DNV·GL

ALLOCATION OF UNIDENTIFIED GAS

Final Allocation of Unidentified Gas Statement for 2019/20

Report No.: 1, Rev. 1.0
Date: 27th March 2019

Project name:	Allocation of Unidentified Gas	DNV GL - Software
Report title:	Final Allocation of Unidentified Gas Statement for 2019/20	Software Consulting
		Holywell Park
Customer:		Ashby Road
Customer contact:	Fiona Cottam	Loughborough
Date of issue:	27th March 2019	LE11 3GR
Project No.:	4011094	Tel: +44 (0)1509 282000
Organisation unit:	Software	
Report No.:	1, Rev. 1.0	

Objective: This document is the Final AUG Statement for 2019/20 and contains details of the methods developed by the AUG Expert for allocating daily UIG between Product Class/EUC including details of the data requested to support this analysis, data received following such requests and any assumptions made.

This document also includes an estimate of the total quantity of permanent Unidentified Gas and identifies each Unidentified Gas source with its proportionate contribution to the total.

This document contains the updated methodology following the consultation period on the Proposed AUG Statement and an updated estimate of the AUG weighting factors based on this updated methodology and using the latest data available.

Prepared by:	Verified by:	Approved by:
Tony Perchard Principal Consultant	Clive Whitehand Head of Section	Anthony Gilbert Global Delivery Manager
Andy Gordon Senior Consultant		
Mark Lingham Group Leader		

DNV GL Distribution:

☒ Unrestricted distribution (internal and external)

☐ Unrestricted distribution within DNV GL Group

☐ Unrestricted distribution within DNV GL contracting party

☐ No distribution (confidential)

Keywords:

Unidentified Gas, AUGE, AUGS

Rev. No.	Date	Reason for Issue	Prepared by	Verified by	Approved by
1.0	2019-03-27	For Publication	Mark Lingham, Andy Gordon & Tony Perchard	Clive Whitehand	

EXECUTIVE SUMMARY

Project Nexus was implemented in June 2017, replacing key IT systems for gas settlement and supply point administration in the gas industry ('UKLink') and changed the way that the gas settlement is handled. Project Nexus introduced individual meter point reconciliation for all meter points including those on CSEPs and a rolling monthly AQ process. It also introduces 4 new meter point 'classes'.

As a result of these changes, there is a requirement to fairly apportion the daily total UIG estimate (daily imbalance calculated from the settlement process) between Product Classes and EUC. Mod 0473 was raised to allow the appointment of an independent expert (AUG Expert) to develop a methodology to do this and provide a table of weighting factors that will target the correct amount of UIG to different classes of meter points, based on an assessment of their relative contribution. The table of weighting factors will be used in the daily gas nomination, allocation and reconciliation processes.

This document is the Final AUG Statement for 2019/20 and contains details of the methods developed by the AUG Expert for allocating daily UIG between Product Class/EUC including details of the data requested to support this analysis, data received following such requests and any assumptions made. This document also describes how the AUG Expert has followed the published guidelines and provides the following additional supporting information

a) identifies each Unidentified Gas Source

b) estimates the total quantity of permanent Unidentified Gas

c) specifies the proportionate contribution of each Unidentified Gas Source to the total quantity of Unidentified Gas

d) provides a first estimate of the AUG weighting factors (based on the latest available data at the time)

Following publication of the Proposed AUG Statement, there was a 21 day consultation period allowing the industry to provide feedback to the AUG Expert and raise any questions/issues. Where appropriate, the methodology has been updated based on this feedback. More detail can be found in the AUG Expert responses to the issues raised by the industry which have been published separately.

This document details the updated methodology including any changes resulting from the consultation period[52][53][54][55]. The AUG factors included within this document are the final factors based on the updated methodology described within this Final AUG Statement and are based on the latest available data.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Background

The majority of gas consumed in Great Britain is metered and registered. However, some gas is lost from the system, or not registered, due to theft, leakage from gas pipes, consumption by unregistered supply points and other reasons. Some elements of the gas that is not directly consumed/measured are currently modelled, and hence the gas consumed by these can be estimated. The gas that is lost and not recorded or modelled is referred to as Unidentified Gas.

The Great Britain gas industry is segmented into two market sectors; Larger Supply Points (LSP) and Smaller Supply Points (SSP). Prior to April 2012 there was no methodology in place to determine the allocation of Unidentified Gas between the LSP and SSP market sectors: Unidentified Gas was ultimately borne by the SSP market sector following reconciliation (an interim amount was allocated for 2011/12). Through the approval of UNC Modification 0229 and the appointment of DNV GL as the Allocation of Unidentified Gas Expert (AUGE) in 2011, a methodology was developed to calculate and apportion Unidentified Gas equitably to the relevant gas market sectors. This approach involved an annual estimate of Unidentified Gas and a monthly transfer of costs between market sectors to address the misallocation of Unidentified Gas that occurred under that regime. DNV GL carried out this annual process until 2014 when it became clear that the requirements for the AUGE would change given the upcoming implementation of project 'Nexus' (Mod 0432).

Project Nexus involved the replacement of key IT systems ('UKLink') for gas settlement and supply point administration in the gas industry. It introduced individual meter point reconciliation for all meter points (previously SSP meters were subject to reconciliation by difference) and a rolling monthly AQ process. It also introduced four new meter point 'classes'. Following Project Nexus implementation, an amended NDM Algorithm (with scaling factor removed) uses actual weather data to derive a bottom-up estimate of NDM Demand. This allows the calculation of a daily Balancing Factor, referred to as UIG. This UIG is shared out to all live sites, on the basis of the AUG weighting factors and their recorded/estimated throughput for the day.

Note this quantity is not the same as the Unidentified Gas figure calculated previously by the AUGE. UIG is calculated 'live' on a daily basis using estimated NDM demand based on prevailing AQs and the NDM algorithm; the AUG Expert estimate of Unidentified Gas was calculated retrospectively (and then projected forward) on an annual basis using estimated NDM demand based on meter reads/volumes for approximately 90% of sites with the remaining 10% based on average EUC demands.

As a result of these changes, the industry noted a requirement to be able to fairly apportion this total UIG between Product Classes and End User Category (EUC). Mod 0473 was raised to allow the appointment of an independent expert (the AUG Expert) to develop a methodology to do this and provide a table of weighting factors that will assign the correct amount of UIG to different classes of meter points, based on an assessment of their relative contribution to Unidentified Gas. This table of weighting factors is used in daily gas allocation processes. Daily measured or estimated gas throughput in each sector are weighted using the AUG table factors to assign daily UIG to Shippers based on their throughput by meter point class and EUC.

DNV GL were appointed to the new role of AUG Expert in July 2016, with Project Nexus go-live occurring on 1st June 2017.

1.2 High Level Objectives

The AUG Expert's high level objectives are:

- To assess the sources of Unidentified Gas and the data available/required from industry bodies to evaluate Unidentified Gas
- To gather data as required from the CDSP, from Gas Shippers, or drawn from other sources, as deemed appropriate by the AUG Expert
- To develop the methodology to assess the relative contribution to Unidentified Gas of different Classes and sizes of sites
- To publish the methodology in the AUG Statement (this document) and present findings to the industry
- To consult with the industry bodies and respond to questions/issues raised, assess the impact of questions on the methodology and update as appropriate
- Produce the table detailing the weighting factors for each Product Class and EUC

1.3 Scope

This document is the Final AUG Statement for 2019/20 and contains the following:

- A detailed description of the proposed methodology including updates following consultation
- A summary of data requested, received and used, and associated assumptions
- The table of weighting factors for apportioning UIG between Product Classes and EUCs
- Details of the database used to hold information associated with Unidentified Gas and used to develop the methodology

This document will be published to the industry for review and comment during the consultation period.

The following are out of scope, but may be referenced where they may impact on the methodology to calculate the AUG factors.

- The AUG Expert is not concerned with issues regarding the deeming algorithm or the former RbD mechanism.
- The AUG Expert is not concerned with resolution of fundamental gas industry business process issues.
- The AUG Expert process is not an opportunity to provide permanent solutions to issues within the gas industry that should be addressed by other workgroups (e.g. Shrinkage Forum). This includes Shrinkage Error (i.e. bias in the Shrinkage and Leakage Model), although it is acknowledged that any such errors feed through into the daily UIG calculation.
- The AUG Expert is not concerned with transportation charges.

2 COMPLIANCE TO GENERIC TERMS OF REFERENCE

This section describes how DNV GL has adhered to the Generic Terms of Reference described in Section 5 of the AUGE Framework[43].

The AUG Expert will create the AUG Statement and AUG Table by developing appropriate, detailed methodologies and collecting necessary data.

The AUG Expert has developed a detailed methodology for estimating factors to apportion UIG between EUC and Product Classes. To calculate the factors, total Unidentified Gas is also estimated using meter read and consumption data for all meters, which has been obtained from the CDSP. Further detailed datasets are used to directly estimate some components of the total Unidentified Gas where this is possible e.g. Shipperless sites. The AUG Expert has also developed a methodology to account for elements of Unidentified Gas which are Temporary in nature.

Additional data regarding theft of gas (the SPAA Schedule 33 data) was sought and obtained from SPAA. Additional, more detailed theft data has also been requested via SPAA.

The decision as to the most appropriate methodologies and data will rest solely with the AUG Expert taking account of any issues raised during the development and compilation of the AUG Statement and AUG Table.

For the avoidance of doubt although UIG includes any LDZ Shrinkage Error, the AUG Expert acknowledges that the process for determining LDZ Shrinkage is laid out in the relevant DNO licences. To avoid dual governance of any LDZ Shrinkage Error, the AUG Expert's role in respect of any LDZ Shrinkage Error is therefore limited to confirming that there are controls in place to ensure that DNOs discharge their licence obligation (that is that there is a methodology and that it is periodically reviewed for confirmation that the methodology remains relevant). The AUGE will present any comments or observations on the LDZ Shrinkage model through the annual consultation carried out by the DNOs.

The proposed methodology and assessment of what constitutes Unidentified Gas has been decided solely by the AUG Expert based on available information. Comments raised by shippers will be considered and additional analysis undertaken where necessary. Having considered all views, the final decisions are the AUG Expert's own.

LDZ Shrinkage Error is not considered by the AUG Expert. However, the AUG Expert has made an assessment of IGT Shrinkage as this is not included within the current LDZ Shrinkage models.

The AUG Expert will determine what data is required from Code Parties (and other parties as appropriate) in order to ensure appropriate data supports the evaluation of Unidentified Gas.

The AUG Expert has assessed what data is required to support the proposed methodology and has requested information from relevant parties. Updated data sets have been requested from the CDSP together with any relevant new datasets available following Nexus implementation.

The AUG Expert will determine what data is available from parties in order to ensure it has appropriate data to support the evaluation of Unidentified Gas.

The AUG Expert has determined what data is available following discussions with the CDSP. A request has also been made to SPAA for further information regarding theft. Information has also been provided by ICoSS on AMR uptake not included within the BEIS dataset used by the AUG Expert.

The AUG Expert will determine what relevant questions should be submitted to Code Parties in order to ensure appropriate methodologies and data are used in the evaluation of unidentified error.

The AUG Expert issued a request for information to the industry in September 2018 and raised a number of questions at the AUG Technical Workgroup Early Engagement Meeting. Further communication will take place as and when necessary.

The AUG Expert will use the latest data available where appropriate.

Most of the data requested by the AUG Expert has been provided, and the latest data has been used where possible. Further data updates will be used in future re-calculations of the table of factors.

Unfortunately, due to issues with the meter read data, the consumption calculations have not been re-run and the factors are therefore based on consumption estimates from the last AUG year (the latest available data).

Where multiple data sources exist the AUG Expert will evaluate the data to obtain the most statistically sound solution, will document the alternative options and provide an explanation for its decision.

For the Consumption Method of estimating total Unidentified Gas, both meter reads and metered volumes are provided. Over time LSP metered volumes may be corrected, but the meter reads are not. The AUG Expert's analysis has shown that metered volumes can be erroneous, particularly for non-corrected SSP data. The decision was therefore taken to use meter reads for SSP and metered volumes for LSP.

