Modelling Results – 01BND

Action: TWG0501

Demand Estimation Sub Committee

Technical Workgroup

27/05/2021

Provided by:

Results – EUC 01BND – Introduction

At DESC TWG meeting on 24th May 2021, the following options were discussed in relation to which approach should be used when considering the most recent analysis period (20/21), in order to produce EUC gas demand models for EUC Band "01BND" for Gas Year 2021/22

- Option 1 –Use the '20/21 model data in this year's process for deriving smoothed demand models, along with 18/19 and 19/20

- Option 2 –Use the '20/21 model data but remove specific days where majority of LDZs displayed negative outliers in April 2020 (namely Gas Days - 15th and 20th April), along with 18/19 and 19/20

- Option 3 –Do not use 20/21 model data and refer back to last years smoothed demand models – this used data from 17/18, 18/19 and 19/20

DESC TWG decided that Option 2 would be the preference, with the caveat that further analysis will be done to investigate a 'hybrid' version which removes all gas days from April 2020 and assess the merit of using this updated model – for the purposes of this material, this has been labelled as "Option 2a"

The following slides provide the results for both Option 2 and Option 2a

Results - EUC 01BND Summary – Options

Where:

- Option 1 – Include all days

- Option 2 - Remove Gas Days 15/04/20 and 20/04/20

- Option 2a - Remove all April 2020

Indicative Load Factor:

- There are minimal differences between the ILFs on each variant of the model

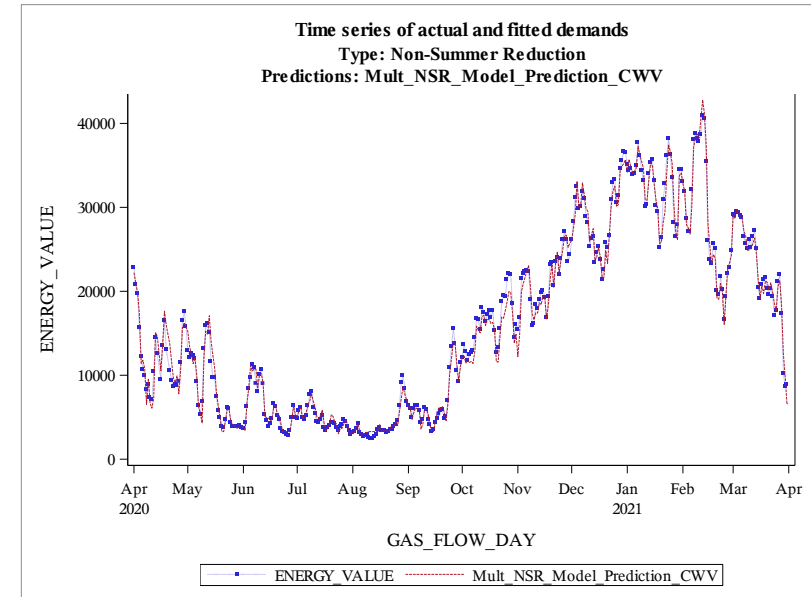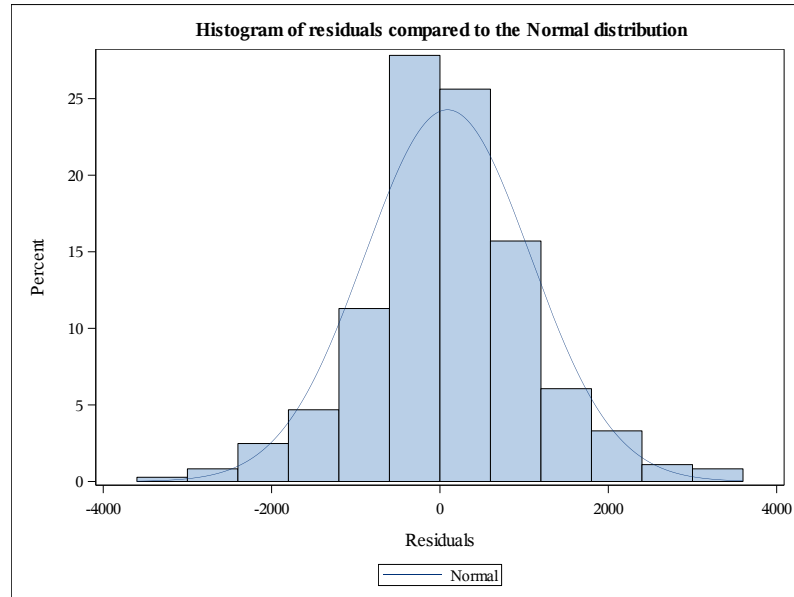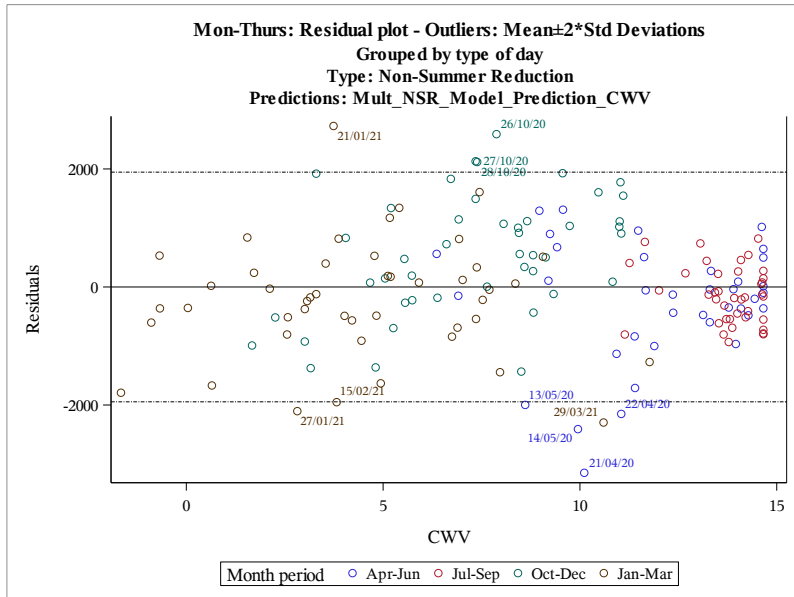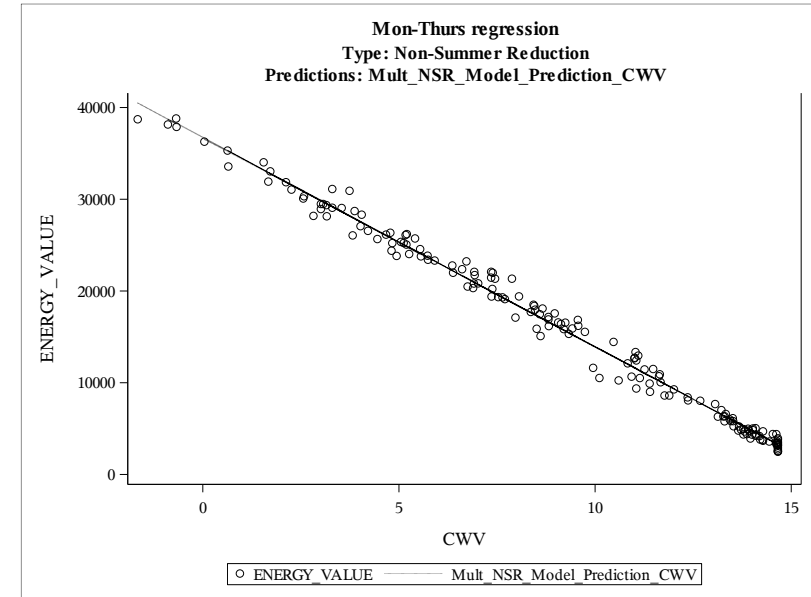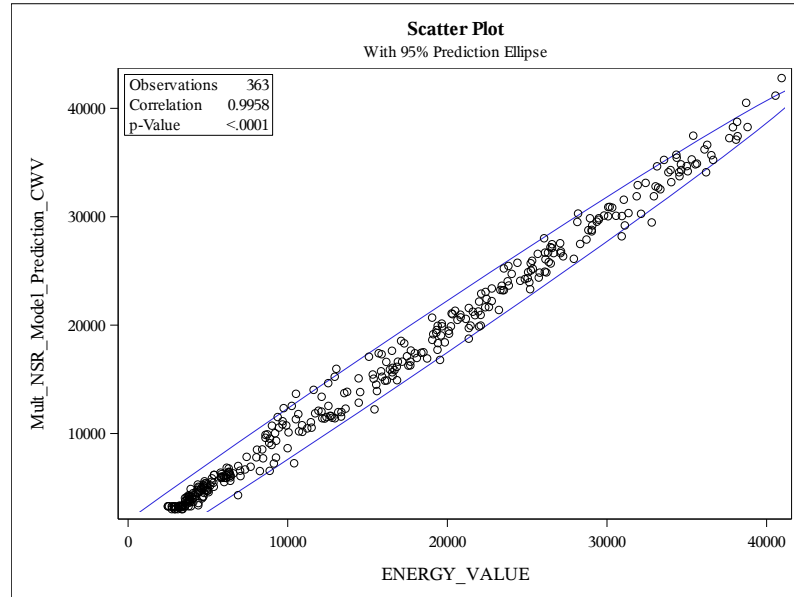**R^2 Results:**

