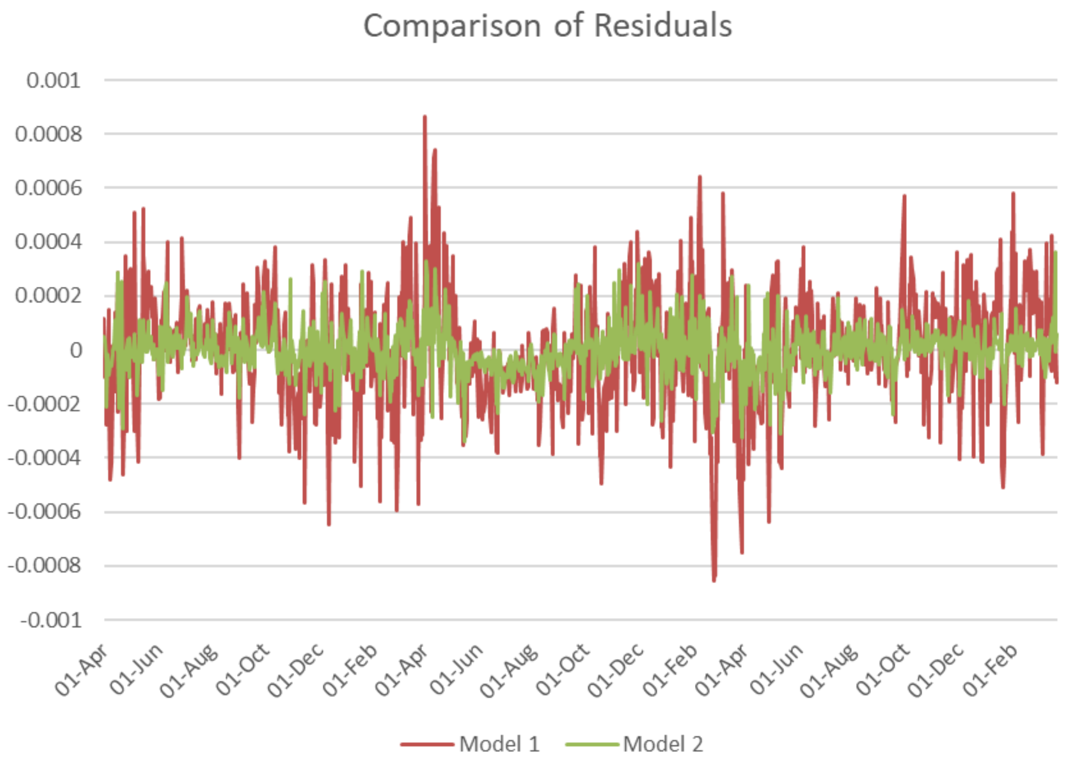


Fit Statistics
Model Selection based on Train: Average Squared Error (_ASE_)

Selected Model	Model Node	Model Description	Train: Average Squared Error	Train: Misclassification Rate
Y	EndGrp	End Groups NN_GLM	.000000035	.
	EndGrp3	End Groups NN_MultiLayer	.000000260	.
	EndGrp4	End Groups NN_O_Rad_Equal	.000000328	.
	Neural4	Neural Network O_Rad_UnEqual	.000000712	.

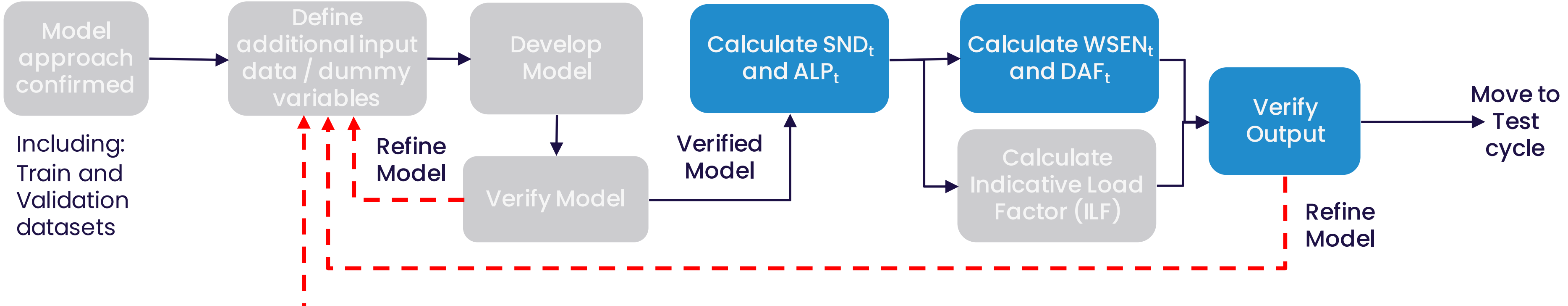
Model Verification Methods

- Comparing residuals can give an indication of how good a model has performed, however this doesn't always give a true picture of the results
- The residual may be larger due to a more accurate DAF and only working the model through shows this
- Both looking at the shape of the forecast and the value of the residuals gives an good indication of which models are worth moving forward with

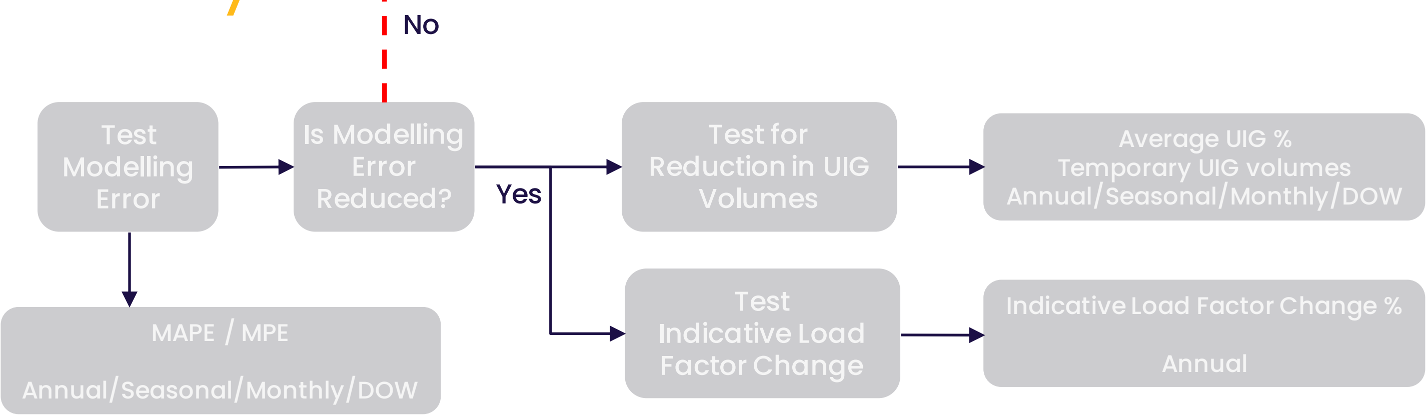


Model Development

Development Cycle



Test Cycle



Evidence

For each model collate:

Approach details,
Modelling Error,
UIG outcomes,
ILF changes
Process observations

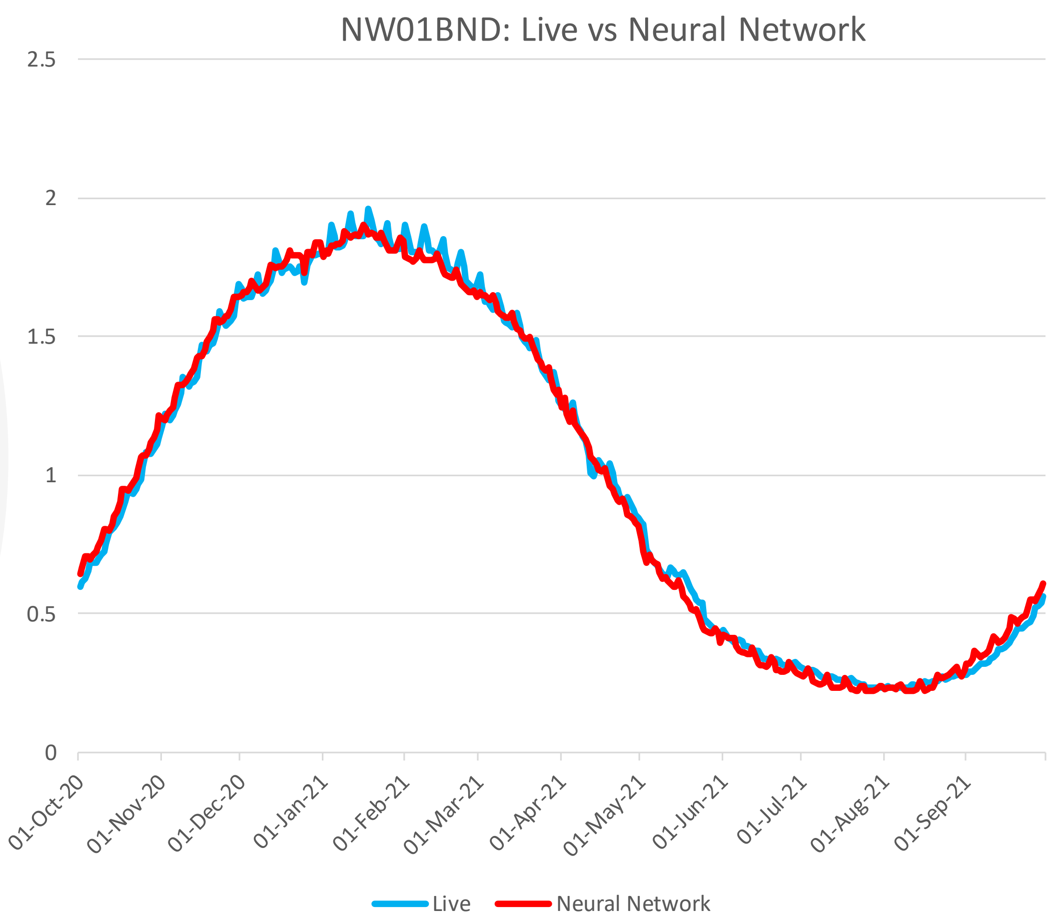
Neural Networks

We noted in meeting 4 that the results we were getting for Neural Networks were quite poor compared to other methods.

This was despite Neural Networks producing some promising results in the task force analysis (although their target measure was UIG).

- The chart on the right shows the ALP results from the Neural Network Machine learning using the Generalised Linear Modelling method.

The Neural Network method has produced a much smoother ALP than the Gradient Boosting method with fewer unusual spikes.



DAF Calculation

We noted in meeting 4 that to date we had been unable to calculate the Weather Sensitivity (WSENs) and therefore the DAFs required to test the accuracy of the new models against the current algorithm.

We took away an action to bring progress on DAF development to meeting 5.

We have now been able to calculate the weather sensitivity using the machine learning output for Seasonal Normal CWV compared to actual CWV in a similar way to the current methodology.

The information on the right is a summary of the current calculation method from Section 3 and section 9 of the Algorithm Booklet.

The gas demand model for an EUC can be expressed as:

$$D_t = SND_t + WSENS_t * (CWV_t - SNCWV_t)$$

Which can be rewritten as

$$WSENS = (D_t - SND_t)/(CWV_t - SNCWV_t)$$

Where D_t is the machine learning demand forecast using the actual CWV.

The DAF_t shall be determined as:

$$\mathrm{DAF_t} = \frac{\mathrm{WVCE_t}}{\mathrm{SNDE_t}}$$

where

DAF_t is the DAF on day t;

$WVCE_t$ is the value of the Weather Variable Coefficient in the Demand Model for the EUC (i.e. the sensitivity to weather or WSENS);

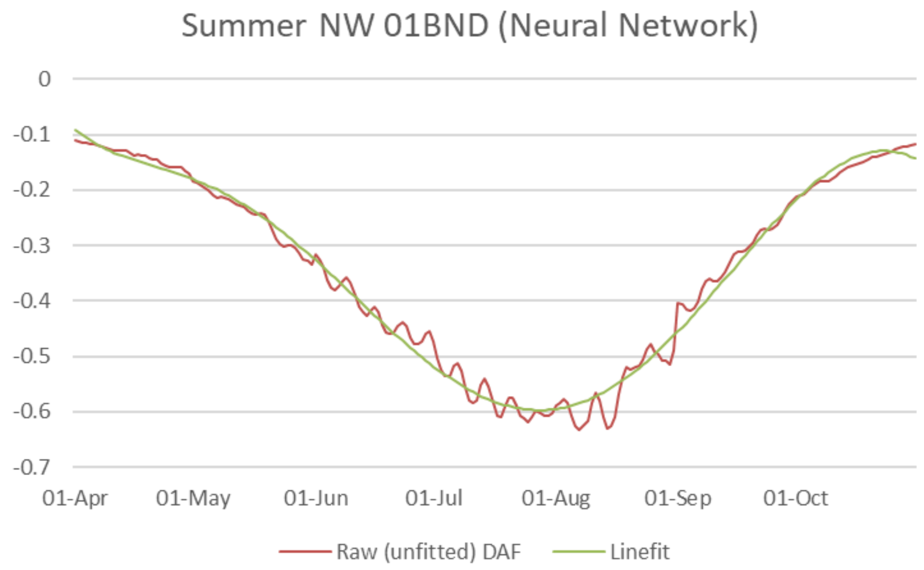
$SNDE_t$ is the value of the seasonal normal demand for the EUC;