Where data is open to interpretation the AUG Expert will evaluate the most appropriate methodology and provide an explanation for the use of this methodology.

Throughout the statement the AUG Expert has described how data will be used and why.

Where the AUG Expert considers using data collected or derived through the use of sampling techniques, then the AUG Expert will consider the most appropriate sampling technique and/or the viability of the sampling technique used.

The Consumption Method for estimating the Unidentified Gas total is the only part of the analysis where a sample rather than the full dataset is used. This calculation will be at its most accurate when the largest possible representative subset of the meter point population is used. To achieve this, a validation process was developed that was designed to maximise the sample size whilst removing any meter points with invalid data. Appropriate methods are then applied to scale up for any meter points which have been excluded.

The AUG Expert will present the AUG Statement, including the AUG Table, in draft form (the "proposed AUG Statement"), to Code Parties seeking views and will review all the issues identified submitted in response.

The AUG Expert has documented and reviewed all feedback resulting from previous versions of the AUG Statement. Section 9 of this document refers to these publications with details of the issues raised, with the full text of the comments from the Code Parties and the AUG Expert responses contained in separate documents published on the Joint Office of Gas Transporters website.

The AUG Expert will provide the proposed and any modified AUG Statement, including the AUG Table, to the Gas Transporters for publication who will then provide the AUG Statement and Table to the CDSP.

This Final AUG Statement for 2019/20 is provided to the Gas Transporters in advance, for publication on 1st April 2019. A Proposed AUG Statement and a Modified AUG Statement have already been provided and published.

The AUG Expert will ensure that all data that is provided to it by parties will not be passed on to any other organisation, or used for any purpose other than the creation of the methodology and the AUG Statement and Table.

On receipt of data, the AUG Expert stores the data in a secure project storage area with limited access by the consultants working on the project. The AUG Expert can confirm data used in the analysis has not and will not be passed on to any other organisation. The data used will be made available to all bona fide industry participants in order to review the methodology, and in this dataset all MPR information has been replaced by 'dummy' MPR references by the CDSP so that the anonymity of the consumer is protected.

The AUG Expert shall ensure that all data provided by Code Parties will be held confidentially, and where any data, as provided or derived from that provided, is published then it shall be in a form where the source of the information cannot be reasonably ascertained.

Data is stored in a secure project storage area with access limited to those working on the project. Any data that contains market share or code party specific information has been and will be made anonymous to ensure the source of the information cannot be ascertained.

The AUG Expert will act with all due skill, care and diligence when performing of its duties as the AUG Expert and shall be impartial when undertaking the function of the AUG Expert, ensuring that any values derived will be equitable in their treatment of Code Parties.

When developing the methodology, the AUGE does not consider the implications of the weighting factors generated from a Code Party impact perspective i.e. it is an independent, evidence based approach. This is noted by OFGEM when describing the role of the AUGE [52].

The AUG Expert will compile the methodology and AUG Statement and AUG Table in accordance with this Framework.

The AUGE has, following the revision of the AUGE Framework earlier in the year modified its approach to the development of the methodology. This has included early engagement meetings with Code Parties and more regular progress updates to ensure the industry are fully up to date with progress. The activities for the remainder of the process are also planned on the revised AUGE Framework.

3 TERMINOLOGY

This section aims to provide clarity surrounding the use of terminology relating to Unidentified Gas.

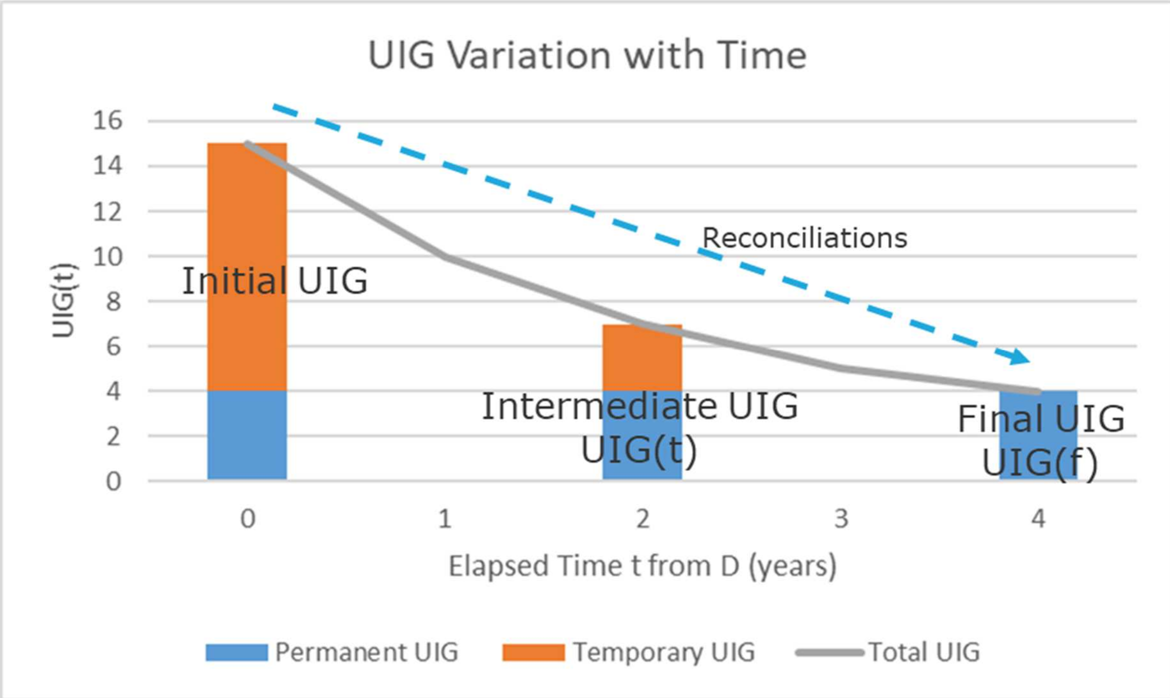

Figure 1: Unidentified Gas Definitions

Unidentified Gas refers to gas which is consumed in an unrecorded manner at or beyond the meter. Estimates of Unidentified Gas for a given gas day (D) are initially made on the day as part of the settlement process. This "initial" value of Unidentified Gas is defined in UNC Section H 2.6.1 as UIG. It is calculated by taking the total LDZ input and subtracting all of the LDZ outputs (either known values or estimates).

At the time of calculating initial UIG, the vast majority of meter reads are not available. This initial value therefore includes an element relating to the uncertainty in the estimate of demand (model error). Ideally this error would be random and have a zero bias, but this is dependent on the estimation algorithm and its inputs being unbiased. Over time, as meter reads are received, this part of UIG will reduce towards zero.

The model error described above is considered as temporary Unidentified Gas as it is corrected for over time through reconciliation. There are further sources of Unidentified Gas which are also temporary in nature e.g. LDZ metering errors which are identified and corrected for. UIG can therefore be recalculated at any point in time after the initial gas day taking into account this updated information. A Code Cut Off Date or 'line in the sand' is defined in UNC, beyond which no further reconciliations will be applied. Once a gas day falls beyond this date the value of UIG at that point is referred to as Final UIG or UIG(f) and consists of only permanent Unidentified Gas.

Intermediate UIG or UIG(t) refers to the value of UIG calculated at some point in time (t) between D and the line in the sand. This value will contain a mix of permanent and temporary Unidentified Gas.

Throughout this document, the term Unidentified Gas will refer to the general concept, i.e. any gas burnt in an unrecorded manner. Where the reference needs to be more specific, the following terms will be used:

- UIG – Initial UIG as defined in UNC Section H 2.6.1
- Intermediate UIG or UIG(t) – UIG recalculated at some point in time, t between D (initial allocation) and the line in the sand
- Final UIG or UIG(f) – UIG as at line in the sand

The methodology described in this statement aims to provide an estimate of the Final UIG and provide factors to apportion this between EUC and product classes.

4 HIGH LEVEL OVERVIEW OF METHODOLOGY

This section provides a high-level overview of the methodology. For each of the areas of Unidentified Gas presented here a more detailed discussion is given in Section 7.

Under the pre-Nexus settlement regime, an independent forecast of permanent Unidentified Gas split by market sector was made to allow correct allocation of Unidentified Gas costs. Following implementation of Project Nexus, UIG is calculated daily as part of the settlement process. This UIG amount is then apportioned based on the weighting factors calculated by the AUG Expert using the methodology described in this document.

As part of the AUG methodology, the AUG Expert makes an independent estimate of Final UIG. The Initial UIG value calculated by the CDSP includes elements which are temporary in nature and are removed over time through the reconciliation process. The Final UIG calculated by the AUG Expert to derive the AUG weighting factors is an estimate of the value of UIG following all reconciliations.

4.1 LDZ Load Components

The Unidentified Gas calculations for the 2017/18 gas year were complicated by the fact that the UIG Factors, which are the ultimate output of the work, are required to be split by post-Nexus market sector definitions, whilst the available data was from the pre-Nexus regime with its associated market sectors. The load components are different in these two scenarios, and this created a requirement to map from one to the other in as accurate a manner as possible as part of the calculation process.

This is no longer the case for the 2019/20 analysis, for which actual Product Class data is available. For completeness, and to provide a link with the historic Unidentified Gas analysis, both pre-Nexus and post-Nexus regimes are described below.

4.1.1 Pre-Nexus Regime

Daily load (as measured or calculated at the Supply Meter Point) fell into three relevant categories as far as the reconciliation process was concerned. These are as defined in Section A of the Uniform Network Code (UNC) [2]:

1. **Smaller Supply Point Component Load**

   Load from Supply Point Components (SPCs) which are part of a Smaller Supply Point (SSP). This is defined as a supply point where the AQ is not greater than 73,200 kWh.

2. **Larger Non-Daily Metered Supply Point Component Load**

   Load from Non-Daily Metered (NDM) SPCs which were part of a Larger Supply Point (LSP). This is defined as a supply point where the AQ is greater than 73,200 kWh but less than the mandatory daily metering threshold of 58,600,000 kWh. Note that historically (prior to the implementation of Mod 0428), Larger NDM SPCs may have contained individual meters that fell below the SSP AQ threshold.

3. **Daily Metered Supply Point Component (DM SPC) Load**

   Load from Daily Metered (DM) SPCs. This includes Daily Metered Mandatory (DMM) sites, which are those above the 58,600,000 kWh threshold, Daily Metered Voluntary (DMV) and Daily Metered Elective (DME) sites.

4.1.2 Post-Nexus Regime

Following Project Nexus implementation, the population of supply points is instead split into four different Product Classes, each of which have different meter read frequency requirements and reconciliation rules. A list of products and associated details (including approximate equivalence to historic services) is shown in Table 1 below. Information in this table is taken from UNC Modification 0432 [3].

Process Description	**Basis of Energy Allocation**	**Basis of Energy Balancing**	**Shipper Read Submission**	**Market Sector**	**Historic Service Equivalent**
Product 1: Daily Metered Time Critical Readings	Daily Read	Daily Read	Daily by 11 am on GFD+1	DM	DM Mandatory
Product 2: Daily Metered not Time Critical Readings	Daily Read	Daily Read	Daily by end of GFD+1	DM	DM Voluntary / DM Elective
Product 3: Batched Daily Readings	Allocation Profiles	Allocation Profiles	Periodically in batches of daily readings	NDM	Non-Daily Metered
Product 4: Periodic Readings	Allocation Profiles	Allocation Profiles	Periodically	NDM	Non-Daily Metered

Table 1: Product Classes

Each site is classified as subscribing to one of these products, and meter read submissions, settlement and reconciliation is then carried out for each site in the manner suitable for its Product Class.