- As expected, R-Squared values have slightly improved when removing outlier days in options 2 and 2a

The following slides show a view of this years model for LDZs EM, WN and SE (i.e. LDZs referenced in main material)

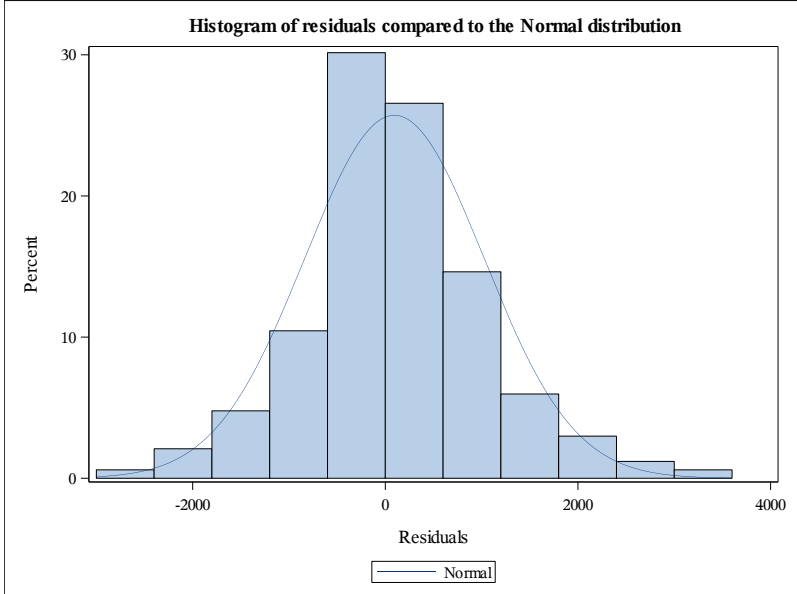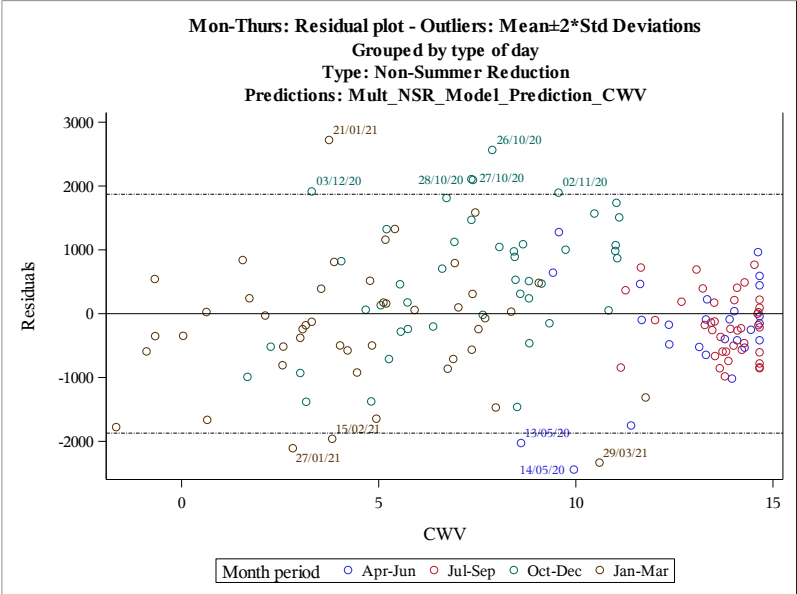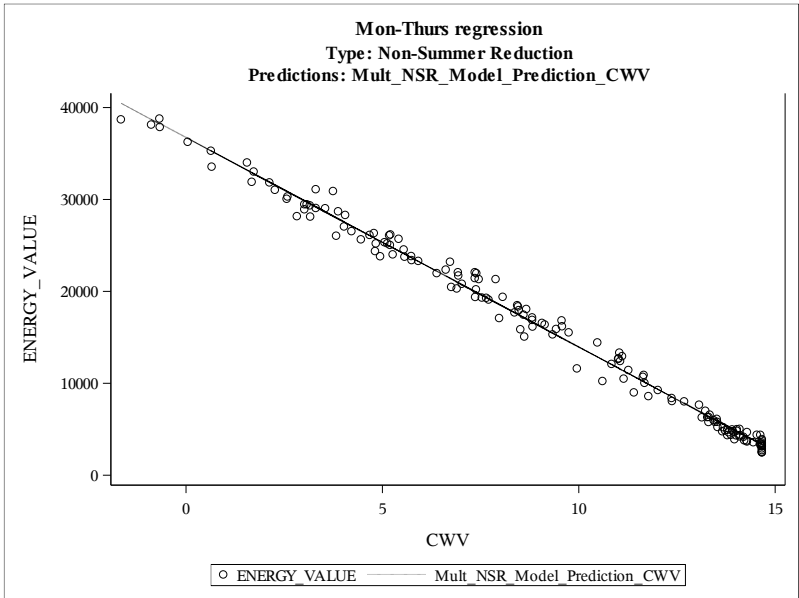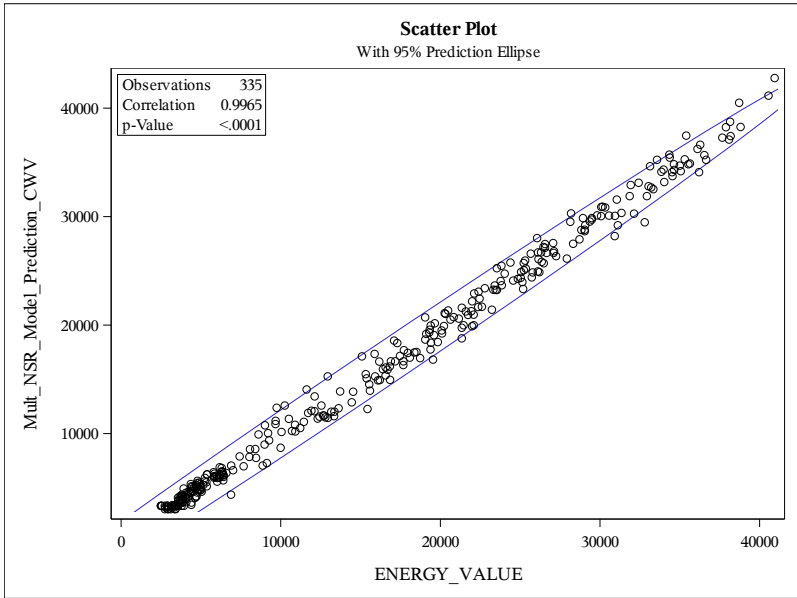| LDZ	Indicative Load Factor			R Squared Value		
	Option 1	Option 2	Option 2a	Option 1	Option 2	Option 2a
SC	35.2	35.2 ◄	35.1 ▼	98.9%	98.9% ◄	99.0% ▲
NO	35.6	35.7 ▲	35.7 ▲	98.3%	98.4% ▲	98.6% ▲
NW	32.3	32.4 ▲	32.6 ▲	98.3%	98.3% ◄	98.5% ▲
NE	34.4	34.3 ▼	34.4 ◄	98.4%	98.5% ▲	98.6% ▲
EM	31.6	31.7 ▲	31.7 ▲	99.0%	99.0% ◄	99.2% ▲
WM	30.9	30.9 ◄	31.0 ▲	99.0%	99.0% ◄	99.1% ▲
WN	31.8	31.8 ◄	31.9 ▲	97.5%	97.6% ▲	97.9% ▲
WS	31.1	31.2 ▲	31.1 ◄	98.5%	98.5% ◄	98.5% ◄
EA	31.5	31.5 ◄	31.6 ▲	98.5%	98.5% ◄	98.6% ▲
NT	32.2	32.2 ◄	32.3 ▲	98.9%	99.0%	99.1% ▲
SE	30.8	30.8 ◄	30.6 ▼	98.8%	98.8% ◄	98.9% ▲
SO	28.3	28.3 ◄	28.5 ▲	98.6%	98.6% ◄	98.7% ▲
SW	29.6	29.6 ◄	29.6 ◄	98.4%	98.4% ◄	98.4% ◄

