UNC Workgroup 0754R Minutes
Investigate Advanced Analytic Options to improve NDM Demand Modelling

Tuesday 22 March 2022

via Microsoft Teams

Attendees,

Rebecca Hailes (Chair)	(RH)	Joint Office
Maitrayee Bhowmick-Jewkes (Secretary)	(MBJ)	Joint Office
Chris Syrett	(CS)	E.ON
Ellie Rogers	(ER)	Xoserve
James Doyle	(JD)	Foxglove Energy Supplies
John Jones	(JJ)	ScottishPower
Jordan Bignell	(JB)	Correla on behalf of Xoserve
Joseph Lloyd	(JL)	Correla on behalf of Xoserve
Katherine Uzzell	(KU)	SSE
Luke Reeves	(LR)	EDF Energy
Mark Perry	(MP)	Correla on behalf of Xoserve
Penny Griffiths	(PGr)	Correla on behalf of Xoserve
Sarah Palmer	(SP)	E.ON

Copies of all papers are available at: https://www.gasgovernance.co.uk/0754/220322

The Workgroup Report is due to be presented at the UNC Modification Panel by 17 November 2022.

1.0 Introduction and Status Review

Rebecca Hailes (RH) welcomed all to the Workgroup.

1.1. Approval of Minutes (30 November 2021)

The minutes from the previous meeting were approved.

1.2. Approval of Late Papers

No late papers were submitted. RH thanked Xoserve.

1.3. Review of Outstanding Actions

1101: Joint Office (MBJ) to add a reference to the UIG Taskforce findings to the standard Workgroup Agenda.
Update: RH advised the agenda has been updated and Review of UIG Taskforce findings has been added Workgroup Participants briefly discussed this area agreeing that individual Workgroup Participants may want to raise areas/queries arising from this document but that workgroup did not need to consult this unless requested. **Closed.**

1102: Xoserve (ER) to review the IP rights of the work done for this Workgroup and confirm this belonged to Xoserve.
Update: Ellie Rogers (ER) explained that clarity around whether Xoserve or Correla owned the IP rights for work carried out had already been provided in other Workgroups since this action was initially raised. ER added that in principle, IP rights were dependent on the funding route agreed for any work arising. For example, if the solution was funded by the

DSC, then Xoserve held the IP rights for those works, whilst for any solution funded by Correla, the rights resided with Correla.
ER noted that the DSC Change Management Committee would oversee the funding arrangements and make any necessary decisions.
RH stated that it was not clear yet who was the owner of the work being done under this Review Workgroup. ER advised that it was still unclear as a firm solution had not yet been developed and therefore it was difficult to identify how it would be funded. ER further noted that the work carried out so far had not required any additional funding.
Mark Perry (MP) explained that this Review Workgroup had been established prior to the Xoserve/Correla split and therefore no decision had been made on how it would continue to be funded in the future and the work being done is being scheduled around the core BAU work.
RH asked who would own the IP rights for this work if this Review Workgroup was established now. ER explained that the funding arrangement would have to be discussed at the DSC Change Committee.
RH advised that it was therefore necessary to get clarity on who owned the IP rights for the work currently being developed at the Review Workgroup and asked if this query could be escalated, and a confirmation could be received on this. ER noted that a legal view had been obtained but she would raise the query again for clarity.
Carried Forward.

1.4 Review of UIG Taskforce Findings

MP advised that the link to the UIG Taskforce Findings is in the slide pack and findings are there for the Workgroup to review.

2.0 Area 1: Trial alternative approaches to deriving SNDt

Jo Lloyd (JL) presented the Workgroup slides providing a high-level overview of what the discussions will cover.

JL provided a brief recap of the key discussion points from the previous meetings and what the focus of this meeting will be.

For a full and detailed update, please refer to the published slides on the meeting page.

2.1. Calculation of Indicative Load Factors (ILF)

JL presented a view of the approach taken to investigate, including reviewing ALPs, DAFs and load factors.

For a full and detailed update, please refer to the published slides on the meeting page.

The key points of discussions are captured below:

- Chris Syrett (CS) asked if there was a difference between the mean and median average for demand. JL was unsure of this. CS noted that it would be beneficial to know which side of the curve the mean and the median were on.
- JL accepted the query stating he would investigate this, adding that in terms of application the calculation has stood the test of time, but he would clarify what the value is.
- RH suggested that to confirm that using the mean was appropriate, it was necessary to know how the seasonal normal would work and the whole data set would need to be seen. CS agreed with this view.

New Action 0301: Correla (JL) to provide an indication of how the data set looks and clarify the difference between using the mean and the median figure.

- JL noted that the calculations had been performed using average AQs (Annual Quantity) available on the day and the Indicative Load Factor (ILF) values matched with the Peak Load Factor (PLF).

- CS and JL agreed to discuss this offline to further CS's understanding of the calculation methodology.

- JL noted an anomaly area had been identified in the results and investigation.

- Sarah Palmer (SP) asked whether utilising the data back to 1960 and there not being enough demand information was causing anomalies. SP added that the weather data could also be driving this.

- JL explained that the ILF was indicative and not produced in the same way as PLF. Scanning weather information from the 1960s helps in searching for lower Composite Weather Variable (CWV) each year and finding extremes. The PLF takes the SND_t and tests it against the CWV entered.

- SP noted that very cold temperatures would also be outside the calculations. JL accepted this.