In addition to splitting the UIG figure between the four products, Mod 0473 also includes a requirement to include a split between EUCs [4]. The final output of the AUG process will therefore be a table of UIG allocation adjustment factors with the following structure:

Supply Meter Point Classification	**Unidentified Gas Weighting Factor**			
	Class 1	**Class 2**	**Class 3**	**Class 4**
EUC Band 1				
EUC Band 2				
EUC Band 3				
EUC Band 4				
EUC Band 5				
EUC Band 6				
EUC Band 7				
EUC Band 8				
EUC Band 9				

Table 2: Example AUG Table

4.1.3 Unidentified Gas

Daily Metered load (i.e. Product Classes 1 and 2) is, by definition, metered and known on an ongoing daily basis. Like all metered load it can be subject to metering error, and data for known errors is used to correct it. NDM load (Product Classes 3 and 4) for a given day is estimated using the NDM deeming algorithm, to which reconciliation is applied when the meter readings become available. The estimation process is described in Mod 0432 for the post-Nexus regime [3].

Even after reconciliation and meter corrections are applied, the sum of these load components does not equal the gas intake into the LDZ due to the presence of two further factors:

1. **Shrinkage**

   LDZ shrinkage occurs between the LDZ offtake and the end consumer (but not at the Supply Meter Point - the LDZ shrinkage zone stops immediately before this point). It covers:

   - Leakage (from pipelines, services, AGIs and interference damage)
   - Own Use Gas
   - Transporter-responsible theft

   The majority of shrinkage is due to leakage, and the overall LDZ shrinkage quantity is calculated using the standard method defined in the UNC [2].

2. **Unidentified Gas**

Unidentified Gas occurs downstream of shrinkage, i.e. at the Supply Meter Point. It potentially covers:

- Unregistered and Shipperless sites
- Independent Gas Transporter Connected System Exit Point (IGT CSEP) setup and registration delays
- Errors in the shrinkage estimate
- Shipper-responsible theft
- Meter errors – this includes both LDZ offtakes and consumer meters
- Volume conversion errors

Unidentified Gas is currently unknown and hence must be estimated.

The relationship between these components of daily load can therefore be expressed as follows:

Total Unidentified Gas = Aggregate LDZ Load – Product 1 to 4 Load – Shrinkage *(4.1)*

4.2 Permanent and Temporary Unidentified Gas

Unidentified gas can be divided into two categories:

Permanent Unidentified Gas is consumed in an unrecorded fashion and costs are never recovered e.g. undetected theft.

Temporary Unidentified Gas is initially consumed in an unrecorded fashion, but volumes are later calculated directly or estimated, and the cost is recovered via back billing or reconciliation. This includes:

- Errors in DM meter reads which are later corrected
- Errors in estimated NDM consumption which are later corrected when new meter reads are available

4.3 Unidentified Gas Methodology

4.3.1 Estimation of Total Unidentified Gas for Historic Years

The overall concept of calculating total Unidentified Gas using metered consumption data is simple and is centred on the basic principle of the allocation process. Note that all historic consumption data used by the AUG Expert is still from the pre-Nexus period at this stage, and hence the calculation is based on the principles of the old regime. Given that the method outputs an aggregate national figure for Unidentified Gas, this output can safely be extrapolated to the post-Nexus 2019/20 gas year for which the factors are being produced. The NDM Allocation is calculated as follows:

NDM Allocation = Aggregate LDZ Load – DM Load – Shrinkage

This is shown graphically in Figure 2.

As the NDM load in equation 3.1 is the sum of all metered NDM consumptions, this allows us to rewrite it as

Total Unidentified Gas = NDM Allocation – Metered NDM Consumption (4.2)

This formulation of Unidentified Gas means that the total will include any Permanent Unidentified Gas arising from unresolved errors in aggregate LDZ load, DM load and shrinkage.

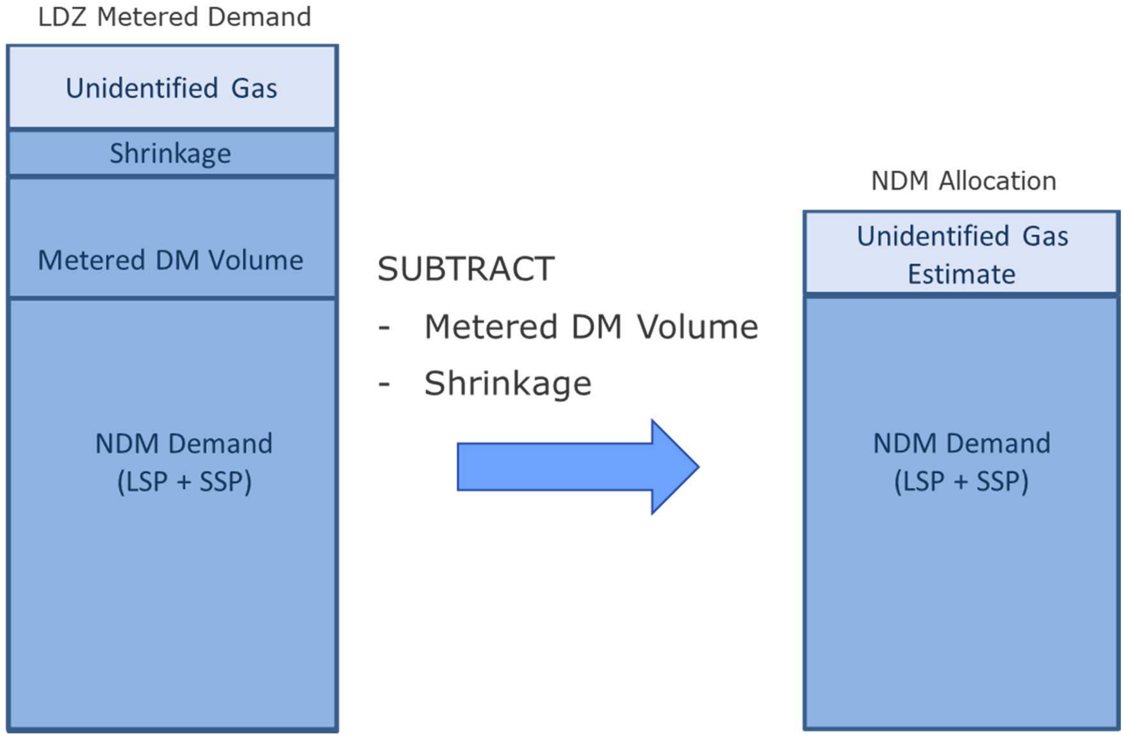

Figure 2: NDM Allocation and Unidentified Gas

The metered NDM consumption is calculated for each meter point and gas year using meter reads or metered volumes, and meter asset information. There are several complexities in this calculation that must be accounted for in the methodology and a fall-back approach must be developed for those meter points without sufficient data of suitable quality. This is summarised in Figure 3. The full details of the Consumption Methodology can be found in Section 7.2.

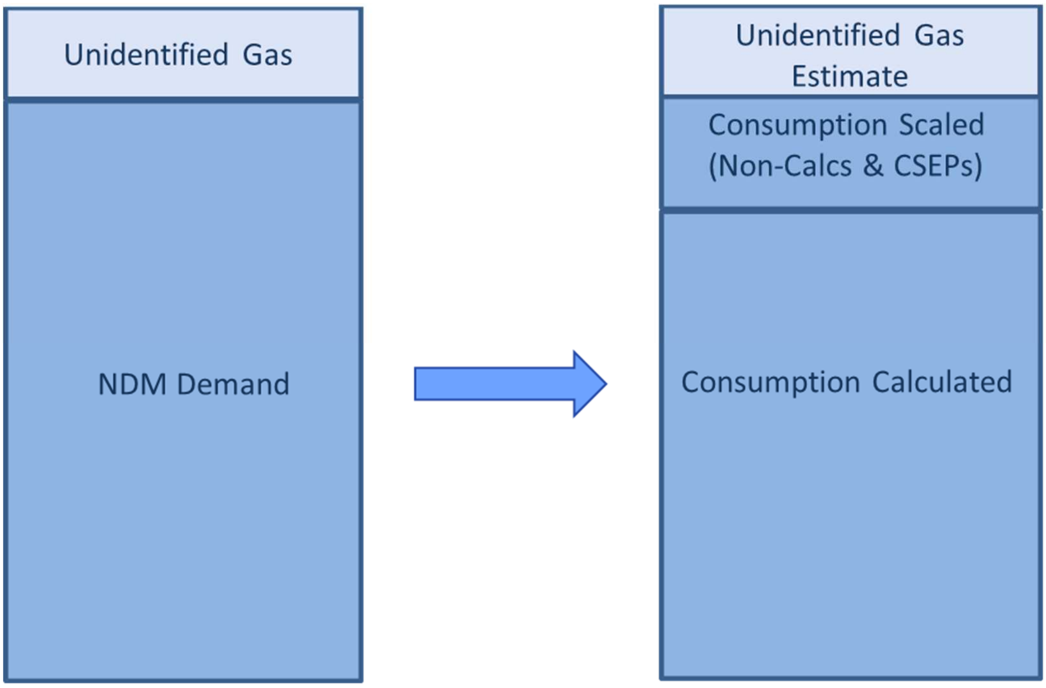

Figure 3: Calculation of Unidentified Gas from Consumptions and Allocations

This method is used to estimate total Unidentified Gas on a gas year basis. This is initially done on an LDZ by LDZ basis due to the very high volume of data required (i.e. all meter reads for all sites). The AUG Expert then determines which years are most suitable for use in the process to estimate the total Unidentified Gas. Recent data is excluded as the number of meters where consumption can be successfully calculated is much lower due to fewer meter reads being available. The use of this data is therefore subject to a large degree of uncertainty. Older data may also be excluded as this is less relevant and may have data quality issues.

Note that at this stage the total Unidentified Gas figure contains both Permanent and Temporary Unidentified Gas.

4.3.2 Calculating Components of Total Unidentified Gas

Having obtained the total Unidentified Gas figure using the Consumption Methodology described above, the value of individual components that make up the Unidentified Gas total are calculated where this is possible. This also includes the calculation of the amount of this Unidentified Gas which is temporary for each component and how the Unidentified Gas is split between market sectors. "Market sector" is now defined in the post-Nexus terms of Product Class and EUC, despite the fact that the training period still lies mostly in the pre-Nexus period. This approach therefore requires a method of applying post-Nexus Product Class information to the pre-Nexus period, and this is defined in Section 4.3.3 below. Splitting each directly calculated source of Unidentified Gas by EUC/Product in this way allows calculations for both the historic period (pre-Nexus) and forecast period (post-Nexus) to be carried out as follows:

1. Pre-Nexus data is required as a total for each Unidentified Gas component (but split into permanent and temporary elements). As described below, however, the Product/EUC split for this time period is still required for training and extrapolating values for the forecast year, and hence the calculations are still carried out on this basis even for the historic years. The output that feeds directly into the calculations for these years is the sum across the categories and so the final figures for the historic

years are not dependent on the categorisation process used. It remains the most efficient approach to work with the Product/EUC split, however, because it is required for the forecast anyway.

2. The post-Nexus Unidentified Gas forecast must be carried out separately not only for each Unidentified Gas component but every individual EUC/Product combination within each component. This is because each EUC/Product combination can follow its own trend over time (depending on market conditions, Mods that have been created to address individual issues, and so on). Therefore, each must be calculated across the whole historic period so that the trend can be identified and extrapolated to the forecast period. A full split into EUC/Product categories is required for this work.

It is known that data for each of the six potential components of Unidentified Gas (Unregistered and Shipperless sites, IGT errors, CSEP shrinkage, Shipper-responsible theft, metering errors and volume conversion errors) is available. The availability and quality of this data varies from component to component, and the AUG Expert has therefore identified the best method of calculating each Unidentified Gas component based on the quality of information available for that component.

Brief descriptions of each Unidentified Gas element are given below.

