UIG Task Force
Machine Learning Update

DESC 9th December 2019

What is Machine Learning?

- Machine Learning is a subset of artificial intelligence
- Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions without being explicitly programmed to perform the task
- For example, our analytics partner trained their Neural Network model on historic consumption and demand data, and it learned to predict gas demand

Machine Learning Benefits

- Machine Learning algorithms can identify relationships in data that might not be obvious
- The algorithms generally get better the more data they have to work with
- The algorithms learn from their own performance to make better predictions over time
- The model is generally used through a web service so we could make it available to customers

What do we want to do?

- The UIG Task force think the industry should consider moving from the current NDM allocation model towards Machine Learning based prediction

- As the custodians and experts of the NDM model, we want DESC’s input in to how the industry can move forward

Phase 1 Progress

- The UIG Task Force and the UIG Analytics Partner have demonstrated that using advanced Neural Network Machine Learning models can reduce Base Level UIG by up to 70% and Volatility by up to 30%
  - See Investigation Item 13.2.6: NDM Algorithm - Advanced Machine Learning
- We tested this on EUC 01 for 12 LDZs

Phase 2 Work & Findings to Date

A number of workstreams are now underway:

1. Develop improved estimation models for EUCs 2-8 to demonstrate whether the ML benefit can be seen across the whole market.
   - Transfer learning used to train existing model on EUC 02 sites
   - At this early stage, we see a marked reduction (c. 2.5%) on day-on-day volatility from using this new ML model on EUC 02 only.
   - The low number of Sample sites in EUCs 03 and above is a challenge. New modelling approach that allows smaller groups of sample sites to be dynamically aggregated for modelling is now underway.

Phase 2 Work & Findings to Date

1. Improvements to Modelling targeting volatility
   - Incorporating LDZ input energy as an input to the model does increase base UIG slightly but also reduces day on day volatility.
2. Sample Meter Point Uncertainty
   - A deep dive of the NDM sample has identified a small number of Meter Points that pass validation and so are used for modelling but have atypical demand patterns for their EUC.
   - Excluding these from the Machine Learning training set reduces volatility.
3. Physical Investigation
   - While in the early stages, other physical factors including pressure differentials and soil moisture could contribute to day on day volatility.

Upcoming Work

- Explore using the models and findings developed to date to look at building a UIG predictive model.
- Rerunning the Neural Network performance comparison against the existing NDM algorithm for more recent Gas Days.