# Demand Estimation Sub Committee

## NDM Sample Update

### 10th February 2020

# Objective

- Provide DESC with the latest news relating to the collection of Daily Gas Consumption Data ahead of this years production of Gas Demand Profiles for Gas Year 2020/21

- This will cover:
    1. Reminder of the priority EUCs we require Daily Gas Consumption Data for
    2. An update on Third Party Data Quality Issues we experienced on our recent validation exercise for 2018/19 Algorithm Performance
    3. Latest on the File Format discrepancy
    4. Xoserve Managed Sample

# Background

- The Daily Gas Consumption Data used in the analysis is derived from 3 sources:
    - Xoserve Managed
    - Transporter's Managed
    - Third Party (UNC Mod 654S)


- UNC Modification 654S, which went live on 1st March 2019, requires Shippers with a portfolio of >25k to submit Daily Gas Consumption Data to Xoserve to support Demand Estimation processes

# How is Daily Gas Consumption Data used?

- Once the data has passed validation, it's used in 2 processes:
  - Gas Demand EUC Modelling (Spring)
    - Used to develop Gas Demand EUC Models which enable us to develop Gas Demand Profiles which are used in daily NDM allocation

  - Algorithm Performance (Autumn)
    - Used to assess the performance of the Gas Demand Profiles for the most recent Gas Year

- It is critical that data errors are removed prior to the Gas Demand EUC Modelling process in order to ensure daily NDM allocation is as accurate as possible and therefore reduce the levels of reconciliation

- Modelling errors will always be present but this should be restricted to imperfections in the modelling approach and not caused by issues with Daily Gas Consumption Data

# Target Sample Size (as at January 2020)

| EUC Band | AQ Range (mWh) | Customer Type | LDZ | | | | | | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | SC | NO | NW | NE | EM | WM | WN | WS | EA | NT | SE | SO | SW | |
| 01 | up to 73.2 | Domestic | 384 | 384 | 384 | 384 | 384 | 384 | 384 | 384 | 384 | 384 | 384 | 384 | 384 | 4,992 |
| | | I&C | 381 | 379 | 382 | 380 | 381 | 381 | 364 | 377 | 381 | 382 | 382 | 381 | 380 | 4,930 |
| | | PrePayment | 384 | 383 | 384 | 383 | 383 | 383 | 380 | 383 | 383 | 384 | 384 | 383 | 383 | 4,978 |
| 02 | 73.2 to 293 | I&C | 374 | 367 | 377 | 371 | 375 | 375 | 322 | 358 | 373 | 377 | 376 | 373 | 371 | 4,789 |
| | | Domestic | 348 | 316 | 353 | 337 | 348 | 344 | 160 | 272 | 348 | 367 | 364 | 342 | 332 | 4,231 |
| | | PrePayment | 141 | 85 | 155 | 98 | 145 | 171 | 17 | 77 | 104 | 186 | 144 | 76 | 86 | 1,485 |
| 03 | 293 to 732 | All | 355 | 333 | 358 | 338 | 353 | 352 | 220 | 297 | 347 | 362 | 355 | 347 | 338 | 4,355 |
| 04 | 732 to 2,196 | All | 325 | 270 | 324 | 285 | 315 | 317 | 145 | 235 | 307 | 340 | 315 | 302 | 284 | 3,764 |
| 05 | 2,196 to 5,860 | All | 220 | 150 | 227 | 169 | 207 | 214 | 50 | 113 | 183 | 257 | 196 | 167 | 147 | 2,300 |
| 06 | 5,860 to 14,650 | All | 118 | 78 | 128 | 97 | 137 | 111 | 24 | 47 | 89 | 138 | 81 | 82 | 89 | 1,219 |
| 07 | 14,650 to 29,300 | All | 66 | 38 | 80 | 48 | 80 | 55 | 16 | 22 | 47 | 47 | 35 | 35 | 40 | 609 |
| 08 | 29,300 to 58,600 | All | 35 | 21 | 63 | 24 | 63 | 41 | 7 | 21 | 35 | 28 | 18 | 19 | 28 | 403 |
| | | Total | 3,130 | 2,803 | 3,215 | 2,915 | 3,172 | 3,128 | 2,089 | 2,586 | 2,981 | 3,251 | 3,033 | 2,891 | 2,862 | 38,056 |

- We are currently short of data in all areas, but more noticeably for:
  Pre-Payment (all LDZs) and Domestic (all LDZs)

- No Pre-Payment data has been received this year and in order for us to produce Gas Demand Profiles on up to date behaviours we NEED data from these consumers.