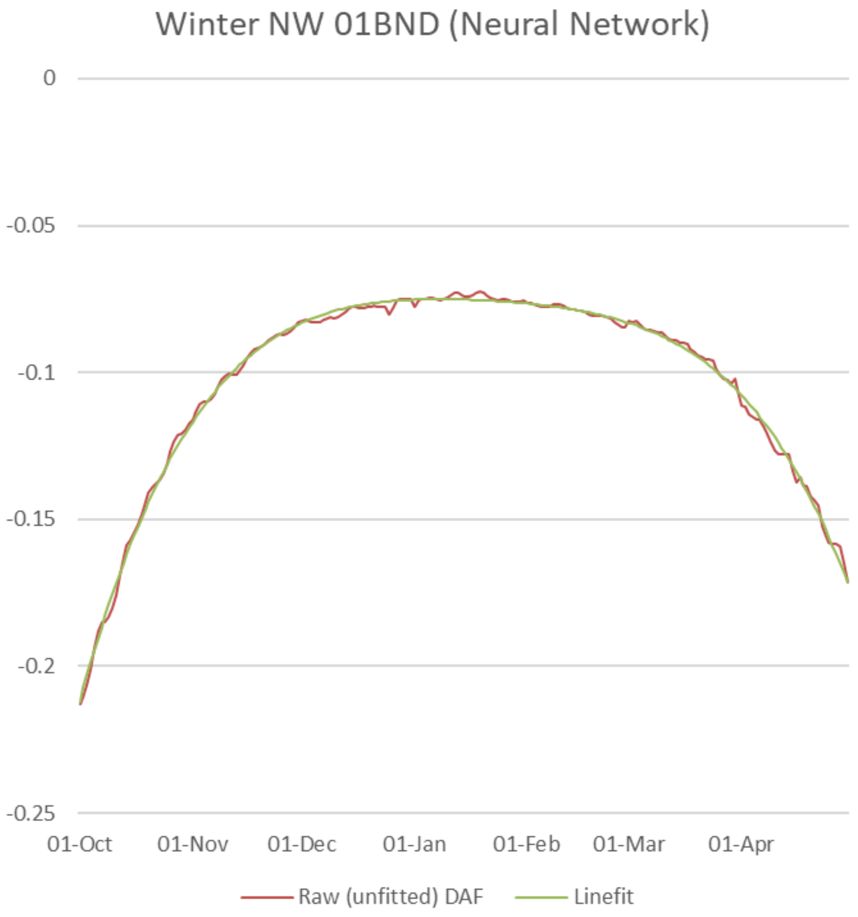
DAF Calculation

In order to get a formula for the DAF we have fitted a line to the calculated daily WSENS values seasonally (summer and Winter).

Shoulder months have been calculated with a blend of the two formulae.

The summer values show a clear difference in DAF for weekends and potentially days with a holiday code, which will be investigated.



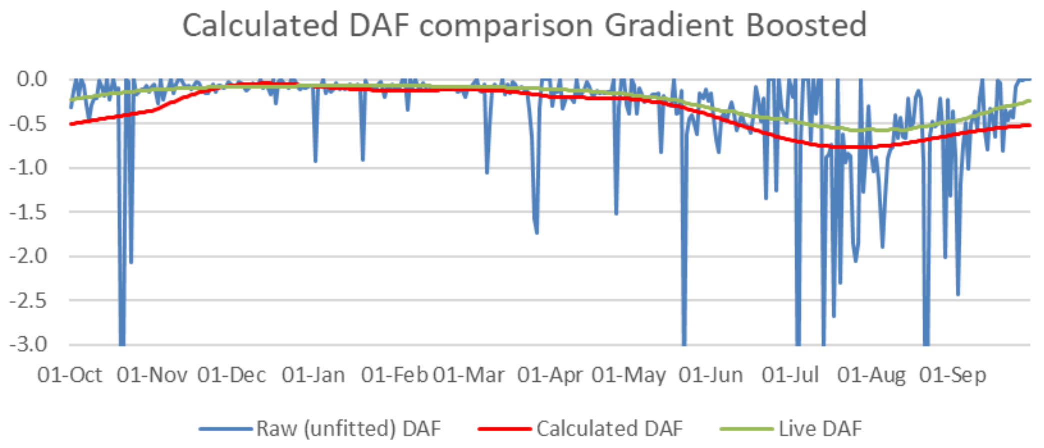


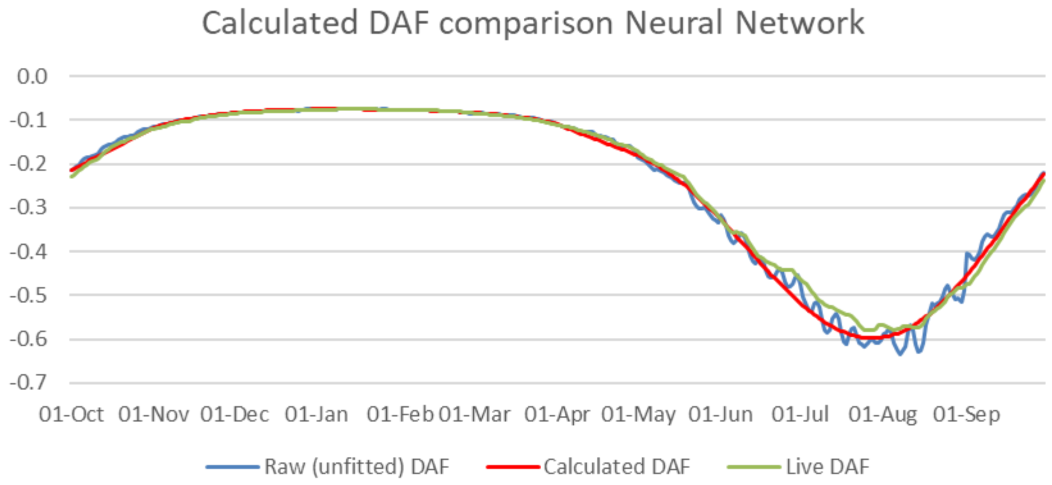
DAF Calculation

The calculated DAF has a similar shape to the live DAF, however the live DAF has slightly more shape to the summer months where as the calculation produces a smooth line.

As with the ALP calculation, there is a lot more noise in the Gradient Boosted DAF calculation, which the Neural Network machine learning methodology has been able to remove.

Live DAF shown is adjusted for new seasonal normal.