1. Unregistered and Shipperless Sites

The data available for this element consists of the details for every site that is either Shipperless or Unregistered at a given point in time. This point in time is the snapshot date, and snapshots are provided on a monthly basis, allowing the trends in each such Unidentified Gas category to be monitored over time. The details for each site include AQ, which allows each site to be assigned to the correct EUC and also allows its gas usage whilst Unregistered/Shipperless to be estimated. Unregistered and Shipperless sites that contribute to Unidentified Gas are split into the following sub-categories:

- Shipper Activity
- Orphaned Sites
- Unregistered <12 Months
- Shipperless Passed to Shipper (PTS)
- Shipperless Shipper Specific Report (SSrP)
- Sites Awaiting GSR Visit

2. IGT CSEP Setup and Registration Delays

Gas consumed in an unrecorded manner due to IGT CSEP setup and registration delays is also included in the Unidentified Gas calculation. Unidentified Gas from this source is due to gas networks owned by IGTs but not present in the CDSP’s records, and also comes from Unregistered sites on known CSEPs. The data available for this analysis consists of the number and composition of these unknown projects (number of sites and AQ split by EUC), and the number and AQ of each Unregistered site associated with a known project. Unknown Project data is again provided in monthly snapshots, allowing trends over time to be established.

3. Shrinkage Error

Shrinkage errors affect the total Unidentified Gas calculation in that estimated Shrinkage is deducted from the LDZ input total (along with DM load) to give the total NDM allocation from which metered load is then removed to calculate total Unidentified Gas. The Shrinkage estimate comes from the Shrinkage Model, and if this is biased it will affect the Unidentified Gas estimate.

Following a recent change to the AUGE Framework[43] the estimation of Shrinkage Error by the AUGE has been ruled out of scope. Therefore, the Shrinkage estimate provided by the GTs will be used without further adjustment.

CSEP Shrinkage is not included in the Shrinkage model. This is relatively small but non-zero and is not taken account of anywhere in the settlement process, and therefore filters through into the daily UIG figure. The value of CSEP Shrinkage must therefore be estimated and split by EUC and Product Class as appropriate.

4. Shipper-Responsible Theft

Whilst detailed information for investigations and confirmed theft is available, theft by its nature is often undetected. Undetected theft levels are very difficult to quantify accurately, and estimates from different sources vary widely, from 0.006% of throughput (based on detected theft only) to around 10%. As it is difficult to accurately estimate theft levels directly, undetected theft is calculated by subtraction once known levels of detected theft have been accounted for. Undetected theft is part of the Balancing Factor (see 7 below), and considered over time, it forms the vast majority of that figure.

5. Meter Errors

Meter errors affect Unidentified Gas in different ways depending on their source. Errors in LDZ offtake metering and DM supply metering affect the estimate of total NDM demand including Unidentified Gas for the training period (which is all pre-Nexus), whilst NDM LSP and SSP metering errors contribute to Unidentified Gas by affecting the NDM metered total. Corrections are applied to LDZ offtakes, DM and unique site meters using detected error data supplied by the CDSP. In addition, the effects of consumer meters (all EUC/Product combinations) under- and over-reading due to operating at the extremes of their range are modelled and included in the calculations.

6. Volume Conversion Errors

The volume conversion errors referred to here are the result of using assumed static values for the atmospheric pressure (1013.25mbars) and temperature (12.2°C) at the meter. In reality, both quantities will vary daily. If the assumed values do not represent the true average (consumption weighted) annual values, permanent Unidentified Gas will result. Meters with volume conversion equipment will not be subject to this error. An estimate of the permanent Unidentified Gas can be made based upon

- Details of meters with volume conversion equipment
- Historical atmospheric pressure data
- Historical meter temperature data

7. **Balancing Factor**

The Balancing Factor is calculated by taking the difference between the calculated total Unidentified Gas and the sum of the directly estimated components. The Balancing Factor is comprised of Unidentified Gas elements that cannot be calculated directly because data is either unavailable or unreliable, and is believed to be mostly undetected theft.

The Permanent component of total Unidentified Gas is then given by the sum of the Balancing Factor and the Permanent components of the directly calculated components (see Figure 4).

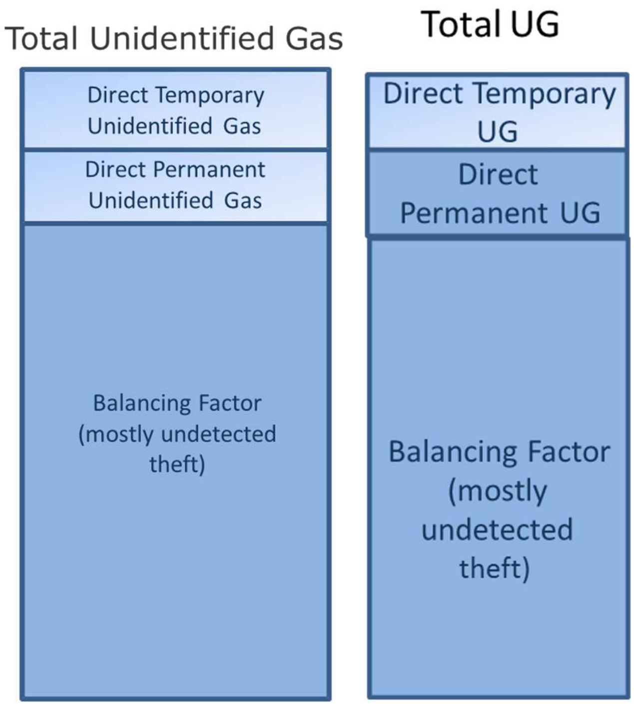

Figure 4: Unidentified Gas Components

4.3.3 Mapping to Post-Nexus Product Classes

The classification of Unidentified Gas into EUC/Product classes is referenced in several places in the descriptions above. The reason for having to create this split is described in Section 4.3.2. The need to do this creates a challenge, however, in that the majority of the training data comes from historic (i.e. pre-Nexus) years, during which time the Product Classes did not exist.

Therefore, a method is required for mapping sites as they were pre-Nexus to post-Nexus Product Classes. For the first AUG Statement of the new AUG period (i.e. the statement for 2017/18), the analysis took place before the implementation of Nexus and so all assignments of sites to Product Classes were handled using a rule set developed for this purpose. For subsequent years actual Product Class information is available for each site. This information can therefore be used to map sites to Product Classes individually, which provides far more accurate output. In some cases, this approach is not possible (e.g. Unregistered Sites, which by definition do not have a Product Class assigned): for these, the actual Product Class data can be used to create an approximate mapping that supersedes the rule sets developed for 2017/18 and provides a much higher level of accuracy.

The rules used to assign Product Classes to each site for each Unidentified Gas category are therefore as follows:

Detected Theft, Consumer Meter Errors, Volume Conversion Errors:

Map each site to its actual Product Class using the most recent data available.

Shipperless and Unregistered Sites, IGT CSEPs:

These sites do not have actual Product Classes recorded. Therefore, for each EUC, calculate the split by Product Class for all sites in the total population (for which data has been provided). This is done by number of sites to prevent unusual AQs from skewing the figures. For each EUC this creates a percentage split by Product Class. The next step is to use as many instances of this population dataset as are available to create a trend over time. Data currently exists for five points in time from Nexus go-live to the present. The calculated trend is then used to extrapolate the population split to the forecast year. Finally, this split is applied to the Unidentified Gas calculated as arising from each EUC to give the estimated Product Class split.

Balancing Factor:

The Balancing Factor is assumed to be largely composed of undetected theft, for which a direct Product Class mapping is not available. The calculation method for the UIG factors for the 2019/20 gas year contains a new method of estimating the split of undetected theft by EUC and Product Class. This is based on historic data for theft investigations and confirmed thefts and includes a newly developed technique for removing the effect of Supplier theft detection activity from these figures. The output resulting from this removal of bias from the confirmed theft figures is a set of theft levels by EUC/Product Class based on population and propensity to steal only – i.e. one that is representative of the split of undetected theft.

Population is a key element of this calculation, and the data used is necessarily from the past. Theft levels will change over time in proportion with any change in population of relevant categories of meter point, and hence in order to provide an accurate split by EUC/Product Class for the forecast year, the outputs from the analysis of historic theft data must then be extrapolated to this year.

In particular, theft levels from Smart Meters/AMR and traditional meters are different and so each of these populations must be extrapolated individually to the forecast year. Smart Meter installation rates and population totals are taken from the latest BEIS Quarterly Smart Meter Report [18], and these are used to calculate a best estimate of the Smart Meter populations for the midpoint of the forecast year based on the trend in installation rates and the current populations. These figures are used to estimate the Smart Meter/traditional meter mix in Product Class 4 for EUCs 01B-03B. The Standard Conditions of Gas Supply Licence [5] state that all sites in EUCs 04B and above must have an advanced meter. Therefore, for these EUCs, the entire population is assumed to be AMR. In addition, all sites in Product Classes 1-3 must also have a Smart Meter or AMR. AMR installation figures for Product Class 4 EUCs 02B and 03B are supplied by ICoSS. All of this information is used together to create a map of the population by meter type and Product/EUC for the forecast year, which is then used to extrapolate the calculated unbiased theft split to this point in time. This process is described in detail in Section 7.9.

4.3.4 Projection of Permanent Unidentified Gas to Forecast Period

Having calculated the best estimates of Permanent and Temporary Unidentified Gas for each historic year for which reliable data is available (the training period), it is then necessary to calculate the projected values of Permanent Unidentified Gas for the forecast year (see Figure 5). Note that the estimated values for the forecast year are calculated based on seasonal normal weather. The projection is carried out individually for each Unidentified Gas component category and EUC/Product class, in each case using the most suitable data and extrapolation technique. Extrapolation to the forecast period is carried out for each of:

- Shipperless and Unregistered
- IGT CSEPs
- Consumer Metering Errors
- Volume Conversion Errors
- Balancing Factor

The methods used differ based on the observed behaviour of each category of Unidentified Gas and are in many cases affected by a number of UNC modifications introduced in order to address various Unidentified Gas issues. The Balancing Factor is calculated for each of the five most recent historic years with reliable meter read data (2011/12 to 2015/16) and projected forward based on the pattern observed in this time period. Input data for the directly estimated components of Unidentified Gas is reliable throughout and so all available data is used. Properties of the Balancing Factor and full details of the extrapolation techniques used in all cases are described further in Section 7.

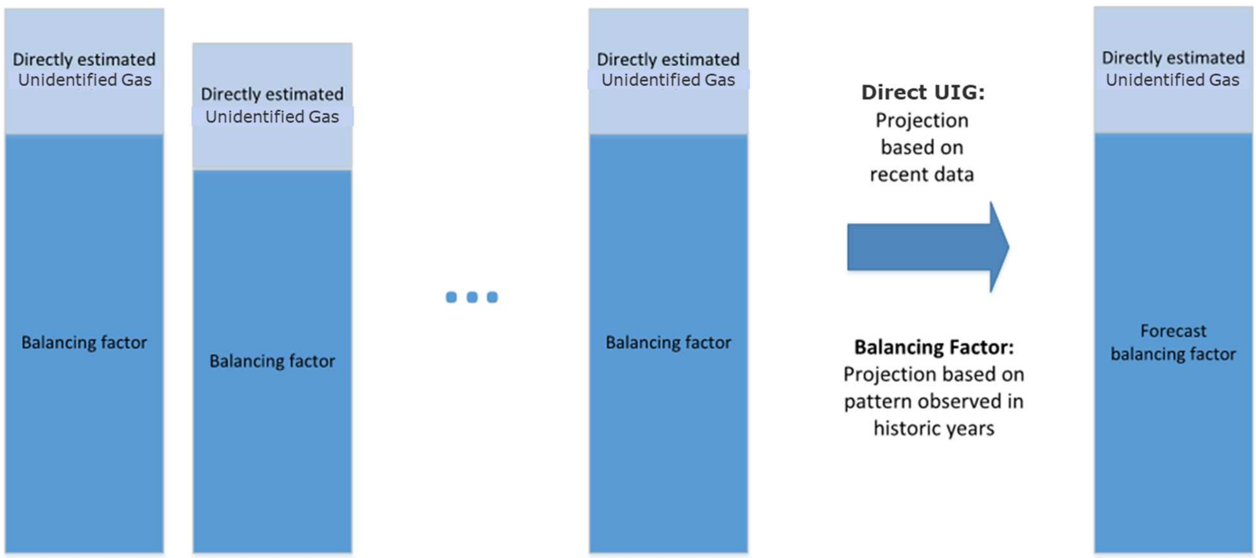

Figure 5: Projecting Unidentified Gas

As part of the estimation of the directly calculated Unidentified Gas components for the training years, an estimate of the amount of Temporary Unidentified Gas for each component is made as described above. The values projected forward to the forecast year are the permanent part of the Unidentified Gas only, however, for each EUC/Product combination. Note that detected theft (up to the 3-4 year cut-off date) is treated as a directly measured component of Unidentified Gas (100% Temporary and hence not taken forward to the forecast year).