# Data Quality

- Unfortunately, the quality of data being received from Third Parties is generally quite poor, this means that the data cannot be used in this process.

- This results in:
  – inaccurate data could be used in the processes (if not removed by the validation process)
  – for many End User Customer (EUC) groupings there isn't a sufficient number of data points available to produce the most appropriate results

- Some of the more common data errors are explained in the next few slides.

# Incorrect Daily Volumes

- Day of the Week
  - This is where the volume received on an individual record does not correctly reflect the volume consumed for the stated read date
  - This is the most common issue we see across the industry and one of the most important items to get right. The file format states that each daily volume should be provided against the Meter Read Date (relevant to the closing 5am meter read)

- When assessing Domestic Supply Points (in aggregate), the typical consumption pattern tends to show a significant increase on a Saturday and Sunday compared to the Monday to Friday period, as shown below.

Domestic Data
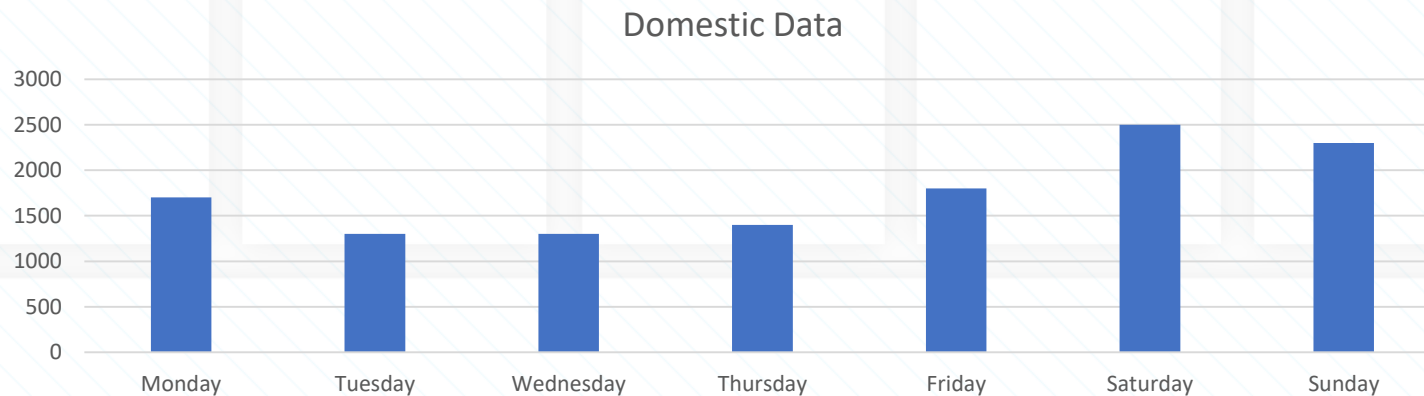
# Incorrect Daily Volumes (continued…)

- Several data errors are regularly identified that cause significant shifts in consumption patterns, these are:

  - Gas flow day supplied instead of meter read day
  - Grouping multiple days of consumption against a single day (if missed from previous days)
  - Smart meters sending reads at midnight a few minutes either side of this will impact the read date
  - An incorrect Market Sector Code on UK Link

- Day of the week checks for I&C sites are more difficult to analyse, due to the varying types of consumers in this sector.