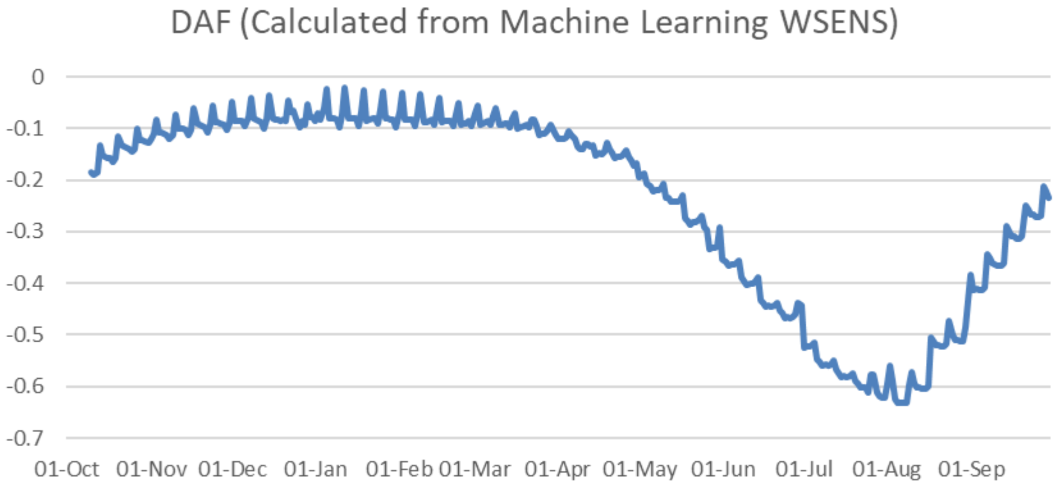
DAF Calculation

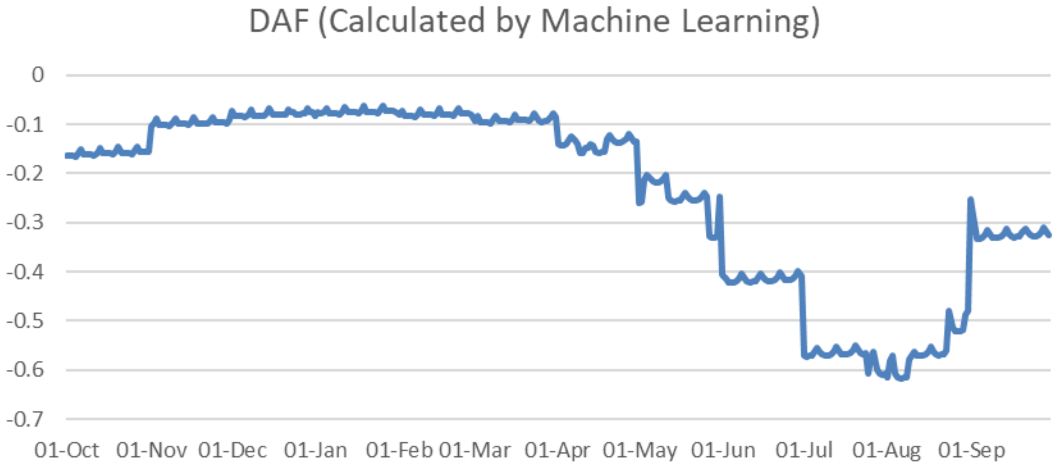
We have also looked at using Machine Learning to calculate the WSENS and DAF values.

The first chart shows the DAF where machine learning has been used to calculate the WSENS and then the DAF has been calculated using the formula on slide 17.

The second chart shows where machine learning has been used to calculate the DAF after the WSENS has been calculated (again using the formula on slide 17).

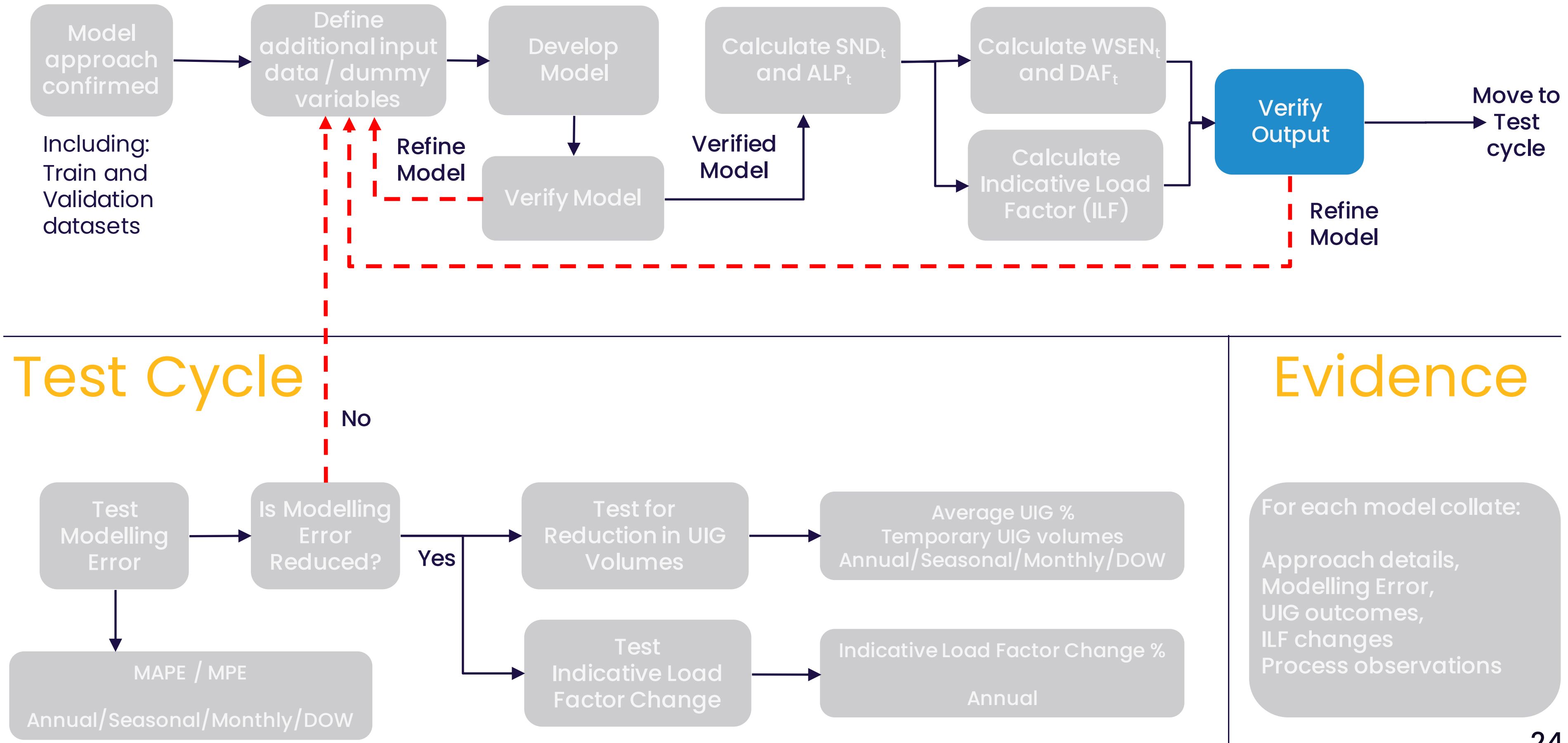
This has produced some interesting results which will be investigated further.





