UIG Task Force

Overview of Options for Use of Machine Learning

v1.0 21/02/2020

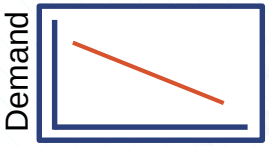
Background

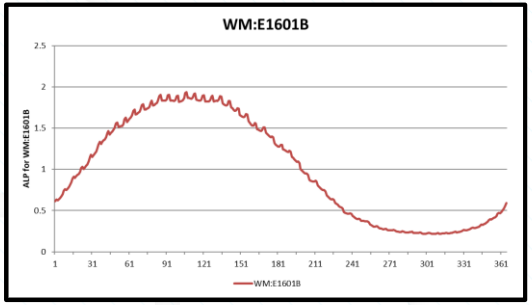
- The Unidentified Gas (UIG) Task Force undertook several phases of Machine Learning investigation
- Aims were to better understand drivers of UIG and identify options to reduce levels/volatility of daily UIG
- As the Balancing Figure in each LDZ each day, UIG is dependent on all the inputs to the calculation
- After initial investigations the focus of Machine Learning has been on improving the Non-Daily Metered (NDM) estimation algorithm to reduce UIG, as this was shown to be a major contributor to daily UIG – Task Force findings 13.2.6 suggest a reduction in base UIG of up to 70% on average

Current State

- The NDM Algorithm uses the following key inputs
  a. Assessment of Seasonal Normal weather
  b. Actual within-day weather observations
  c. Annual Load Profile (ALP) – daily gas usage patterns under seasonal normal conditions, by End User Category
  d. Daily Adjustment Factor (DAF) – measure of daily sensitivity to weather fluctuations, by End User Category
  e. Annual Quantity (AQ) – uses ALP, DAF and actual weather to correct actual consumption to a Seasonal Normal position (i.e. relies on a. to d. above)

Development of ALPs and DAFs – Current State



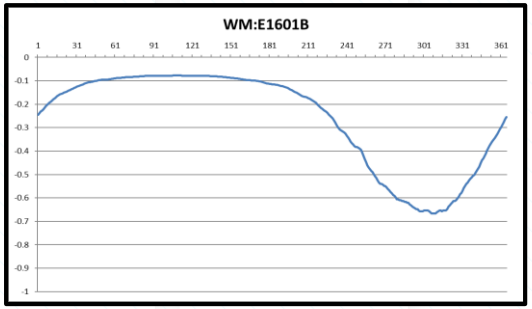


ALP – Annual Load Profile

Simple regression formula between NDM Sample Demand and daily Composite Weather Variable

Input data and output parameters are published – users can recreate the ALPs and DAFs from the data

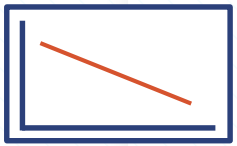
ALPs and DAFs published and consulted on before the start of the Gas Year



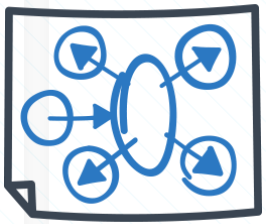
DAF – Daily Adjustment Factor

Options for use of ML in Allocating NDM Energy

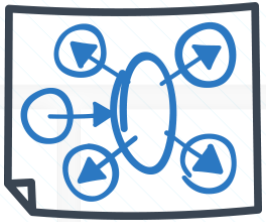
Increasing accuracy of estimates



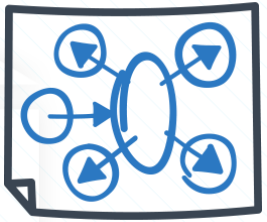
Current State:
Simple regression formula to produce ALPs and DAFs.
LPA factors record the actual allocation after the day.



Machine Learning "black box" creates new ALPs and DAFs every [12] months.
LPA factors record the actual allocation after the day.



Machine Learning "black box" uses agreed input data to predict NDM Energy.
Sends either daily allocations or WAALP factor straight to Gemini.
No ALPs and DAFs published.
LPA factors record the actual allocation after the day.