# Negative Volume and Volume Spikes

- We regularly receive daily volumes that are less than zero

| METER_POINT_REFERENCE_NUMBER | METER_READ_DATE | METER_SERIAL_NUMBER | UNCORRECTED_VOLUME | CORRECTED_VOLUME | UNITS_OF_MEASURE | MARKET_SECTOR_CODE |
|---|---|---|---|---|---|---|
| 123456789 | 20190531 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190530 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190529 | G4P12345678912 | 4 | | SCMH | D |
| 123456789 | 20190528 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190527 | G4P12345678912 | -2 | | SCMH | D |
| 123456789 | 20190526 | G4P12345678912 | 3 | | SCMH | D |
| 123456789 | 20190525 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190524 | G4P12345678912 | 1 | | SCMH | D |

- We also see a significant number of daily volume spikes
  - Common instances with erroneous daily volumes of 999,999 most likely caused following a meter exchange
  - Although it is also common to see a volume spike of any value but is incorrectly 10 or more times higher than other volumes for the same period

| METER_POINT_REFERENCE_NUMBER | METER_READ_DATE | METER_SERIAL_NUMBER | UNCORRECTED_VOLUME | CORRECTED_VOLUME | UNITS_OF_MEASURE | MARKET_SECTOR_CODE |
|---|---|---|---|---|---|---|
| 123456789 | 20190531 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190530 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190529 | G4P12345678912 | 4 | | SCMH | D |
| 123456789 | 20190528 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190527 | G4P12345678912 | 25604 | | SCMH | D |
| 123456789 | 20190526 | G4P12345678912 | 3 | | SCMH | D |
| 123456789 | 20190525 | G4P12345678912 | 2 | | SCMH | D |
| 123456789 | 20190524 | G4P12345678912 | 1 | | SCMH | D |

# Data Quality – To Conclude

We are asking DESC to help us by encouraging their colleagues to perform checks on their Daily Gas Consumption Data before submitting to Xoserve, specifically looking at the scenarios mentioned

# File Format

- Please ensure that the files sent to Xoserve match the published file format exactly, including the file name.

| RECORD/FIELD NAME | OPT | DOM | LNG | DEC | DESCRIPTION |
|---|---|---|---|---|---|
| METER_POINT_REFERENCE_NUMBER | M | N | 10 | 0 | A unique numeric reference associated to the supply meter point |
| METER_READ_DATE | M | D | 8 | 0 | The date that the meter/converter read relates to in order to derive the volume (i.e. the date of the closing 5am read). For example, 'Meter_Read_Date' of '20181029' would relate to the Gas Day '28/10/2018'. **FORMAT: YYYYMMDD** |
| METER_SERIAL_NUMBER | M | T | 14 | 0 | The manufacturers meter serial number from which the meter read/consumption was taken |
| UNCORRECTED_VOLUME | M | N | 12 | 0 | The uncorrected metered volume calculated for the relevant gas day (in Cubic Metres or Cubic Feet) |
| CORRECTED_VOLUME | O | N | 12 | 0 | The corrector (converter) volume calculated for the relevant gas day (in Cubic Metres or Cubic Feet) |
| UNITS_OF_MEASURE | M | T | 5 | 0 | Indicator identifying the unit of measurement of the stated uncorrected / corrected volume. Allowable values: **SCFH = Standard Cubic Feet per Hour** **SCMH = Standard Cubic Meters per Hour** |
| MARKET_SECTOR_CODE | M | T | 1 | 0 | A code that specifies that the site is used for domestic or industrial and commercial purposes. Allowable values: **D = Domestic** **I = Industrial** |
|  |  |  | 62 |  |  |

File naming should be:

SM01_SSC_yyyymmddvv.CSV

e.g
SM01_ABC_2019122901.CSV

Please ensure that file headers are included in every file

As per the DESC meeting in December 2019, we are changing the file format to allow for 2 decimal places in Uncorrected_Volume and Corrected_Volume.

Going forward, Xoserve will accept files that contain this change to the volume.

# Xoserve Managed Daily Gas Consumption Data

- Sample provides Daily Gas Consumption Data for Band 1 and 2 supply points – majority of which are Domestic

- As reported to DESC before, numbers are being impacted by SMART rollout

- Currently there are c. 2300 in the sample

- It is anticipated there will be <2000 by April 2021 and so we are hopeful we shall have reasonable numbers for modelling for this year and next year, but these will still be significantly lower than the 200-250 sample points per LDZ we used to receive

- The 5 year guaranteed battery life is due to start being passed from September 2020 onwards

- Decisions around replacing batteries need to be taken given background of SMART meter rollout and availability of 3rd party data

# Contacts

Should you have any queries for this process, please contact us using the details below:

Xoserve.Demand.Estimation@Xoserve.com

Simon Bissett – 0121 229 2640