Questions from NDM Algorithm Consultation Background Briefings

Please find below a record of the questions raised by the attendees at the background briefings held on 26th October and 3rd November 2020:

Q1. Will slides be available after the session?

All material relating to the NDM Algorithm Consultation, including Background Briefing slides and details of any questions raised will be published on dedicated page on DESC's area on the Joint Office website - link here

Q2. What are the impacts to customers of using Machine Learning, for example on Distribution Network's peak demands?

At this stage the consultation process is focussing on gathering industry views in order to establish what a future NDM demand estimation framework may look like (and equally not look like). This means there is no impact assessment available for any given option as this will come later. Any changes to the current regime would need to be proven to improve accuracy and be delivered in a framework that is supported by the industry. There are assumptions in the background document that point out which parameters / calculations would still be needed, which includes Peak Load Factors.

Q3. Will any subsequent follow up UNC Review Group (established after the consultation) be separate to Demand Estimation Sub-Committee (DESC)?

It would be for the Joint Office to decide but we would support/encourage that the UNC Review Group meets on the same day as DESC (before or after) to ensure there is continuity, review and scrutiny by DESC

Q4: Why would a UNC Review Group be required after the consultation? Why can't the consultation results be handed back to DESC to work through given they have the UNC obligation to review the NDM Algorithm?

It depends on the outcome of consultation. If there is clear appetite for significant change then establishing a UNC Review Group ensures that awareness and engagement is maintained throughout the development of any new approach and means, its conclusions (and any subsequent MOD proposal) are more likely to be supported. Clearly the expertise that DESC holds will also be needed as a key input to any industry discussions and recommendations.

Q5. What does LPA stand for? / Do you know where I can find more information on Load Profile Allocation data? I believe I have access to the files internally but it would be useful to know the definition of them, and how they can be used.

LPA stands for 'Load Profile Allocation'. This is a generic file (i.e. not customer-specific data) which is issued to Shippers after D+5 close-out for the Gas Day and includes a Volume Factor, Energy Factor and SAP Factor for every End User Category (EUC) for that Gas Day. Those factors represent the amount of volume (m3), energy (kWh) and financial value of gas (at System Average Price) allocated per 1 kWh of AQ in that EUC for the day. This can be used by Shippers to calculate the amount of energy which was allocated to them at meter point or EUC level for the day, or aggregated with other Gas Days for reconciliation purposes.

Please find attached link to the LPA file format here.

Q6. Will Machine Learning forecast NDM Energy for each EUC or MPRN?

The expectation is that the estimation of NDM energy will continue to be performed in aggregate rather than at meter point level. The aggregation currently achieved via End User Categories, which is composed of data items on the Supply Point Register (UK Link), is likely to continue to be used.

Q7. Will more frequent inputs be used?

It depends on the approach preferred by the industry. If the feedback from the consultation concluded that having a set of parameters ahead of the new gas year was important or at least some time in advance, then this would influence how frequent and timely the inputs used in the Demand Model deriving the NDM estimate could be. If this was deemed not to be an issue, then more frequent 'real-time' inputs into the Demand Model may be possible. The Machine Learning paper provided as a link in the Background document covers these points.

Q8. There is a heavy focus on a Machine Learning approach, but this is just a tool, isn't AQ still the most critical data item?

The AQ will continue to be a critical value in the accuracy (or not) of the NDM estimate. The submission of reads by the industry will therefore continue to be critical in ensuring the AQ reflects the current energy usage.

Q9. Will Machine Learning use 'live data' or still rely on the sample data?

For the foreseeable future it is envisaged that sample data as currently collected from Shippers, Xoserve's and DN's sample providers will continue to be used as the input. This will be until such time that alternative sources become both available and reliable.

Q10. How do we assess the financial benefits?

We would like parties to quantify the financial benefits to them of lower daily UIG (on average) but bearing in mind that NDM allocation is likely to increase. What are the financial benefits to you if there is less movement between NDM allocation and reconciliation?

Q11. With the current NDM Algorithm, assuming you have a view of likely weather, we're able to use the parameters, such as ALPs and DAFs, to predict gas demand a week in advance. Will Machine Learning also provide this option?

In the event that an approach which did not produce parameters ahead of time was preferred (e.g. full Machine Learning), then it is envisaged that part of the solution developed by the industry should include the ability to access a 'Sandbox' area where customers could 'plug-in' details for their portfolio ahead of the gas day in order to understand the forecast NDM demand, and model their own weather scenarios. We would expect this requirement to be discussed in more detail as any potential solution is developed.

Q12. For the more advanced machine learning proposals (where the current NDM algorithm is replaced), would this be using machine learning to help create a new model to replace the existing one, or using an actual machine learning model in place of the existing one?

The Machine learning paper referenced in the Background briefing document looks at both of these options. At this stage there is no preference from an CDSP perspective, we have set out the Pros and Cons as we see it. The consultation will hopefully help DESC work out which the industry would like to develop as a potential approach.

Q12a. The reason I ask the above question is that from a forecasting perspective it is quite important to know if the model is mechanistic or empirical, in addition to how often it is subject to change (something we are quite comfortable with currently is the annual provision of the ALP etc.).

For example, if I am trying to forecast a subset of customers for a year – am I trying to predict the outcome of say a neural network model? Where a) the relationships between inputs and output are not clearly defined (i.e. a black box as described in your presentation), and b) it may be re-trained frequently within the year, or even re-trained on the same data and produce a different output (due to things like randomly seeding starting weights between nodes in the model).

With both of these points I think what I am getting is that moving away from a mechanistic approach might make the model more accurate for historic allocations, but at a cost of transparency and determinism. Potentially for a supplier/shipper this makes forecasting, trading, pricing, and cost allocation more complex and less accurate over long time periods. Is there any compensation for this within the advanced machine learning proposals? i.e. provision of something akin to ALP factors

Your question raises some very valid points which could form the basis of your response to the different options set out by the Machine Learning options paper. Option 2 in the paper includes provision something akin to ALP Factors whilst Option 3 and 4 move away from that.

Q13. This leads me onto another question, if the models are trained more than annually I assume they will not be accounting for the allocations made by previous models? And thus if you were to use them to calculate the demand of a single meter point it would be possible to allocate more than it's AQ value to it for the year, which doesn't make sense to me. Apologies if I am misunderstanding the allocation process.

The current approach, once weather correction is applied, could allocate more or less energy than the AQ across a gas year. AQ represents a typical annual consumption under seasonal normal conditions whereas the allocation process has to incorporate actual weather conditions into the estimate. We would expect any future approach to follow similar principles to give realistic estimates.