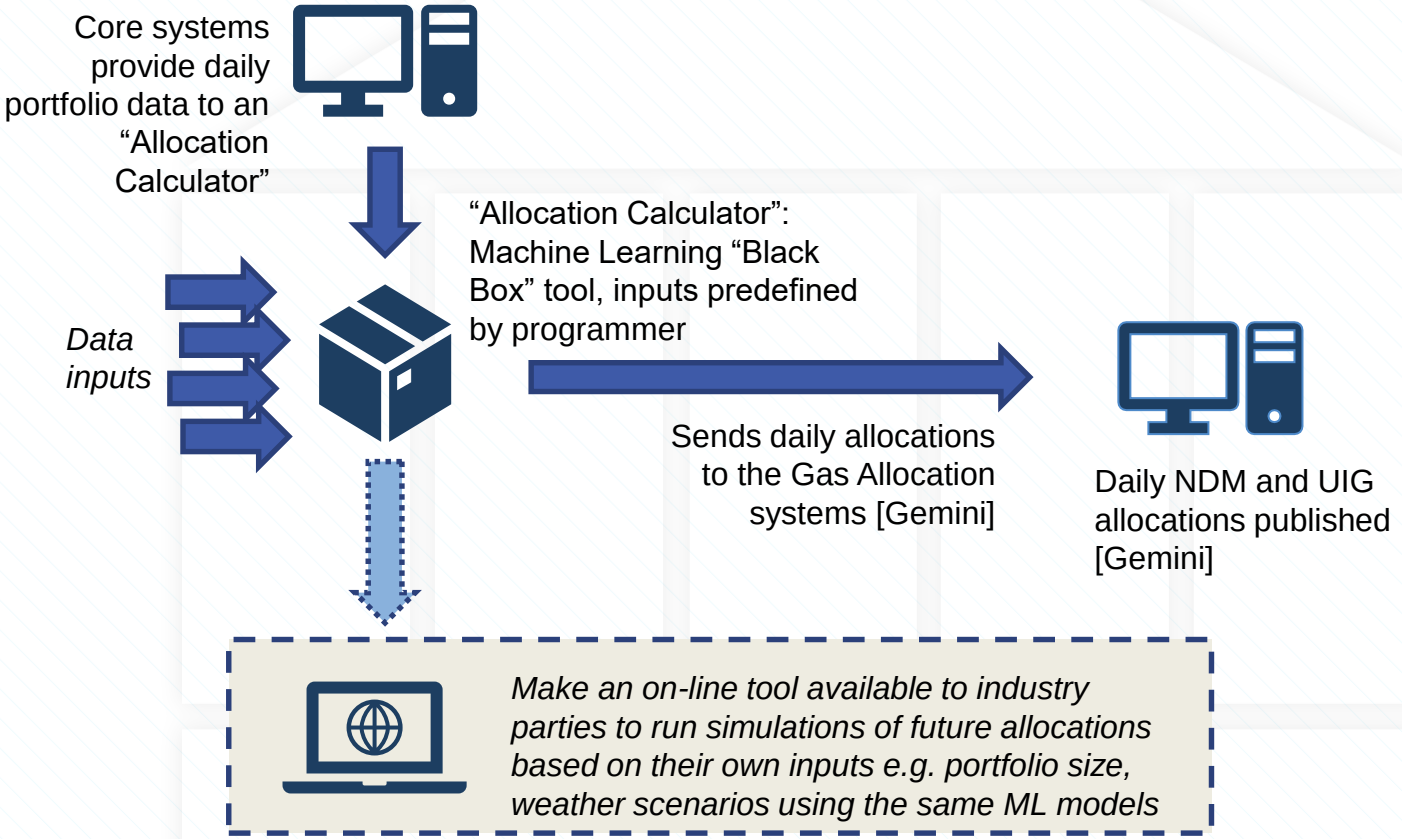
Machine Learning "black box" receives extra data every [month/week] and continually "learns".
Sends daily allocations straight to Gemini.
No ALPs and DAFs published.
LPA factors record the actual allocation after the day.

Increasing complexity

Comparison of Options for NDM Estimation

Option	Pros	Cons
As-Is: Simple Regression Analysis, Annual ALPs and DAFs	Well understood Easy to replicate ALPs and DAFs can be used by all parties for forecasting, estimation etc	Limited set of inputs Not the most accurate way to estimate NDM Demands
Annual Machine Learning development of ALPs and DAFs	ALPs and DAFs produced each year, can be used by all parties for forecasting, estimation etc. Harnesses benefits of ML, using wider range of inputs	Outputs much harder to replicate or explain Only updated once a year Can only learn based on historic observations
Machine Learning model outputs daily energy allocations to the Balancing system	Harnesses more ML benefits Same process can calculate consistent Nominations and Allocations	No ALPs and DAFs produced before the start of the year Needs a new solution to support users' forecasting processes prior to D-1
Ongoing within-year Machine Learning, outputs daily energy allocations to the Balancing system	Model regularly updated for new trends/behaviours	No ALPs and DAFs produced before the start of the year Needs a new solution to support users' forecasting processes

How could Users model future usage in a full Machine Learning environment (with no ALPs and DAFs)?



Observations on Use of Machine Learning for NDM Estimation

- Overall accuracy of the NDM estimates will still be reliant on portfolio data, especially AQs – there will always be an element of NDM model error
- There will still be an inherent level of UIG in every LDZ due to other causes such as theft, missing sites, metering errors, unexpected consumer behaviours
- Daily UIG will still vary depending on other factors such as LDZ Offtake and DM Measurement accuracy
- NDM Allocation will still be an interim position – final positions will still require regular, timely meter read submissions

Suggested Next Steps for Machine Learning

- DESC's autumn/winter workplan is due to include a review of the current NDM Algorithm – options for use of Machine Learning will be considered
- Recommend consultation between Demand Estimation Sub-Committee and the wider industry – feedback on appetite for scale of change v. benefits
- Review UNC Section H and Demand Estimation Methodology to identify what updates are needed depending on chosen solution – is UNC Mod required?
- Identify options for systems solutions to the enhanced calculations

Other Initiatives already implemented/in progress

- Already implemented as a result of findings/recommendations from Machine Learning:
  - Push to improve quality of Winter:Annual Ratios and take-up of WAR Band EUCs (Mod 0652)
  - Additional End User Categories for Domestic/I&C/Prepayment customers (October 2019)
  - Increased focus on NDM sites over the DM Threshold (since Jan 2019)
  - Mandatory provision of NDM Sample data (Mod 0654)
- Soon to be implemented:
  - Use of solar radiation in the Composite Weather Variable (October 2020)


xoserve