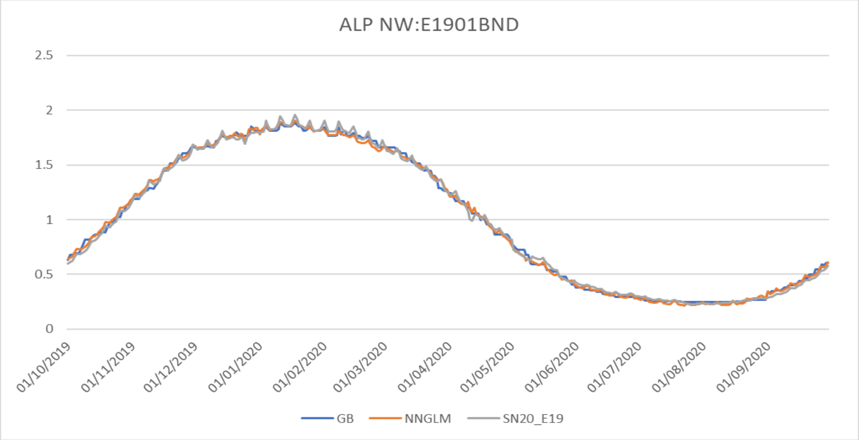
Verify Output

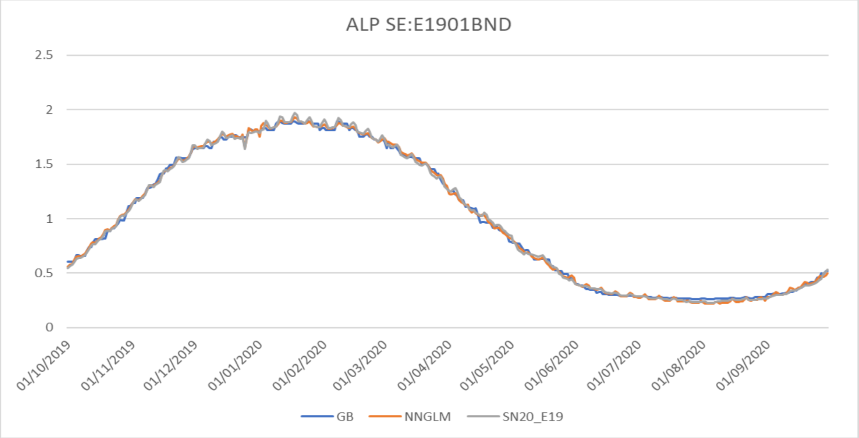
Development Cycle

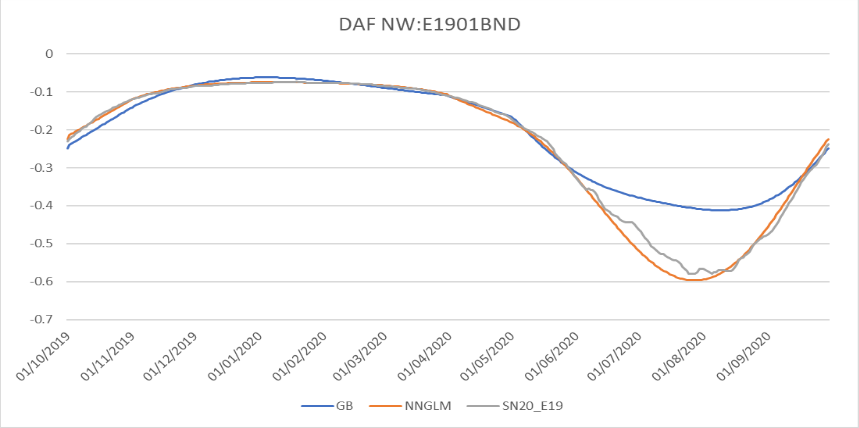


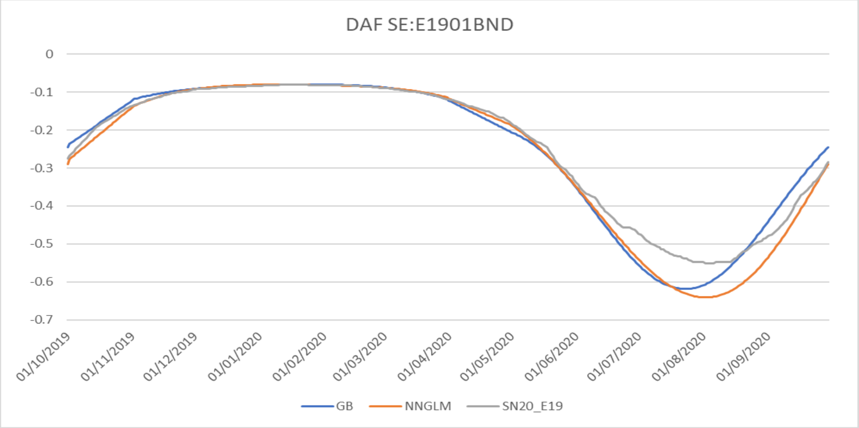
Verify Output

ALP and DAF Profile for 01BND

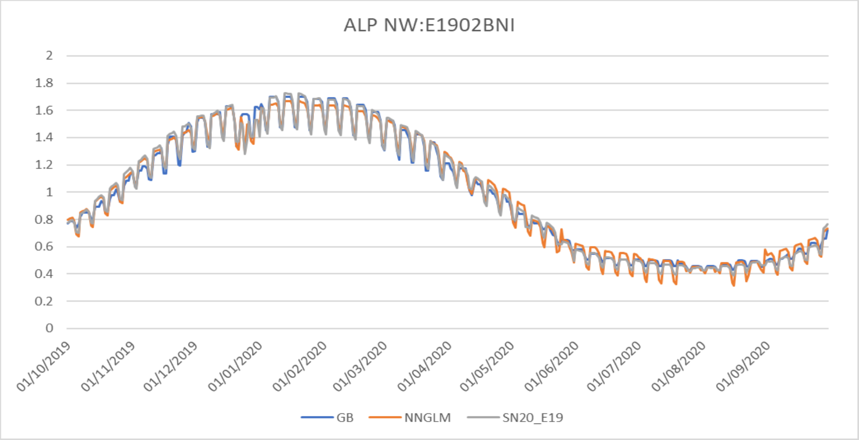