4.3.5 Unidentified Gas Factors

The final output of the Unidentified Gas analysis is a set of UIG factors rather than direct estimates of the magnitude of Unidentified Gas itself. These factors can be applied to the population (defined in terms of the aggregate AQ for each EUC/Product) to give the relative magnitude of Unidentified Gas from each: these relative figures can then be applied to the independent daily UIG estimate to give the final breakdown (in energy terms) by EUC and Product Class.

The advantage of this approach is that this allows the effect of changing population to be taken account of in the UIG split without the need for the factors to change: when the number (and hence AQ) of sites for a particular EUC/Product category goes up or down, the fact that this AQ is then multiplied by the relevant factor ensures that the value of UIG from this source also goes up or down accordingly. This means that fixed factors can be generated that last a full year, until the results of the subsequent AUG Expert analysis become available, with the effects of changing population during that year still taken into account.

The factors themselves are a fundamental link between population and the Unidentified Gas from it, however, and so they must be calculated using the detailed estimates of the value of Unidentified Gas (for the year in which the factors will be in force) described above. Once the Unidentified Gas for each EUC/Product combination for the forecast year has been estimated, this is converted into a factor by dividing by the relevant aggregate AQ (i.e. the best estimate of the AQ for that EUC/Product combination for the forecast year):

$$UIG\ Factor_{PRODUCT,EUC} = Unidentified\ Gas\ (GWh)_{PRODUCT,EUC}\ /\ Aggregate\ AQ\ (TWh)_{PRODUCT,EUC} \times 100 \quad (4.3)$$

Note that in this equation the Unidentified Gas and Aggregate AQ have different units, and the raw output is also scaled up by a factor of 100. These steps ensure that the resulting factors give sufficient precision when expressed to 2 decimal places as required.

The final step in the UIG factor calculation is to apply a smoothing process. This is done on a Product by Product basis to ensure a smooth transition from each EUC to the next and avoid any step changes or up-and-down effects. The Unidentified Gas calculations involve a large amount of statistical modelling, and like any such models these will always contain a certain amount of common cause variation (i.e. natural variability that cannot be removed). This variability can, in some cases, cause the factors within a single Product Class to exhibit minor changes from one EUC to the next that fall outside the underlying pattern. A cubic fit (one per Product Class) is used to remove any such random effects and ensure a smooth pattern for the factors.

It should be noted that UIG factors are calculated by Product Class and EUC. As such they represent the average contribution to Unidentified Gas from any site within that Product Class and EUC. It is recognised that when a site changes Product Class and/or EUC it will be deemed to contribute a different quantity of Unidentified Gas as a result, even though there may be no physical change at the site. This is a consequence of the use of Product Class and EUC groupings, which creates an assumption that each site within any given group contributes Unidentified Gas in the typical manner for that group.

5 SUMMARY OF ANALYSIS

This section contains details of analyses carried out and updates made to the Unidentified Gas calculation methodology since the previous AUGS for 2018/19 was published. It also summarises new UNCC modifications and industry initiatives that could potentially impact on the Unidentified Gas calculation methodology in the future.

5.1 Theft Analysis

As referenced in Section 4.3.3 above, a new method has been developed for splitting undetected theft by EUC and Product Class based on the removal of the impact of Supplier detection strategy from the detected theft figures. This analysis requires detailed line-by-line data, not only for confirmed thefts but also for all investigations carried out (regardless of whether a theft was detected), and all leads (including ETTOS and TRAS qualified outliers) regardless of whether they were investigated. The additional line-by-line information on leads and investigations allows the effect of Supplier theft targeting to be estimated and removed.

As such, anonymised line by line theft records as provided by the Gas Suppliers (the contents of the TRAS Outcome files) were requested from SPAA and supplied to the AUGE. The data requested included the following fields for each lead/investigation/theft:

- (Dummy) MPRN
- Meter Serial Number
- EUC
- Product Class
- Meter Type (traditional/Smart/AMR)
- Meter installation date
- Source of lead (MAM, MRA, GT, TRAS, own analysis, tip-off)
- Lead investigated? (Yes/No)
- Theft detected? (Yes/No)
- Assessed Losses

Operational issues with the data request and its authorisation resulted in the data for the 2019/20 analysis being incomplete. Whilst all requested fields were present for the theft records supplied, only records involving actual investigations were included, and hence the following were missing:

- ETTOS leads that were not investigated
- TRAS Qualified Outliers that were not investigated

The loss of this data does not prevent the new theft analysis from being carried out, but it does limit its ability to remove the effects of Supplier theft detection targeting from the data. There are two steps in the theft detection process where bias is introduced by the Supplier's chosen theft strategy:

1. Generation of leads
   Leads from some sources result from a standard process applied equally across the whole population of meter point and hence are unbiased: they treat each meter point in the same manner and flag suspicious behaviour. The best example of this type of lead is the TRAS qualified outlier. Other leads, such as those generated by the Supplier's own analysis, are likely to be targeted towards certain elements of the population because of the composition of the Supplier's portfolio and any specific

theft detection strategy they are following. These leads are biased and would skew the undetected theft split.

2. Investigation of leads
   Each Supplier chooses what leads to follow up, and each has their own strategy. Some generate large numbers of their own leads and base their investigations on these, whilst others do relatively little of their own analysis and base their investigations on TRAS leads. The choices Suppliers make in what leads to investigate can again lead to bias towards certain market sectors.

The data as supplied for the 2019/20 factors (i.e. investigations only) allows the first type of bias to be removed but not the second. This is because we have the lead source for each and so those from unbiased sources (e.g. TRAS) can be included and those from biased sources (e.g. own analysis) can be discounted. The fact that we only have data from investigations and not the full set of leads they came from means that we have no information about the Suppliers' strategy for the selection of investigations (i.e. the Lead→Investigation step) and cannot remove bias from this source.

Despite this limitation, the new theft method is still a significant improvement on the old one, and the consultation responses were mostly supportive. However, some specific concerns were raised, particularly as not all required data was made available for calculation of the UIG factors. The method has been implemented, subject to the limitations of the unavailable data.

The method is described in full in Section 7.9 below.

5.2 Unidentified Gas from PC2 sites

Product Class two contains two distinct sub-populations of sites from which Unidentified Gas arises in different patterns and quantities:

- Sites that were previously non-mandatory DM
  These sites have been subject to compulsory transfer from Product Class 1 to Product Class 2 and now make up a large proportion of the total Product Class 2 population. It has been confirmed in the past that there has never been a documented theft incident from such sites, and so no theft (detected or undetected) should be assigned to them. In addition, such sites have a relatively high proportion of volume converters and so they attract relatively little Unidentified Gas from this source. The Balancing Factor (assumed to be mostly theft) accounts for the vast majority of the total Unidentified Gas, meaning that these sites should have only small amounts of Unidentified Gas assigned to them.

- Sites that were previously NDM
  Sites that have moved up from the NDM population rather than down from the DM population are subject to theft and are also less likely to have volume converters. They therefore attract more Unidentified Gas than those from the ex-DM sub-population.

The Unidentified Gas factors are split by EUC and Product Class, and therefore the factors for Product Class 2 should reflect the balance of the two sub-populations. This is an equivalent situation to Product Class 4, which is a mix of sites with Smart Meters and those without.

The asset data has been used to trace the origin of each site in Product Class 2, i.e. to classify it as ex-DM or ex-NDM. The following approaches are then used for each category of Unidentified Gas:

- Unregistered and Shipperless Sites, IGT CSEPs, Detected Theft, Undetected Theft (Balancing Factor)
  Ex-DM sites do not give rise to any Unidentified Gas, whilst ex-NDM sites do.
- Meter errors, Volume conversion errors
  These are carried out on a site-by-site basis and so calculations specific to each individual site are carried out.

The UIG factors that are the final output of the analysis represent the amount of Unidentified Gas per unit throughput (AQ), as described in Section 4.3.5 above. The AQ figures used as the denominator in this calculation for PC2 are always those for the full combined Product Class (i.e. the sum of the two sub-populations). This approach ensures that the final factors represent the correct amount of Unidentified Gas arising from this source, but also the full population of the Product Class across both sub-populations.

5.3 Replacement Values in Consumption Calculation

In previous years where the consumption calculation fails for a particular meter point, the average consumption for its EUC is used. This method was chosen because there is evidence that prevailing AQ values suffer from a bias. By using an average value, although it may not be accurate for a particular meter the differences should cancel each other out.

This year the AUGE has carried out analysis to look at the possibility of using another replacement value. Specifically, in the case where the consumption calculation fails for one year but is successful either side, using the average of these two values.

Initial investigation showed that this produced some suspect results. Following further analysis, the AUG Expert has decided to apply some additional validation:

- Tighten the AQ validation check from 5x to 4x the prevailing value
- Only use interpolation if the consumptions either side differ by less than 40%.

The impact of applying these new rules is summarised in Table 3:

Year	Number of meter points	Consumption change (GWh)
2012	41,453	122.98
2013	33,108	-134.70
2014	43,533	10.00

Table 3: Replacement Value Impact

The overall impact is minimal, but a significant amount of consumption is shifted between years 2012 and 2013. Only a small number of sites are affected as most failures don’t have a valid consumption either side.

5.4 Uncertainty in Conversion from Metered Volume to Energy

The process of converting metered gas volumes to energy is key within the gas industry, as settlement processes and billing are based on deemed energy usage. The process for conversion from volume to energy is covered by the "The Gas (Calculation of Thermal Energy) Regulations 1996" (GTER)[26]. Despite the importance of energy conversion, the vast majority of GB meters have their energy calculated using fixed factors. These factors may no longer be appropriate for the post-Nexus settlement regime.

The issues surrounding energy conversion have been investigated previously, mostly with regard to the impact on consumer billing. Ray Cope (an ex-director of the Gas Consumers Council) provided written evidence to the Energy and Climate Change Committee in Jan 2013[30] following a paper he had written in 2001 about the inaccuracies in gas billing.

In 2000, Ofgem published a consultation document[34] seeking comments about potential changes to the way energy is calculated. Ofgem subsequently contracted Dave Lander Consulting to carry out an assessment into the effect of assumptions used in energy conversion on domestic consumer billing[24]. In an open letter published in 2014, Ofgem concluded that "the prescribed method for converting metered volumes of gas into thermal energy remains appropriate"[44]. Dave Lander Consulting also raised issues around inaccuracies in energy conversion in a 2012 report for the Energy Market Issues for Biomethane Projects (EMIB) expert group[25].

BG Technology/Advantica carried out studies between 1998 and 2001 looking at how environmental factors affect metering accuracy which included an in-depth study of meter temperatures[35][36][37][38].

Temperature and pressure corrections were at one time included in LDZ Shrinkage. In 2000, a network code modification (NC0396)[31] was introduced which allowed for the temperature and pressure correction included in the LDZ shrinkage calculation to be reconciled through RbD. It also allowed for a retrospective correction following the results of the domestic meter temperature survey which was ongoing at the time. Initially, temperatures derived only from I&C sites were used in the calculation of LDZ shrinkage, but concerns about different temperatures in domestic premises had prompted the survey of domestic meter temperatures[35] (including a survey of internal vs external meters).

It was later deemed that LDZ shrinkage was not the appropriate place to deal with these corrections and they were removed from the shrinkage calculation from 2004 onwards. Northern Gas Network's final shrinkage report states "The continued removal of Temperature and Pressure correction greatly facilitates the establishment and operation of Distribution Network specific transportation charging formula (which is an Ofgem objective). For this reason, in 2004 Transco proposed to move to a regime that did not include Temperature and Pressure Correction. Northern Gas Networks concur with this proposal."[32].