Results – EUC 01BND – Charts for LDZ EM – Option 2

Model:
No Summer Reduction
EUC: 01BND
LDZ: EM
Demand: EM
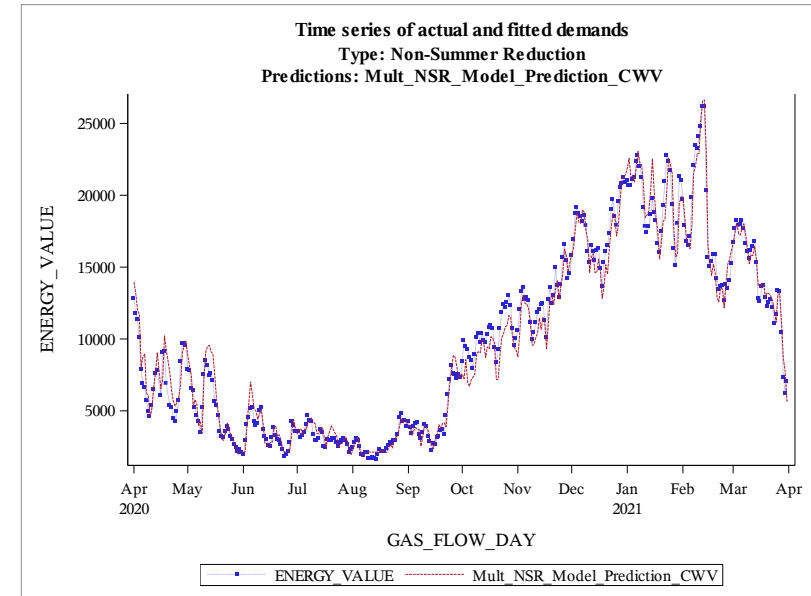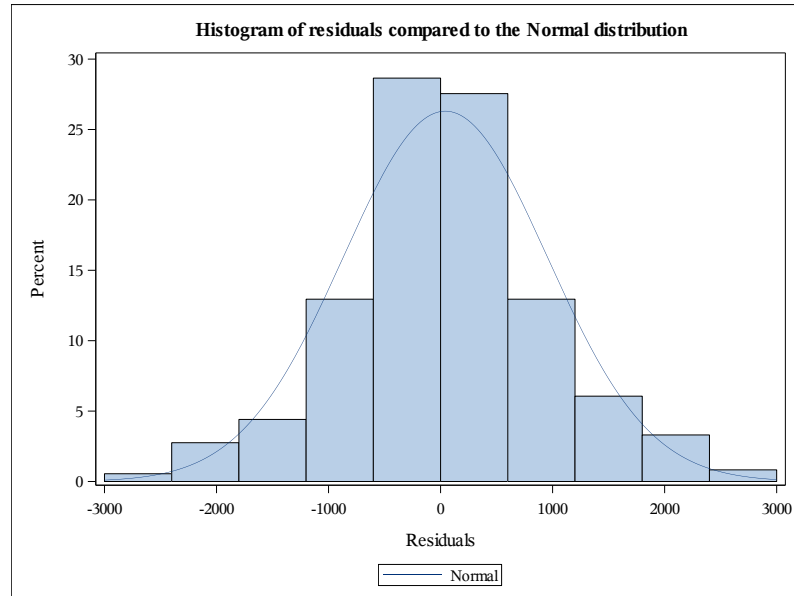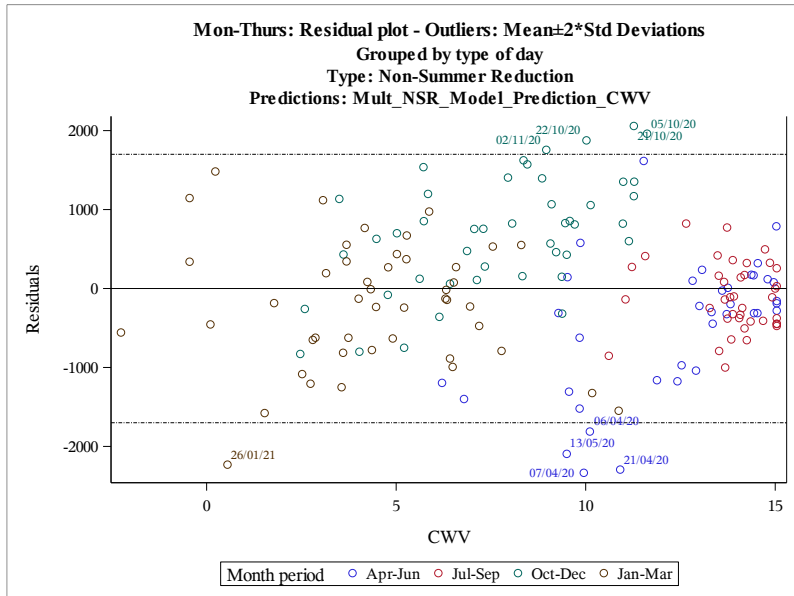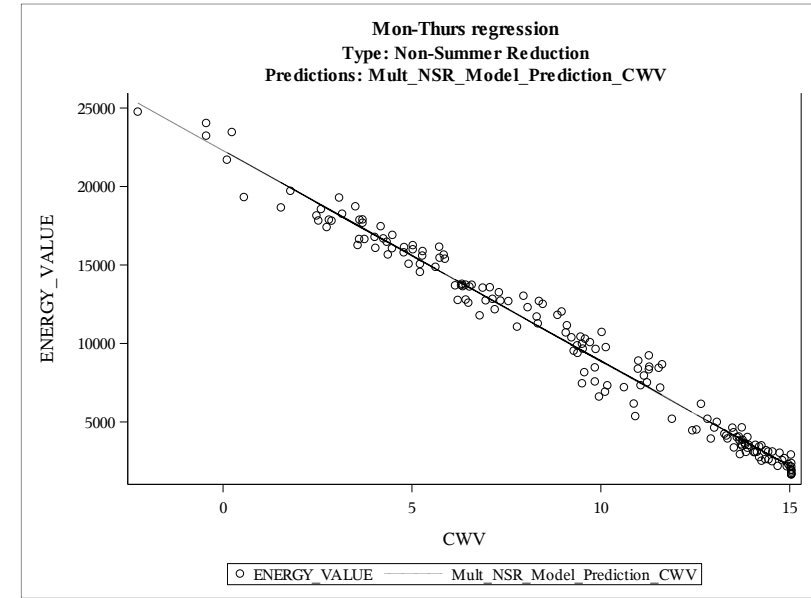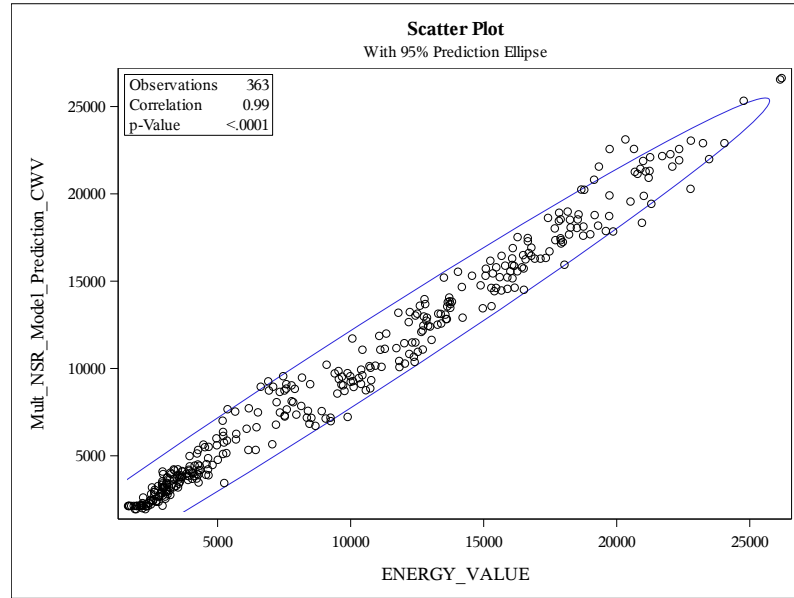R2 = 99.0%
ILF = 31.7
Sample Points = 385

Results – EUC 01BND – Charts for LDZ EM – Option 2a

Model:
No Summer Reduction
EUC: 01BND
LDZ: EM
Demand: EM
R2 = 99.2%
ILF = 31.7
Sample Points = 385

Scatter Plot
With 95% Prediction Ellipse

Observations	335
Correlation	0.9965
p-Value	<.0001

Mon-Thurs regression
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Mon-Thurs: Residual plot - Outliers: Mean±2*Std Deviations
Grouped by type of day
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Histogram of residuals compared to the Normal distribution

Time series of actual and fitted demands
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Results – EUC 01BND – Charts for LDZ WN – Option 2