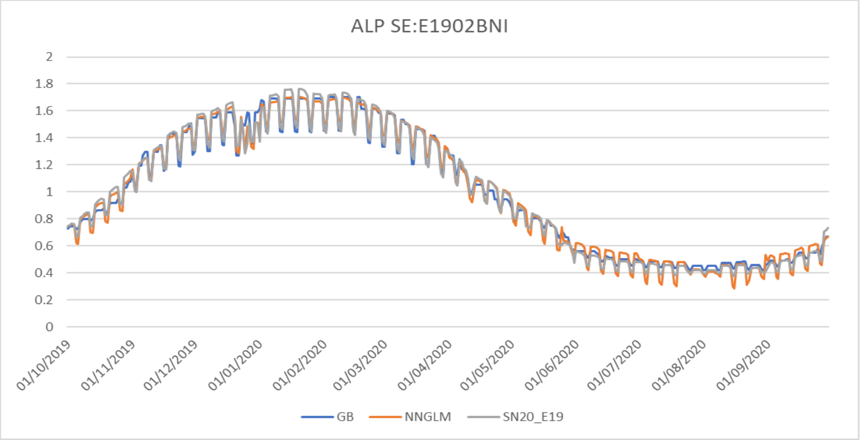


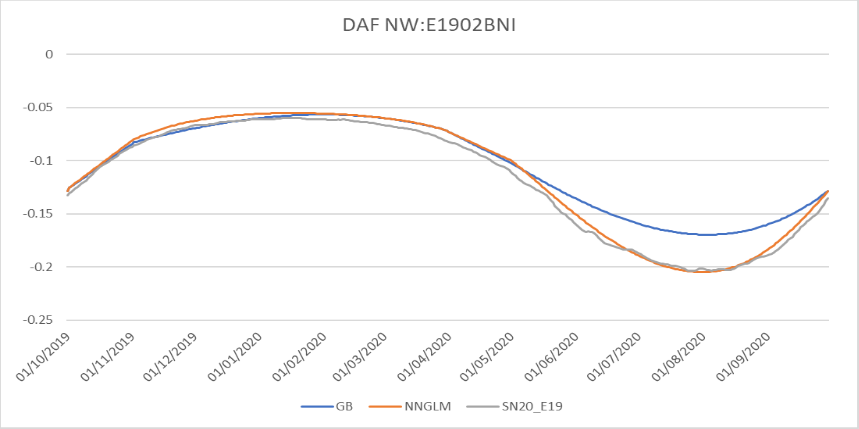


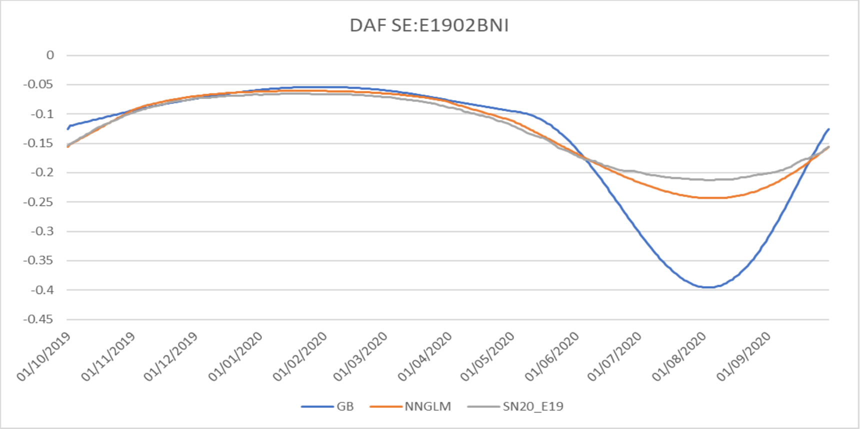


ALP and DAF Profile for 02BNI

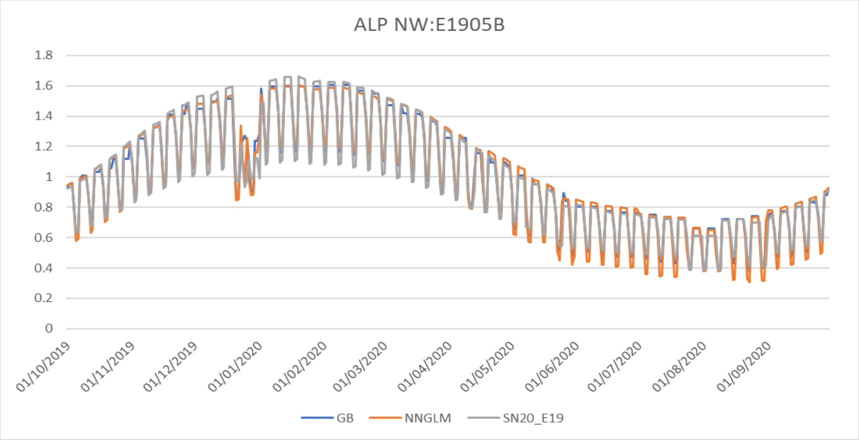


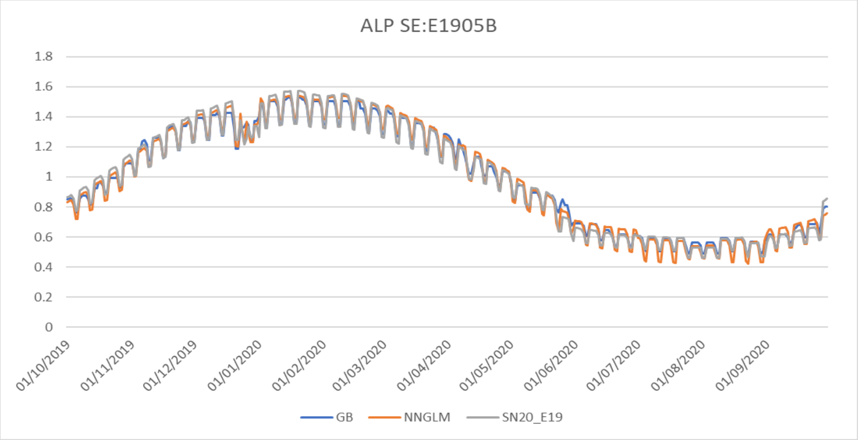