Despite all of the analysis work carried out over a period of about 20 years, the approach to volume conversion in the GTER remains unchanged. Given the changes implemented to the gas settlement regime under project Nexus and the large-scale deployment of Smart meters, the AUG Expert believes an industry-wide review of the calculation of energy from metered volume is now warranted.

5.4.1 Technical Overview

Conversion from volume to energy is a two-stage process as given by equation 5.1 below. Firstly, the volume of gas measured (V_m) must be converted to standard conditions. The aim of this 'Volume Conversion' is to calculate what volume the amount of gas used would have occupied at 15 °C and 1013.25mbar (standard conditions) rather than the actual conditions at which the gas volume was measured at the meter. VCF in equation 5.1 is the Volume Conversion Factor which is used to convert to standard conditions.

The second stage of the process is to convert this standardised volume to an energy using the CV (which is known at standard conditions). This is represented by ECF (Energy Correction Factor) in equation 5.1.

$$Energy = V_m \times VCF \times ECF \qquad (5.1)$$

Where

V_m = Metered Volume

VCF = Volume Conversion Factor

ECF = Energy Conversion Factor

The VCF is in turn made up of three components representing pressure, temperature and compressibility corrections as given by equation 5.2.

$$VCF = \frac{P_m}{P_b} \times \frac{T_b}{T_m} \times \frac{Z_b}{Z_m} \qquad (5.2)$$

Where P is Pressure, T is temperature and Z is compressibility. The subscript m refers to the value at the meter and represents the conditions under which the volume measurement was made. The subscript b refers to base (standard) conditions.

The pressure and temperature corrections can be calculated as per equations 5.3 and 5.4. The compressibility factor is more complicated as it depends on pressure, temperature and gas composition. For meter pressures below 2 bar, the compressibility factor is close to 1.

$$CF_{temp} = \frac{T_b}{T_m} = \frac{(273.15+15)}{(273.15+T_{meter})} \qquad (5.3)$$

Where

CF_{temp} = Temperature Correction Factor

T_b = Temperature at standard conditions in K

T_m = Temperature of metered gas volume in K

T_{meter} = T_m converted to degrees Celsius

$$CF_{press} = \frac{P_m}{P_b} = \frac{(P_{msl}+P_{meter}+(A_{meter} \times ACF))}{P_b} \qquad (5.4)$$

Where:

CF_{press} is the Volume Correction Factor relating to gas pressure effects

P_m is the pressure of the gas at the meter when the volume was measured

P_b is the pressure of gas under base or standard conditions. This is 1013.25 mbar according to GTER[26]

P_{msl} is the atmospheric pressure at mean sea level (msl)

P_{meter} is the regulator set pressure i.e. the additional pressure of the gas over atmospheric pressure at the meter (after the regulator)

A_{meter} is the altitude of the meter installation above msl in m

ACF is the Altitude Correction Factor. This is taken to be -0.120208 mbar/m i.e. for every increase in altitude of 1m, the pressure will reduce by 0.120208 mbar.

Given the above, it is clear that conversion from metered volume to energy needs to account for local conditions at the meter (altitude and weather) in order to be accurate. In some cases, a volume conversion device is fitted which automatically converts the measured volume to the volume at standard conditions based on locally measured temperature and pressure. However, for most meters the conversion relies on factors which have fixed values and do not account for either geographical or day to day variations in pressure and temperature. They also do not allow for variations in average temperature and pressure between gas years. Larger consuming sites should have a site-specific correction factor calculated which takes account of the meter altitude and meter pressure.

Use of an inappropriate CF for volume conversion will result in UIG. This will be permanent Unidentified Gas as no adjustment is subsequently made to correct the volume based on a more accurate CF. However, when calculating the UIG factors, the AUG Expert is dealing with annual GB consumption under Seasonal Normal (SN) weather conditions. There will therefore only be an impact on Final (permanent) UIG over a year if the daily UIG values do not cancel out.

5.4.2 Presence of Volume Converters

In order to estimate the quantity of Unidentified Gas resulting from inaccurate volume conversion it is important to understand the amount of energy which can be affected i.e. the energy consumption of meters without volume conversion devices. The CDSP have provided the AUG Expert with an updated dataset containing a list of meters which have volume conversion devices since the Proposed AUGS. This data represents a significant improvement over the previous dataset.

The percentage of energy (based on AQ) for meters using volume conversion has been calculated by EUC and PC. This is shown in Table 4.

	PC1	PC2	PC3	PC4
01B	100.00	100.00	0.02	0.00
02B	100.00	100.00	0.08	0.15
03B	100.00	100.00	0.15	0.86
04B	100.00	100.00	2.33	6.79
05B	100.00	100.00	20.64	31.74
06B	100.00	100.00	29.14	57.99
07B	100.00	100.00	64.65	75.45
08B	100.00	100.00	77.68	80.79
09B	99.86	100.00	100.00	88.06
Overall	99.87%	100.00%	12.85%	6.51%

Table 4: Percentage of Energy from Meters with Volume Conversion Devices

Overall, 18.4% of total throughput energy is recorded by meters with volume conversion devices. Excluding PC1, 9.5% of energy is from meters with volume conversion.

There are no mandatory requirements on the fitting of volume converters. Historically, custom and practice has been to fit volume converters to meters consuming in excess of 2.93 GWh/annum[40]. The decision of whether to use a volume conversion device is driven by a commercial analysis of the benefits vs the costs involved.

5.4.3 Altitude Correction

As natural gas is compressible, the amount of gas (number of molecules) present in a given volume is dependent on pressure and temperature. It is therefore necessary to correct for this when converting from metered volume to energy as described in 5.4.1 above. The correction required for pressure can be considered as two separate corrections. One relates to the altitude of the meter installation as pressure varies linearly with altitude (the higher the altitude, the lower the pressure). The effect of altitude is discussed here.

Altitude will affect the calculation of CF_{press} which is calculated according to equation 5.4 described previously. Given that P_{msl} is ~1013 mbar, the effect of the altitude correction will be relatively small for all but the highest altitudes. It should also be noted that for higher pressure installations (larger P_{meter}), the effect of altitude will be even less significant.

The GTER[26] currently differentiates between meters which consume below 732,000 kWh/annum (EUCs 01-03) and those that consume more. For EUC 03 and below, a standard Correction Factor (CF_{std}) should be used for volume conversion. This CF_{std} is based on an assumption that the average altitude of a gas meter in the UK is 66m above mean sea level. This will lead to an error in the volume conversion for all meters where the altitude is not 66m, but over the UK as a whole there should be no bias, assuming that the 66m represents an accurate average altitude. However, there will be local geographical biases in the volume conversion, as different LDZs will have different average altitudes. This is investigated further below.

For meters which are in EUC 04 or above, GTER[26] states that a site-specific CF should be used. It is the responsibility of the Meter Asset Manager (MAM) to record the meter altitude and to calculate a site specific CF taking into account the altitude using the table provided in GTER[26] as per section 10.6 of the Code of Practice for Meter Asset Managers (MAMCoP)[27].

5.4.3.1 Unidentified Gas Resulting from the Assumption of Average Altitude

As discussed above, meters in EUCs below 04 use a standard CF which assumes an average altitude of 66m. The AUG Expert was asked to assess the likely magnitude of this assumption in terms of its effect on Unidentified Gas.

Freely available OS altitude data was identified[28] by the AUG Expert. This provides altitude (to the nearest 10m) by postcode for the UK. The CDSP mapped this data at the MPR level (98% of meters could be matched) and provided the AUG Expert with the anonymised meter level data. Using this data, the AUG Expert was able to recalculate consumptions for all meters for 2011-2015 using the actual altitude and to determine the impact of the 66m average altitude assumption currently used. The following assumptions should be noted:

- The analysis was based on those meters where a valid consumption could be calculated (>80% of all meters).
- The analysis was based only on those meters assigned the standard CF i.e. it was assumed that a non-standard CF meant that a site-specific CF including the actual altitude had been calculated. It should also be noted that some of the meters included may have volume converters fitted such that the standard CF is not actually used.
- A pressure of 21 mbar at the meter was assumed. No data on meter pressures was available but 21mbar is a good assumption for all domestic sites. It is unclear whether meters in EUCs 01 to 03 could be operating at higher pressures. However, if they are then the significance of any change in altitude will be even less.

Based on these assumptions, an average bias of -0.007% over GB for all 5 years was found. If this is scaled up to allow for meters where consumption was not calculated, the overall discrepancy in terms of energy is of the order of -16GWh/annum which is insignificant in terms of GB as a whole. It is expected that any discrepancy will reduce as standard CFs used for sites in 04B and above are updated to site specific factors.

However, as previously mentioned, this discrepancy will vary geographically. Figure 6 shows the effect of altitude for 2011 on an LDZ by LDZ basis. This will not impact the UIG factors as these are not split by LDZ, but gas consumers will be affected depending on their location and some gas suppliers/shippers may be affected if they have a portfolio concentrated in some areas.

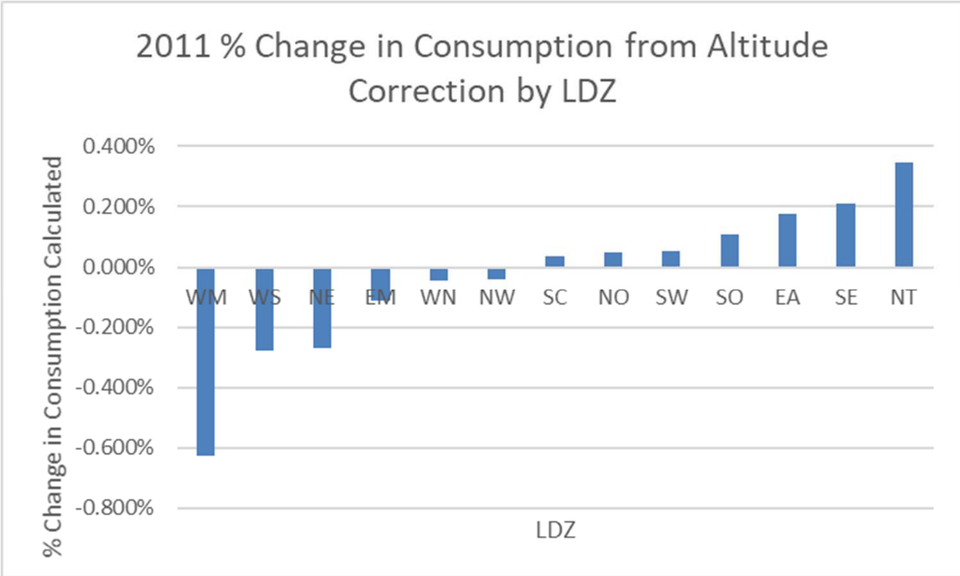

Figure 6: Effect of Altitude Correction on Energy Consumption Calculation by LDZ

The above analysis is based upon the change in pressure due to altitude as this is recognised as a contributing factor in the GTER[26]. However, there is also a relationship between temperature and altitude with temperature decreasing by ~0.65°C per 100m. It can be shown that this is about five times smaller than the effect of pressure and it is an opposing effect. Therefore, the overall (pressure & temperature) size of the altitude effect will be even smaller than quoted above. The AUG Expert therefore believes that no change to the current methodology is required to account for meter altitude.

5.4.4 Use of Static Correction Factors for Volume Conversion

All meters without a volume converter fitted (measuring gas pressure and temperature at the meter) rely on the use of a static volume conversion factor. In some cases, this factor is site-specific. Site-specific factors should take into account the altitude of the meter installation, the meter pressure and the gas composition. However, these static factors assume an average annual gas temperature and atmospheric pressure. Using static factors for volume conversion over an interval significantly different from a year will result in errors due to the daily variation in temperature and pressure.

Figure 7 below shows the sensitivity of the CF to changes in temperature and atmospheric pressure assuming a meter pressure of 21 mbar. It shows the percentage difference between the standard CF (1.02264) and the CF calculated using various combinations of temperature and atmospheric pressure. In cases of lower temperature and/or higher pressure (top-right of figure), the gas will be denser than assumed by the standard CF. The conversion of the metered volume to energy will therefore understate the energy content of the gas resulting in a positive contribution to Unidentified Gas.