Model:
No Summer Reduction
EUC: 01BND
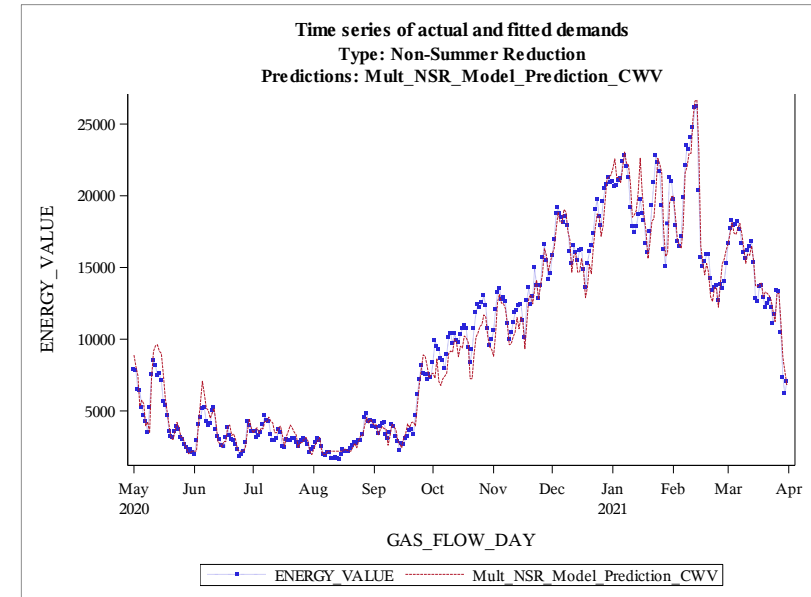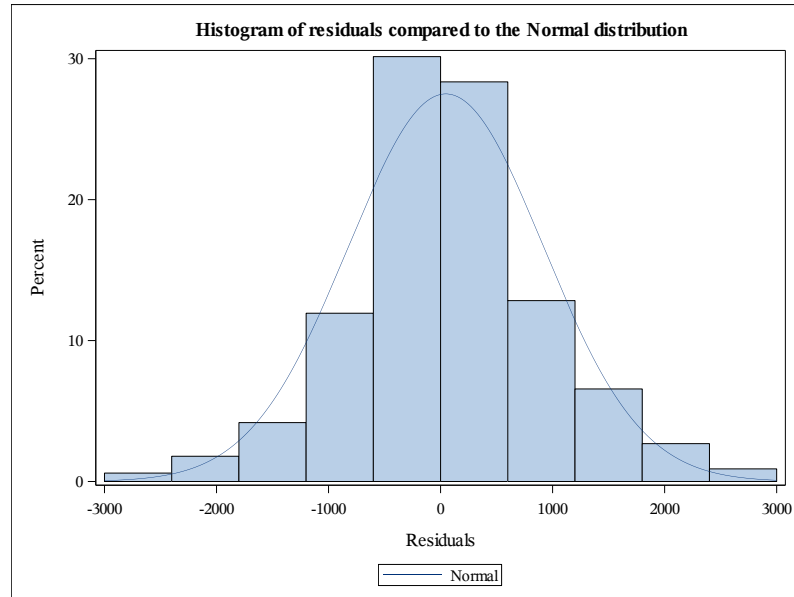LDZ: WN
Demand: WN
R2 = 97.6%
ILF = 31.8
Sample Points = 385

Scatter Plot
With 95% Prediction Ellipse

Observations	363
Correlation	0.99
p-Value	<.0001

Mon-Thurs regression
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Mon-Thurs: Residual plot - Outliers: Mean±2*Std Deviations
Grouped by type of day
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Histogram of residuals compared to the Normal distribution

Time series of actual and fitted demands
Type: Non-Summer Reduction
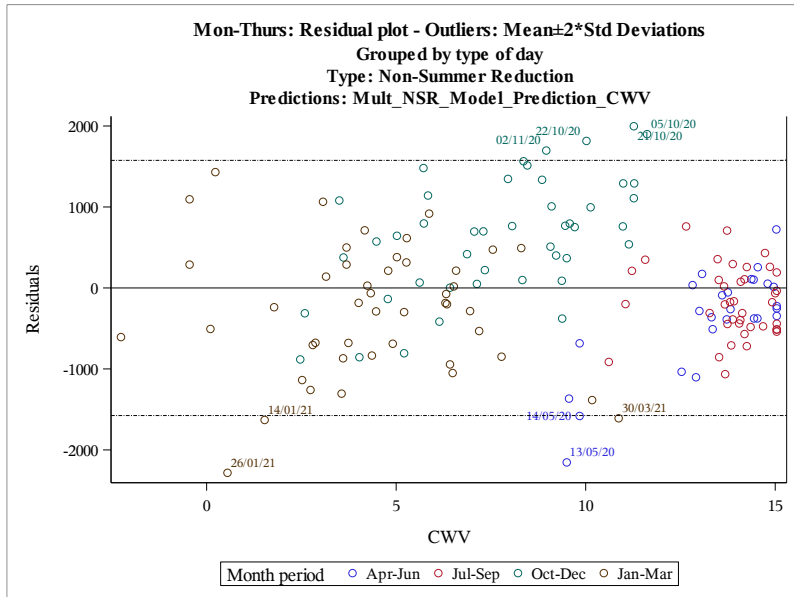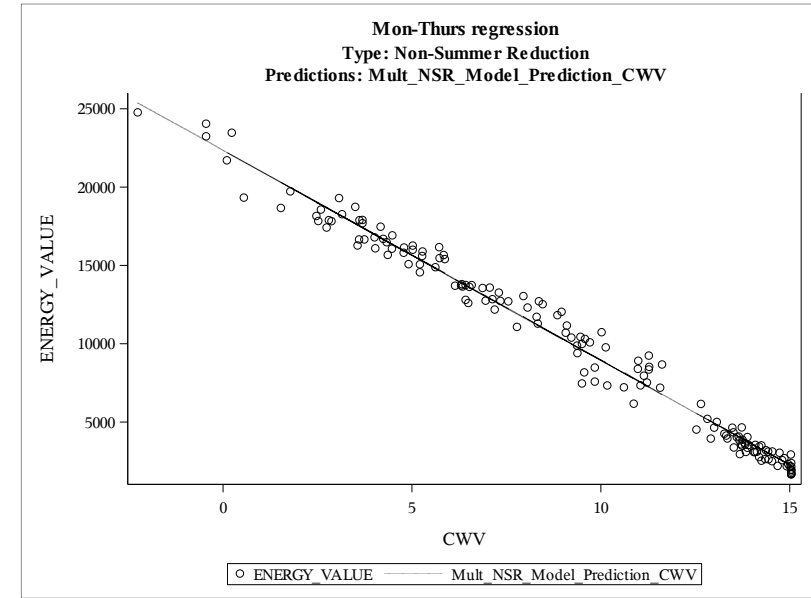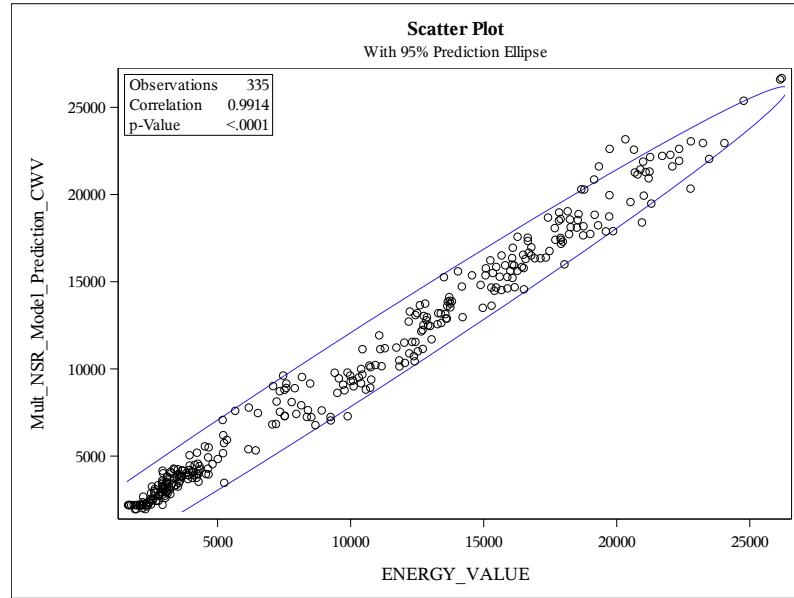**Predictions: Mult_NSR_Model_Prediction_CWV**