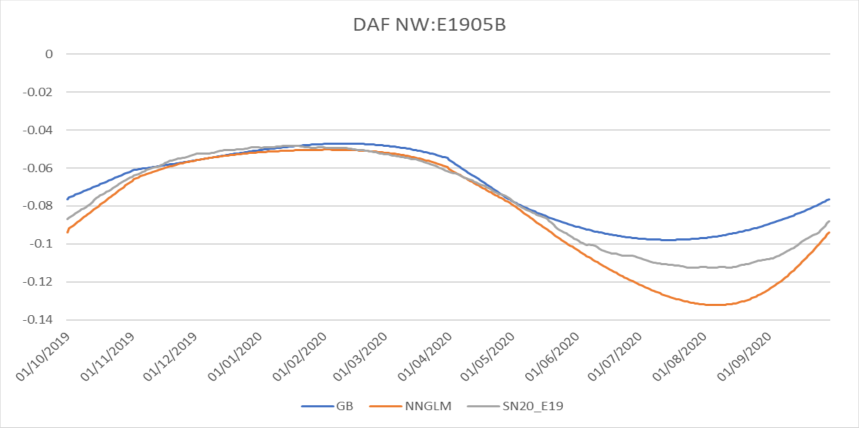


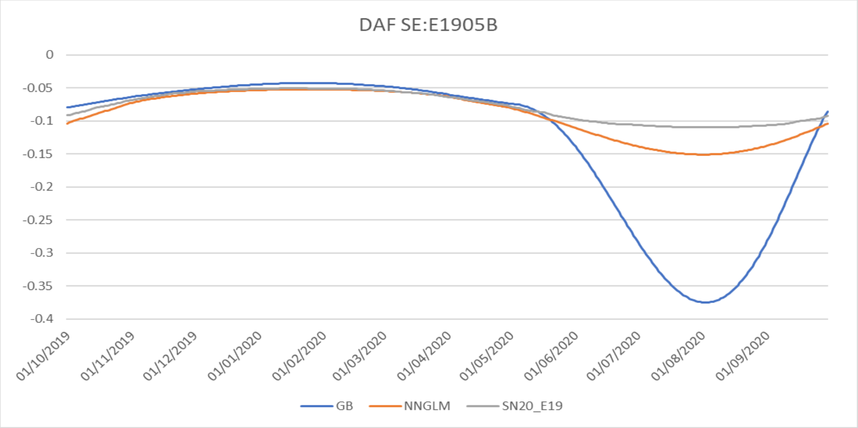


ALP and DAF Profile for 05B









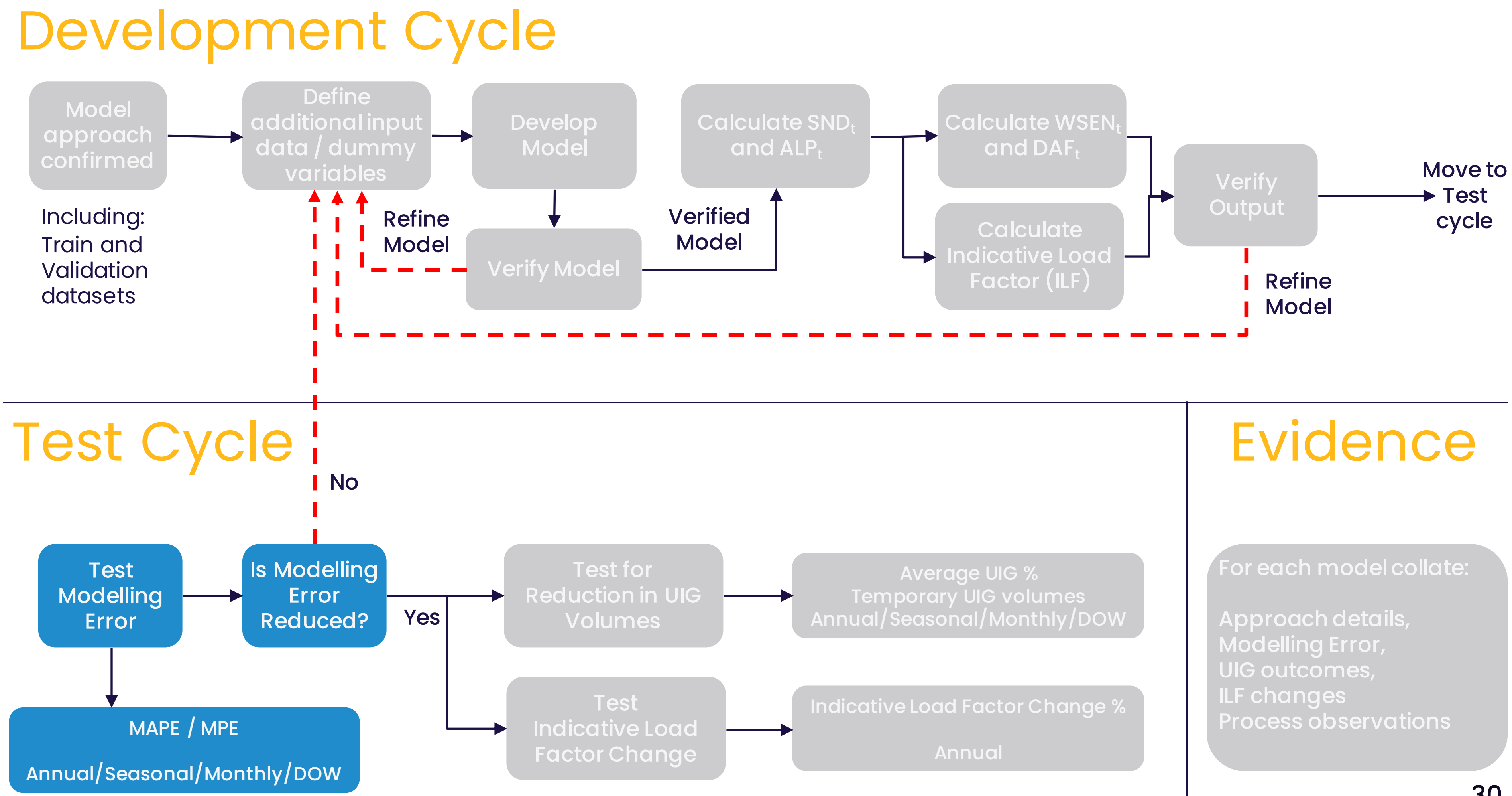
Initial ALPs and DAFs

A key stage in calculating a set of ALP and DAFs for the Test EUCs has been overcome.