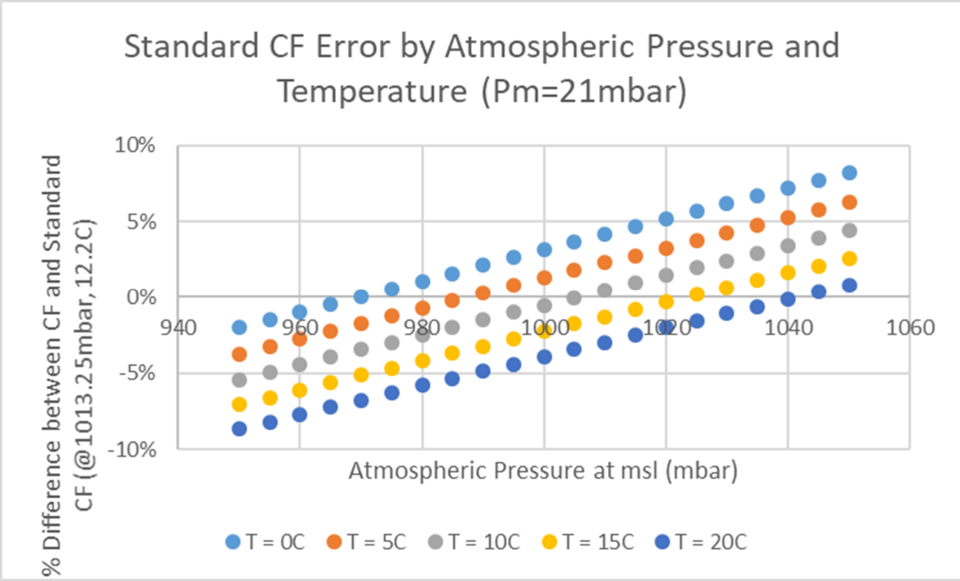

Figure 7: Variation in CF_{std} with Atmospheric Temperature & Pressure

Clearly, day to day changes in temperature and pressure can have a significant impact on Unidentified Gas. For temperature in particular, the contribution to Unidentified Gas will show a seasonal and geographical variation. The following sections look at the effect of pressure and temperature separately and make an order of magnitude estimate of their impact on Unidentified Gas.

5.4.4.1 Atmospheric Pressure

An initial sensitivity analysis was carried out by looking at data for Crownhill (Plymouth). The average daily atmospheric pressure over a 7 year period was found to be ~1016 mbar (assumed to be at MSL) which is slightly higher than the assumed standard conditions of 1013.25 mbar. This equates to a difference in CF of about 0.27%. Although daily pressures varied between 978 mbar and 1043 mbar, the extreme pressures are rare. Figure 8 shows the distribution of pressures at Crownhill 2011-2017.

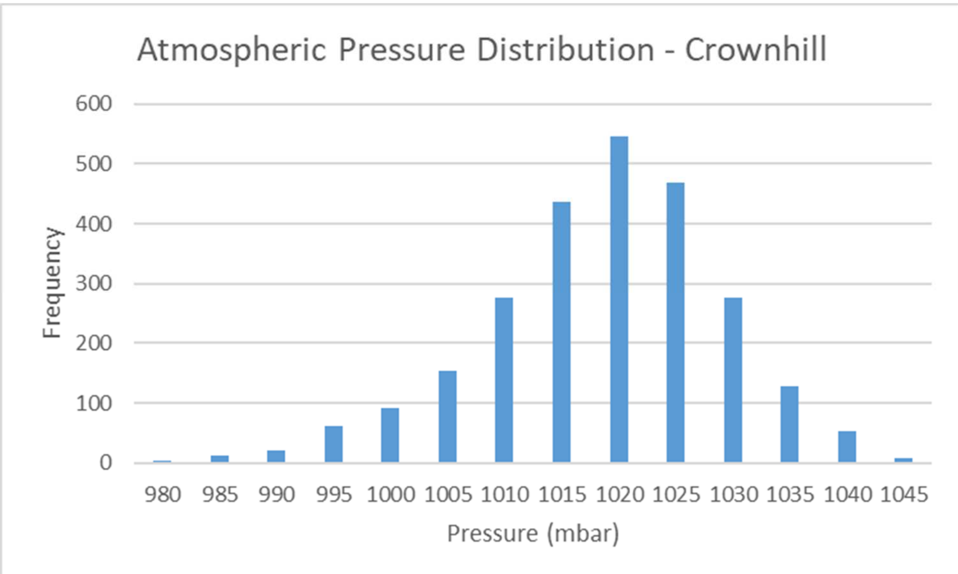

Figure 8: Atmospheric Pressure Distribution for Crownhill

This distribution is approximately normal with 95% of pressures falling in the range 995-1036 mbar. This range in pressures equates to an error in CF between -1.74% and 2.29%. Figure 9 below shows the monthly average atmospheric pressures for Crownhill 2011-2017. This shows that the average pressure varies little month to month and there is no seasonal pattern. This suggests that the use of a fixed average value for pressure is appropriate when looking at time frames of a month or longer. As the UIG factors are calculated using annual consumptions, they should be unaffected by atmospheric pressure variability, assuming that the average pressure of 1013.25 mbar used to calculate CF_{std} is a good estimate of average pressure for GB.

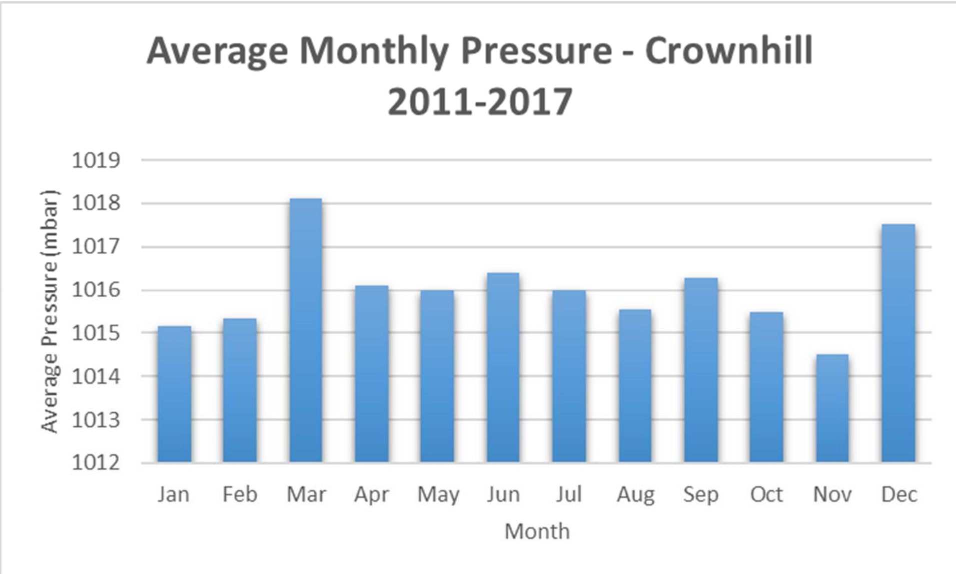

Figure 9: Average Pressure by Month for Crownhill, 2011-2017

In terms of daily initial UIG, day to day pressure variations will add to the volatility. Further work is recommended to assess the impact of pressure on initial UIG, but this falls outside of the current remit of the AUG Expert.

Based on the above analysis for a single location, the AUG Expert recommended that additional pressure data be obtained for a variety of geographical locations to validate the average pressure assumption. The CDSP have since provided hourly pressure data for 8 weather stations covering a period of 6 gas years (2012 to 2017). These measured pressures (simple average of all hourly values over the gas year) are shown in Table 5. The average atmospheric pressure across all locations and all gas years was found to be 1014.0 mbar. This is higher than the assumed average value of 1013.25 mbar and results in an under-estimate of the standard CF by 0.08%. This would result in a corresponding under-estimate of energy consumed and an over-estimate of Unidentified Gas.

Weather Station	2012	2013	2014	2015	2016	2017
COLESHILL	1013.4	1011.9	1015.1	1013.8	1017.0	1014.4
NOTTINGHAM WATNALL	1013.5	1011.9	1015.0	1013.7	1017.0	1014.3
FILTON	1013.6	1012.6	1016.0	1014.7	1017.6	1015.1
YEOVILTON	1014.0	1013.0	1016.4	1015.1	1017.9	1015.4
ST ATHAN	1013.6	1012.2	1015.8	1014.5	1017.4	1015.0
GLASGOW BISHOPTON	1011.5	1008.7	1011.6	1010.7	1014.6	1012.0
ALBEMARLE	1011.9	1009.6	1012.4	1011.3	1015.0	1012.4
HEATHROW	1013.9	1013.0	1016.1	1014.9	1017.7	1015.1
Average	**1013.2**	**1011.6**	**1014.8**	**1013.6**	**1016.8**	**1014.2**

Table 5: Atmospheric Pressure at MSL by Gas Year

Any error in the assumption of average atmospheric pressure will affect only meters without volume converters fitted. Based on 71% of annual energy being consumed at meters without volume conversion (See 5.4.2), the additional Unidentified Gas calculated as a result of assuming a fixed pressure of 1013.25 mbar could be up to 0.06%.

This Unidentified Gas estimate is based on the error in the standard CF only. For meters which operate at higher pressures (>21 mbar), the effect will be smaller. For meter pressures of 1,000 mbar for example, the error in CF approximately halves.

The methodology for the calculation of UIG factors has been updated to incorporate the effect of any bias in the assumed average atmospheric pressure.

5.4.4.2 Gas Temperature

Figure 10 is a schematic diagram to explain the effect of temperature on energy conversion and is not based on real data. The line labelled “LDZ Offtake Energy” (blue) represents a hypothetical annual profile of energy entering an LDZ. LDZ offtake meters measure volume, but also accurately measure and account for temperature, pressure and CV.

The line labelled “Sum of Consumer Energy” is again a hypothetical annual profile of energy but represents the sum of all consumer energy use within the LDZ. Each consumer will have a meter which measures the volume of gas used. However, in most cases energy is calculated by converting the measured volume to energy assuming a gas temperature of 12.2 °C. This will only give the correct energy value for gas at 12.2 °C. If the gas is colder than this, it will be denser and contain more energy per unit volume. In this case, energy will be underestimated resulting in positive Unidentified Gas. The converse is true when the gas is warmer than 12.2 °C.

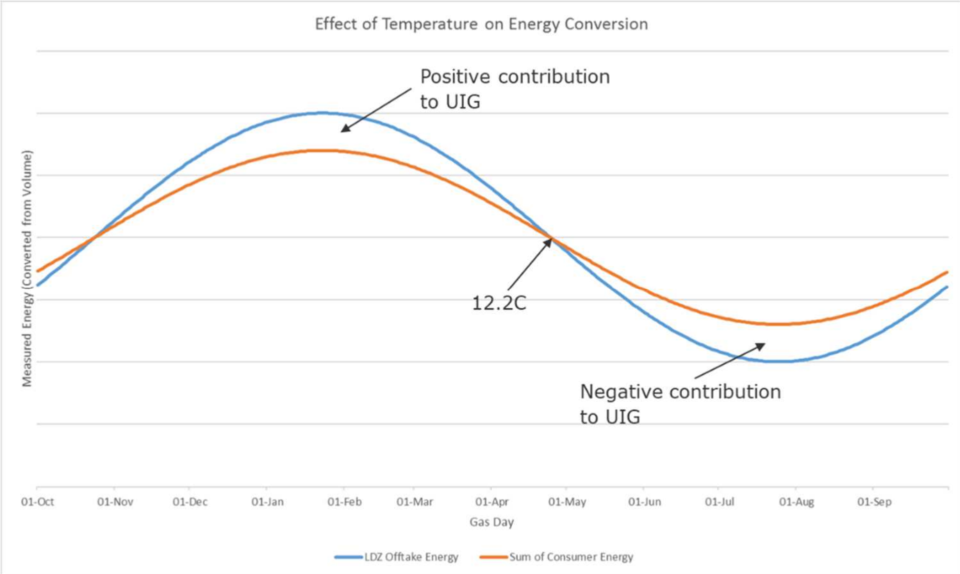

Figure 10: Example to Show Temperature Effect on Volume Conversion

When an energy value is calculated from a measured volume, the calculated energy will only be correct if the consumption weighted temperature over the period of the volume measurement is 12.2 °C. For example, if two meter reads one year apart are used, the conversion to energy should be accurate assuming that the average temperature (consumption weighted) over the year equals 12.2 °C. If monthly reads are used however, the consumption calculated in a winter month will be under-estimated and in summer will be over-estimated.