- RH asked JL to clarify what the 'Gumbel Jenkinson' method was[1]. JL explained it was a calculation analysing probability extremes, noting that Weibull was an alternative, producing similar results.

2.2. Comparison with live models

JL presented live model findings with trends for comparison.

For a full and detailed update, please refer to the published slides on the meeting page.

The key points of discussions are captured below:

- JL stated that for all the models produced, the Annual Load Profile (ALPs) and Daily Adjustment Factor (DAFs) could be identified using the sample AQs.

- Katherine Uzzell (KU) asked if the models are based on the total throughput or the actual usage as the sample data is based on actual usage. JL confirmed a calculated AQ was used.

- RH asked if the data was NDM sample data. JL confirmed that it was adding that the approach and technique was an established way of analysing data.

- Day of week trend SE01BND: RH asked if the Saturday negatives were under allocation. JL confirmed this was the case.

- Day of week trend SE02BNI: SP questioned the data set used and how it was reflected in the modelling. JL agreed that the data may not be perfect and may or may not have been impacted by COVID. However, JL added that new sample data

[1] In probability theory and statistics, the Gumbel distribution (Generalized Extreme Value distribution Type-I) is used to model the distribution of the maximum (or the minimum) of a number of samples of various distributions. See:

https://en.wikipedia.org/wiki/Gumbel_distribution and

https://en.wikipedia.org/wiki/Generalized_extreme_value_distribution

would soon be available to use to train models that would address the current issues. SP accepted this.

- Holiday code (1-16 used) MPE02BNI: MP highlighted that the number of overallocations is large.

- Band 5 sites: JL stated that these are large industrial sites, typically with small population and the modelling was to gain an understanding of how new patterns effect these.

- KU asked for clarity on the influence of sample data on modelling. JL explained that when using Holiday Codes, using actual rather than predicted data was better.

2.3. Understanding principles of the models

JL presented an overview of the principles of modelling, noting that Penny Griffiths (PGr) has provided the slides on understanding the Neural Network models.

For a full and detailed update, please refer to the published slides on the meeting page.

The key points of discussions are captured below:

- RH asked for clarity on the graph influencing demand over the Christmas period. JL explained that the coloured part at the top of the graph shows the influences raising demand.

- JL explained that the dummy variable slices through data and then decides on influences.

- KU asked whether the modelling takes into account holiday codes or just the ALPs and DAFs. JL explained that it considered all core data.

- Luke Reeves (LR) asked whether large errors would still show dummy variables. JL advised that some minor influences can be identified whilst bigger impacts come from CWV.

- CS wondered how modelling might be impacted with the possible trend of people potentially moving to different working weeks (i.e., four day working week) and whether this has been considered in the modelling and how robust would it have to be to be built in.

- JL explained that modelling can include individual days of the week and some of the models work on aggregates. JL noted that it was important to get the best data from the samples.

- CS asked how organic the modelling was to react to these changes. MP advised that that impact on domestic modelling during COVID lockdown showed no major changes except in patterns as models allocating demand remained similar although the AQ changed.

- KU noted that it was interesting that the domestic model had not changed much. MP clarified that the consumption had changed slightly. SP stated that the first COVID lockdown was over summer so it had very little impact on heating and that by winter most of the population was back at work, which would explain the lack of change in the modelling.

3.0 Introduction to Area 2: Improve Validation Processes

JL confirmed the next area of discussion would be to improve the validation processes, noting that the models were only as good as the data being put in. JL added that at present whilst validation is carried out, it is not all encompassing, and sometimes other factors are included.

JL advised these discussions will consider the approach and technique of validation and will include reviews to identify how it can be improved.

JL presented the intended plan on how this will be achieved, including looking at techniques for identifying demand patterns.

It was noted that a Modification may need to be raised to support the work eventually coming out of this Review Group.

4.0 Next Steps

JL noted, the focus of the next Workgroup would be on progressing three areas of work in Area 1 and then moving to work on Area 2:

- Area 1: Investigate Peak Demand calculation for GB model
- Area 1: Investigate the Day of the week trends for the 02BNI and 05B datasets and test it against non-Covid datasets.
- Area 1: Try other dummy variables
- Area 2: Investigate methods to support validation identifying suspicious demand patterns

5.0 Any Other Business

None.

6.0 Diary Planning

Further details of planned meetings are available at: www.gasgovernance.co.uk/events-calendar/month

Workgroup meetings will take place as follows:

Time / Date	Venue	Programme
10.00 am Tuesday 28 June 2022	Microsoft Teams	Area 1 and Area 2 Progress
10:00 am late August 2022	TBC	TBC

Action Table (as at 22 March 2022)

Action Ref	Meeting Date	Minute Ref	Action	Owner	Target Date	Status Update
1101	30/11/21	1.3	Joint Office (MBJ) to add a reference to the UIG Taskforce findings to the Workgroup Agenda.	Joint Office (MBJ)	Next Workgroup	**Closed**
1102	30/11/21	6.0	Xoserve (ER) to review the IP rights of the work done for this Workgroup and confirm this belonged to Xoserve.	Xoserve (ER)	Next Workgroup	**Carried Forward**

Action Table (as at 22 March 2022)

Action Ref	Meeting Date	Minute Ref	Action	Owner	Target Date	Status Update
0301	22/03/22	2.1	Correla (JL) to provide an indication of how the data set looks like and clarify the difference between using the mean and the median figure.	Correla (JL)	Next Workgroup	**Pending**