The charts show:

- A consistency in the ALP profile between current modelling approach and the alternatives.
- The DAF is showing some variation, the full impact of this will be clearer once we review the modelling error in the Test Cycle phase

Test Cycle:
Test Modelling Error


Development Cycle
Model approach confirmed
Including: Train and Validation datasets
Define additional input data / dummy variables
Develop Model
Verify Model
Refine Model
Verified Model
Calculate SND_t and ALP_t
Calculate $WSEN_t$ and DAF_t
Calculate Indicative Load Factor (ILF)
Verify Output
Refine Model
Move to Test cycle
Test Cycle
No
Test Modelling Error
Is Modelling Error Reduced?
Yes
MAPE / MPE
Annual/Seasonal/Monthly/DOW
Test for Reduction in UIG Volumes
Average UIG %
Temporary UIG volumes
Annual/Seasonal/Monthly/DOW
Test Indicative Load Factor Change
Indicative Load Factor Change %
Annual
Evidence
For each model collate:
Approach details,
Modelling Error,
UIG outcomes,
ILF changes
Process observations

Assess Model Error

Testing the modelling error involves assessing the ALP and DAFs against sample data (Oct 2019 to September 2020). Using datasets collected for DESC's Algorithm Performance (Strand 3)

The profiles are in their first draft with further refinement necessary if modelling error is not reduced

This phase involves assessing Sample data against the following profiles:

- Current Approach ALPDAF
- Neural Network ALPDAF (GLM model)
- Gradient Boosting ALPDAF

Test Cycle

Initial MAPE 01BND

- Encouraging initial results with both machine learning models quite close to the current model
- Refining the ALP and DAF will hopefully improve this further

MAPE (Mean Absolute Percentage Error) Comparison

NW:E1901BND

	Summer	Winter	Full Year
Live Model	11.20%	4.05%	7.62%
Gradient Boosted	13.00%	4.10%	8.55%
Neural Network	12.37%	4.06%	8.22%

SE:E1901BND

	Summer	Winter	Full Year
Live Model	10.71%	3.58%	6.89%
Gradient Boosted	11.90%	3.60%	7.15%
Neural Network	11.72%	3.62%	7.32%

Initial MAPE 02BNI

- Note: These datasets have COVID impacted days between April 2020 to September 2020 which explains the poor percentages for all the models.
- The Gradient Boosted model is better than Neural Network for NW but not SE
- The live Model is still giving the best results for both areas

MAPE (Mean Absolute Percentage Error) Comparison

NW:E1902BNI

	Summer	Winter	Full Year
Live Model	32.64%	11.13%	21.89%
Gradient Boosted	34.02%	11.59%	22.80%
Neural Network	35.48%	11.77%	23.62%

SE:E1902BNI

	Summer	Winter	Full Year
Live Model	26.65%	7.94%	17.29%
Gradient Boosted	31.90%	8.84%	20.37%
Neural Network	29.95%	8.70%	19.33%

Initial MAPE 05B

- Note: These dataset have Covid impacted days between April 2020 to September 2020 which explains the poor percentages for all the models.
- The Neural Network model is quite close to the live model for both areas
- The Neural Network model is slightly better for NW and better for Summer in SE
- Gradient Boosted results were not as good

MAPE (Mean Absolute Percentage Error) Comparison

NW:E1905B

	Summer	Winter	Full Year
Live Model	24.14%	10.89%	17.52%
Gradient Boosted	23.39%	10.92%	17.16%
Neural Network	21.71%	10.61%	16.16%

SE:E1905B

	Summer	Winter	Full Year
Live Model	19.83%	6.62%	13.23%
Gradient Boosted	22.07%	7.59%	14.83%
Neural Network	19.77%	7.12%	13.44%

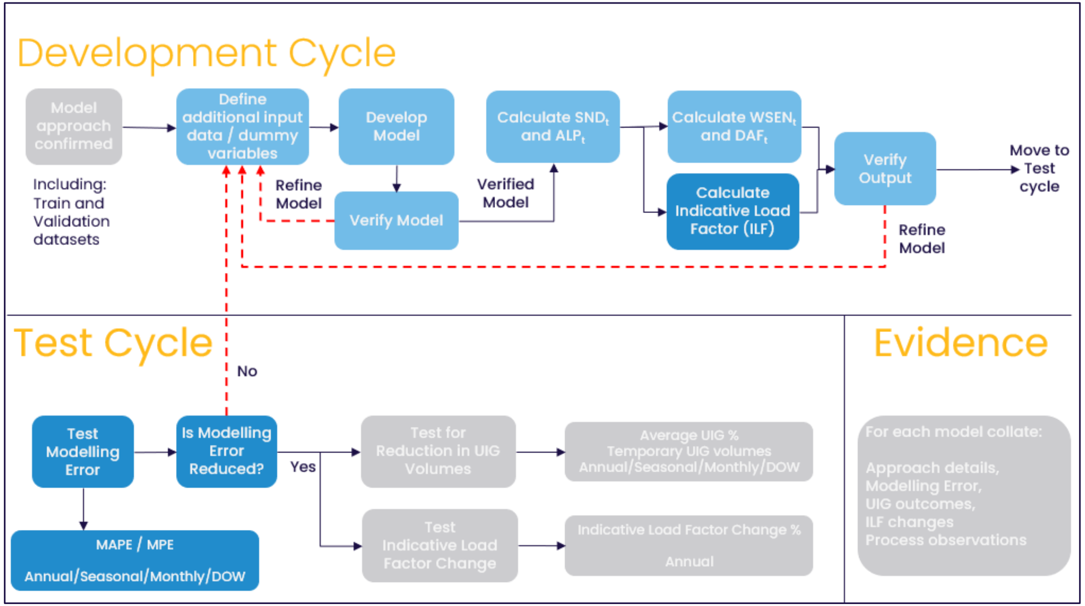
Conclusions and Next Steps

Conclusions

To date we have made good progress with calculating ALPs and DAFs and the results are fair compared to the live model.

Now we are getting good results from Neural Networks we can look at adding in additional factors and tweaking the modelling methodology to see if results can be improved.

This is now moving into the Test Cycle.



Next Steps:

- Deep diving into the output to see areas of success and areas for improvement, e.g. looking at individual months, weekdays, holiday codes
- Complete refinement of models for trial LDZs
- Calculation of Indicative Load Factors (ILF)
- Meeting 6 preparation

Timeline

Workgroup 0754R Timeline

March 2021

WG Meetings 1&2
- Scope/Objective
- High Level Principles
- Resources/Support
- Potential Areas to Investigate

- Top 3 Areas to Investigate
- Data Availability
- Resourcing / Costs

July 2021

WG Meeting 3 (7th)
- Area 1 Focus
- Measures/Success Criteria
- Timescales

Oct 2021

WG Meeting 4 (5th)
- Area 1 Progress - Initial Results

Nov 2021

WG Meeting 5 (30th)
- Develop DAF methodology
- Further refine models

Jan 2022

WG Meeting 6 (25th)
Area 1 progress
Area 2 Intro.

Mar 2022

WG Meeting 7 (22nd)
Area 1 conclusion

Nov 2022

WG Conclusion

Demand Estimation Team

Core Demand Modelling work will take place between April and June, 2022 for the production of Gas Demand Profiles for Gas Year 2022/23

Thank you

ON BEHALF OF xoserve