Results – EUC 01BND – Charts for LDZ WN – Option 2a

Model: No Summer Reduction
EUC: 01BND
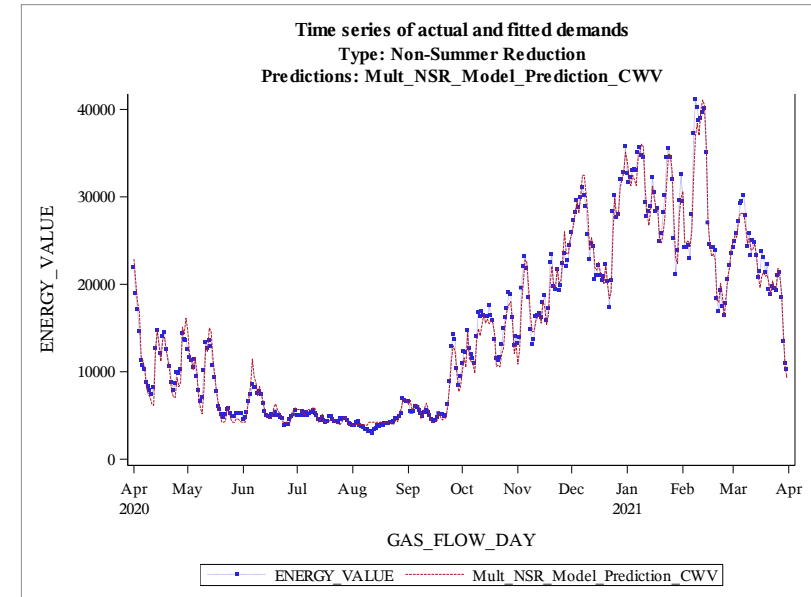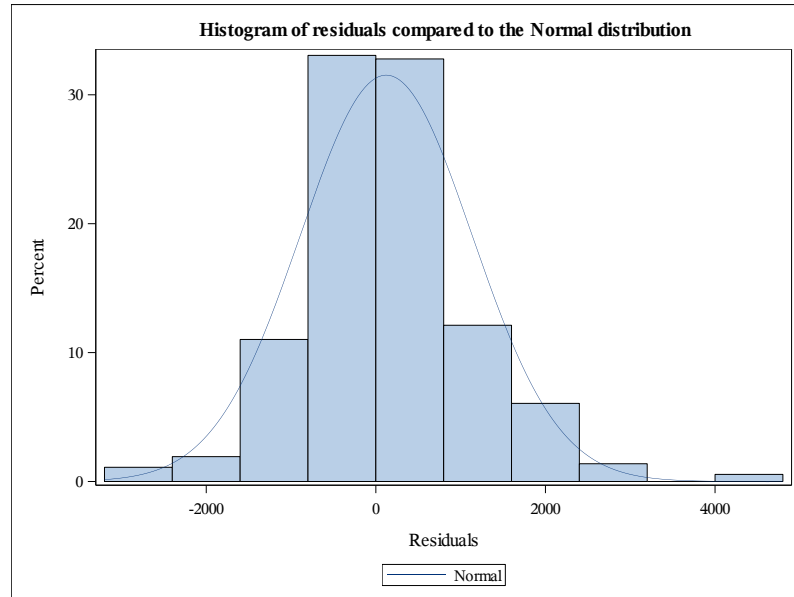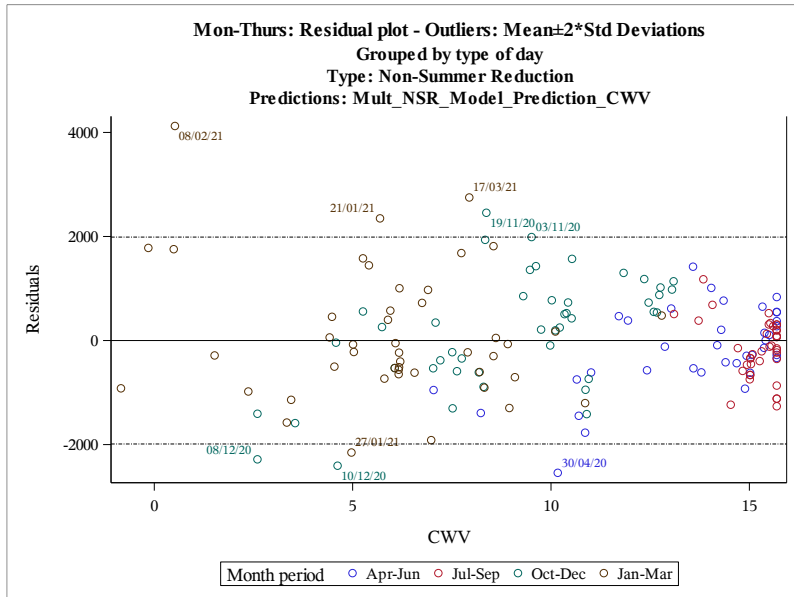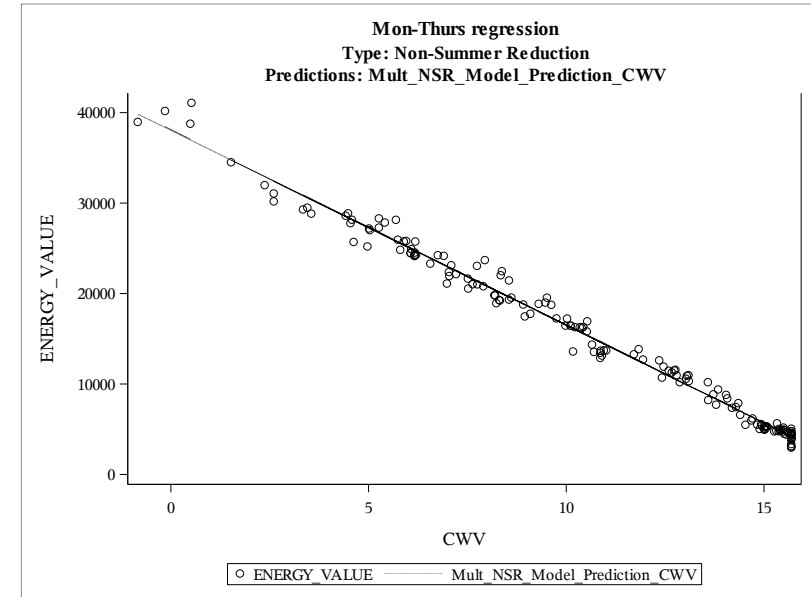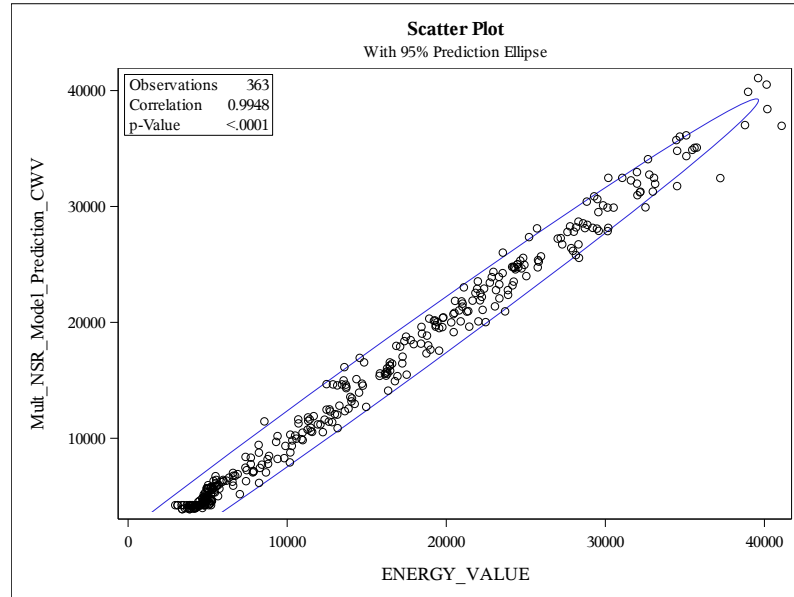LDZ: WN
Demand: WN
R2 = 97.9%
ILF = 31.9
Sample Points = 274

Scatter Plot
With 95% Prediction Ellipse

Observations	335
Correlation	0.9914
p-Value	<.0001

Mon-Thurs regression
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Mon-Thurs: Residual plot - Outliers: Mean±2*Std Deviations
Grouped by type of day
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Histogram of residuals compared to the Normal distribution

Time series of actual and fitted demands
Type: Non-Summer Reduction
Predictions: Mult_NSR_Model_Prediction_CWV

Model:
No Summer Reduction
EUC: 01BND
LDZ: SE
Demand: SE
R2 = 98.8%
ILF = 30.8
Sample Points = 385

Model:
No Summer Reduction
EUC: 01BND
LDZ: SE
Demand: SE
R2 = 98.6%
ILF = 30.6
Sample Points = 385
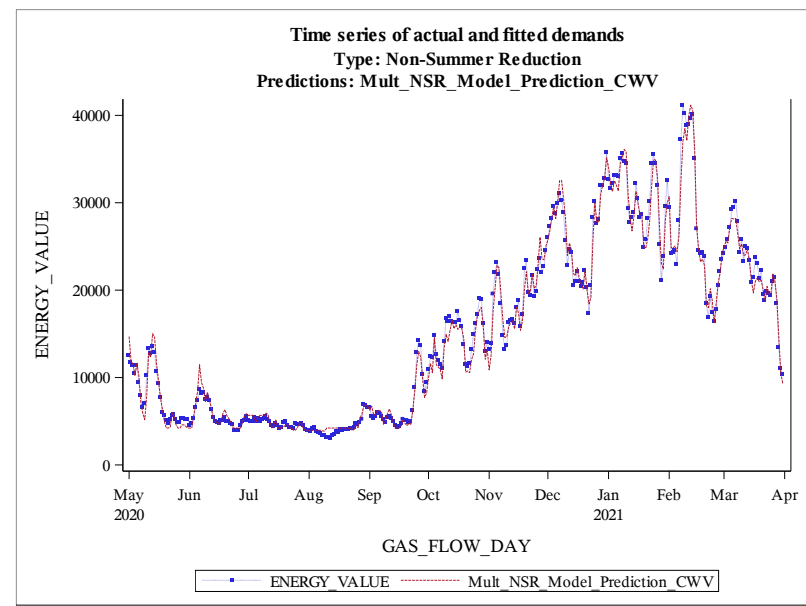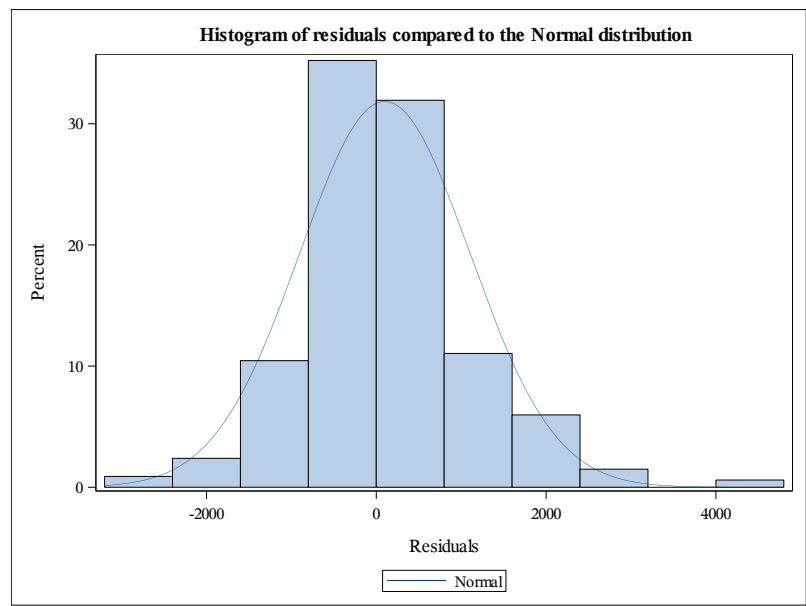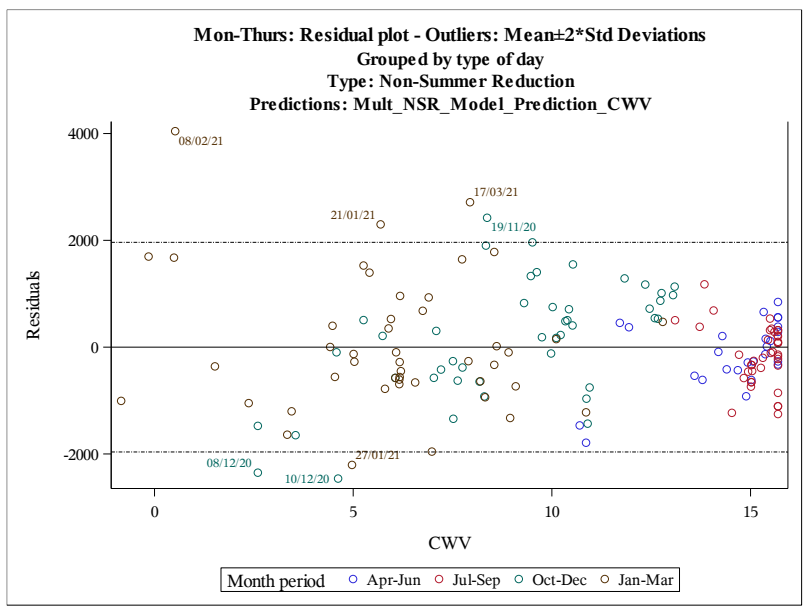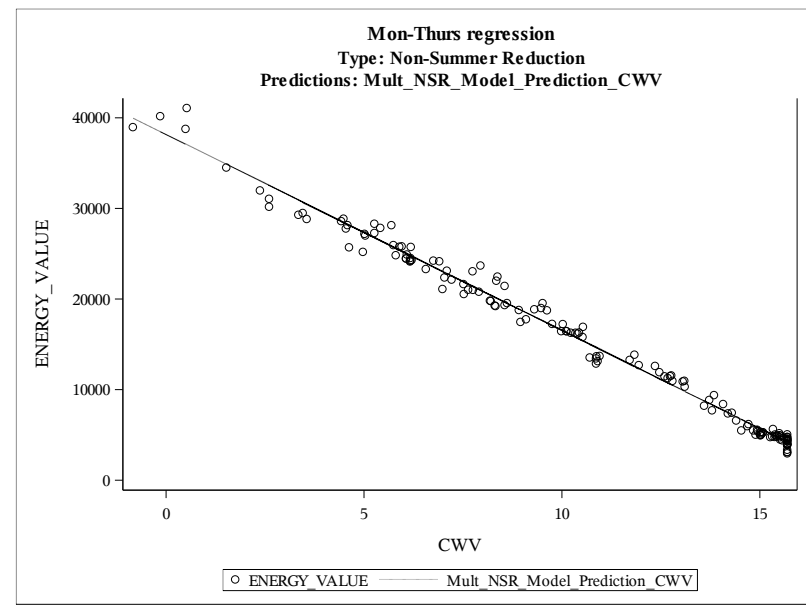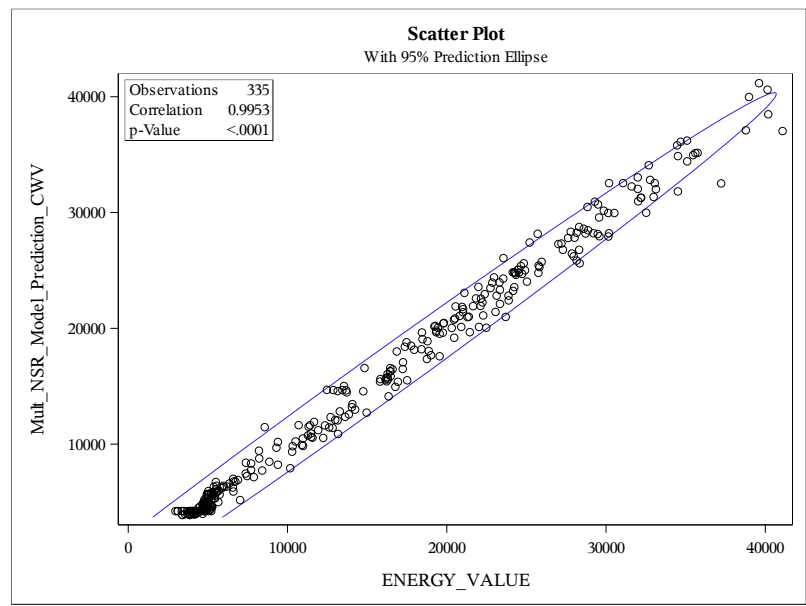
EUC 01BND – Conclusions & Recommendations

Conclusions:
- As discussed during DESC TWG meeting, April 2020 was an unusually warm and sunny month. Analysis as part of Algorithm Performance presented to DESC in December 2020 showed April to be the 3rd warmest April in the past 50 years in terms of GB CWV values. April 2020 has also been described provisionally as the 'Sunniest April on record' by the Met office

- This context helps to explain why April 2020 contained more outliers than other months and suggests they are not as a result of COVID-19 impacts but weather related

- Removing 15th and 20th April has been based on a criteria of these 'outlier days' being present in the majority of LDZs. After this, the outliers in April and other months are more variable across LDZs, however we are confident, following our data validation routines, that these are 'real' demands that should be used in informing the 20/21 demand models and we therefore feel the removal of all of April 2020 is not justified

- The removal of more and more outliers inevitably leads to an improvement in R^2 values but does mean different data points then become outliers and so the cycle continues

Recommendation:
- Option 2 is taken forward i.e. gas days where a majority of LDZs (7+) are showing as outliers are removed from the population, this is limited to gas days 15/04/2020 and 20/04/2020

- Models using the remainder of data from Analysis year 2020/21 are used to produce the smoothed models in conjunction with data from analysis years 2018/19 and 2019/20

- When reviewing the performance of the CWV in the future, particularly the new Solar Radiation term, data from April 2020 can be a useful dataset to inform any future amendments to the optimised parameters