Following project Nexus implementation, Initial UIG is calculated on a daily basis and there is a drive towards obtaining more frequent meter reads. This move towards finer granularity consumption data and the use of rolling AQ calculations necessitates a review of how volume conversion should be performed. The AUG Expert believes that the use of a national fixed factor may no longer be appropriate.

For the calculation of the UIG factors, the AUG Expert is interested in the Final (total permanent) UIG over a gas year. In this case, the positive and negative effects of inaccurate volume conversion do not matter if they cancel out. The key consideration for the AUG Expert is therefore whether the assumed temperature of 12.2 °C represents a good consumption weighted average of SN temperature.

It is important to understand that the temperature referred to here is the temperature of the gas **in the meter**. The gas temperature is likely to start close to the ground temperature but will exchange heat with its surroundings as it travels through the service pipework into the meter. The gas temperature will therefore be highly dependent on local conditions including ground temperature (a function of air temperature, soil composition, depth of pipework etc.), air temperature, site exposure (to wind and sun), length of service pipe and location of meter (external or within a heated space). A number of studies have been carried out to assess the temperature of gas in a meter and these suggest that due to the relatively low flow rate and intermittency of flow there is a good correlation between the gas temperature and the air temperature in the vicinity of the meter[36][42].

Given that the current assumed average gas temperature of 12.2 °C is based on extensive field studies, the AUG Expert believes that this value is appropriate, or at least was when it was derived. However, some factors may have changed (e.g. increased average air and ground temperatures, changes to the distributions of meter locations etc.) which would warrant a review of the assumed value. It would also seem appropriate to consider the use of local factors to account for geographical variations in temperature and different factors by EUC to account for the possibility of more internally located meters in some cases (leading to higher gas temperatures).

As the AUG Expert does not have access to detailed gas meter temperature data which would allow a review of the assumed average value, it is not appropriate to propose a change to it. However, the methodology has been updated to account for a bias in the assumed average temperature should one be proven at a later date. This bias will be set to zero for the calculation of UIG factors this year.

Although the AUG Expert does not have the data to justify an alternative value or assess the current value's potential bias, sensitivity analysis shows that a bias in the assumed average temperature of 1°C equates to Unidentified Gas of ~1.4 TWh (taking into account meters with volume conversion devices fitted). It should be noted that the AUG factors are estimated assuming SN temperature. There can therefore be significant variations in UIG from year to year depending how close the average gas temperature is to that assumed under SN conditions.

The AUG Expert has also identified a number of sources of relevant temperature data which are provided in Appendix C for information. The wide range of values highlights the uncertainty in the temperature.

The AUG Expert has looked at a number of scenarios based on assumptions around temperature. It has been found that the average gas temperature is very sensitive to assumptions regarding the temperature of gas in internal meters. Two scenarios are used demonstrate this. In both cases the following assumptions are made

- Average gas temperature in external meters is 9°C (DMTS measured this as 9.07°C)
- 39% of meters are internally located, 51.5% external and 9.5% in sheltered locations (based on data provided by CDSP)
- The temperature for sheltered meters is assumed to be the average of internal and external temperatures

Scenario 1:

Using the measured internal temperatures from the DMTS (See Appendix C) of 14.72°C with the above assumptions leads to an average gas temperature of 11.5°C.

Scenario 2:

Based on an assumption that the gas temperature for internal meters is equal to the internal building temperature minus 2°C[42] and the average internal temperature is 19°C[56] together with the above assumptions leads to an average gas temperature of 12.5°C.

The AUG Expert believes that both of these scenarios represent realistic assumptions but lead to a wide range in average temperature. Given the sensitivity of Unidentified Gas to temperature, further work is therefore required to understand and resolve this issue.

5.4.5 Other Correction Factor Issues

The GTER specifies that all meters with a consumption which can reasonably be expected to exceed 732,000 kWh (EUC04B and above) should have a site-specific CF calculated by the MAM for use in volume conversion. For meters below this threshold, the standard CF of 1.02264 should be used. Processes are in place to flag meters which cross the 732,000 kWh AQ threshold in either direction to the appropriate shipper.

A UK Link change, "Correction Factor Application"[29] was raised in 2015 which would have automated the process of switching between standard and site-specific CF values when meters crossed the 04B AQ threshold. This change was not implemented in full. The only change which was implemented was to set the CF to the default value above the 732,000 kWh threshold rather than zero which was used previously.

The GTER also allows for the use of a site-specific pressure only CF where a temperature only conversion device is used on the meter. This was discussed at the AUG Technical Workgroup Early Engagement meeting. Neither the CDSP nor any other attendees were aware of temperature only conversion devices being used. The AUG Expert is therefore assuming that there aren't any.

An incorrect CF will cause an error when meter reads are used to calculate energy consumed. The AUG Expert has therefore looked at a range of issues related to the use of CFs.

5.4.5.1 Incorrect Site-Specific Correction Factors

The AUG Expert has carried out an analysis of the latest asset data to identify unusual CF values. These fall into a number of categories. Note that some meters may fall into more than one category.

A list of meters with suspect CF values has been provided to the CDSP for further investigation.

Unusually Low CF Values

Domestic consumers will have meters which operate at a pressure of 21 mbar. This is assumed in the calculation of the standard CF and all meters consuming <732,000 kWh/annum should use the standard CF for calculation of energy. The AUG Expert has therefore calculated the range of CF values which would correspond to 21 mbar operating meter pressure across the range of possible meter altitudes assuming the base temperature of 12.2 °C. This gives a range of CF values from 0.97922 to 1.03193. Any CF values falling below the lower value are likely to be incorrect.

There are 10 instances with CF values below this threshold as follows

- One which can be discounted as it has a volume converter fitted and is in EUC 09B
- Three which appear to be incorrectly entered versions of the standard CF (0.02264, 0.11264 and 0.12264)
- Five values of 0.000001 which are essentially zero. This will lead to a calculation of zero consumption and associated knock-on effects to AQ calculation
- One meter with a value of 0.013058

Apart from the meter in EUC 09B which has a volume converter fitted, all other meters except one (in 03B) are in 01B. Two of the meters have an AQ of 1 which is probably the result of calculating AQ based on the incorrect CF.

The AUG Expert has assessed the potential contribution to Unidentified Gas from these meters by calculating what the consumption would have been based on using the AQ and standard CF. For the 2 meters with AQ = 1, an assumed consumption for 01B of 13,500 kWh was used. The overall contribution to Unidentified Gas from all affected meters was found to be <1 GWh. From a UIG perspective, this is insignificant as it affects such a small number of meters, but these factors should be corrected. The AUG Expert recommends that anomalously low CFs are flagged to shippers for investigation on an ongoing basis to ensure that the number of meters affected remains low.

Unusually High CF Values

It is more difficult to identify meters with a CF which is too high. A minimum CF value can be calculated assuming 21 mbar as a lower bound for meter pressure. For higher CF values it is impossible to know if the high CF value is correct and is the result of higher meter set pressure or if it has been calculated or entered incorrectly. To validate CF values, the meter set pressure data (recorded by the MAMs and provided to suppliers) would be required. This data is not held by the CDSP. Figure 11 shows the relationship between CF and meter set pressure assuming a fixed temperature, atmospheric pressure and gas composition. The graph covers pressures up to 80bar as this equates to a CF of ~100 which is the maximum value currently in use.

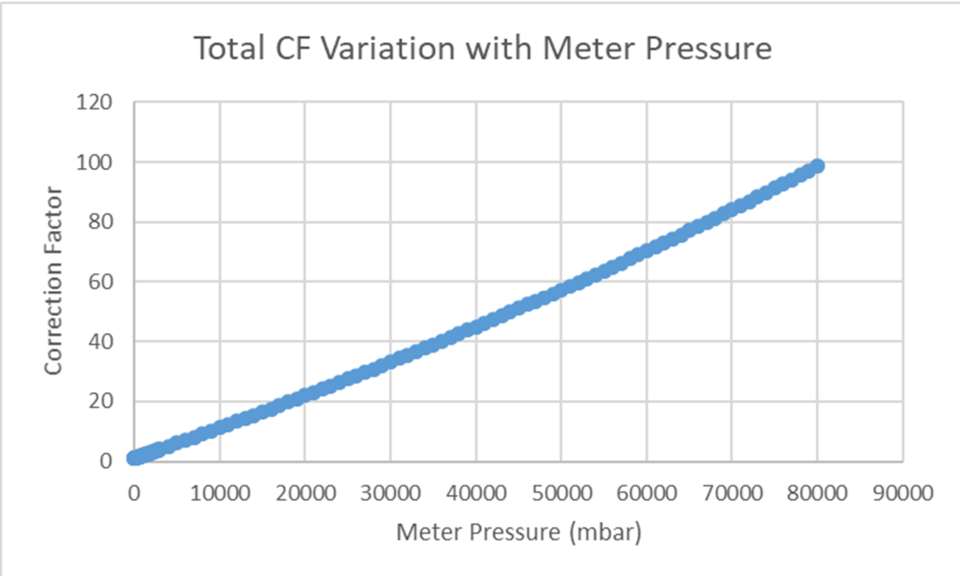

Figure 11: Variation of Volume Correction Factor with Meter Pressure

Notwithstanding the above, the AUG Expert has identified some meters with high CF values which require further investigation.

Meters which fall within EUC 01B would not be expected to be operating at high set pressures. A CF value of 1.5 equates to ~500 mbar set pressure. There are 8 meters without volume converters in EUC 01B which have a CF in excess of 1.5.

Additionally, there is one meter with a CF set to 100. This is the largest CF in the dataset and is more than double the next largest value. It is also for a meter in EUC 02B and so should be using the standard CF. The energy consumption for this meter will therefore be over-stated by a factor of ~100. The CDSP have recently confirmed that this CF is incorrect and has been updated by the shipper. The CDSP will be providing updated meter asset data to the AUG Expert to reflect all CF updates.

The small number of meters identified as potentially having an anomalously high CF is small so the effect on Unidentified Gas will be negligible. A list of affected meters has been sent to the CDSP for further investigation.

Variations of the Standard CF

A number of examples of CF values which are very similar to the standard CF have been identified which are likely to be the result of typing errors or entering the values to a different number of decimal places. The number of meters affected is small and in most cases the CF value used is close to the standard CF value (anomalously low and high CF values are covered separately). The effect of these errors will therefore be insignificant.

Use of Integer Values for CF

It is unlikely that the calculation of CF would result in an integer value. Currently there are 2,085 meters with an integer CF value. Of these, 1093 have volume converters. In these cases, the CF value will not be used.

Of the remaining 992, most (989) have a value of 1. This leaves just 3 meters with integer CF values. Details of these have been sent to the CDSP for further investigation. From an Unidentified Gas perspective, the impact will be negligible due to the small number of meters affected.

For meters with a CF of 1, assuming that the CF used should be the standard CF, this would result in an under-estimate of energy consumption of 2.21% for these meters. This will have negligible impact given the small error and low number of meters affected.

5.4.5.2 Incorrect Use of Standard Correction Factors

The GTER mandates that all meters reasonably expected to consume in excess of 732,000 kWh (EUC 04 or higher) have a site-specific CF. Analysis carried out for the Proposed AUGS estimated a potential impact of using standard CFs rather than site specific CFs of around 136GWh. The AUG Expert also assumed that this would be wholly temporary as the issue was raised at the MAMCoP forum where the MAMs confirmed that site specific factors would be provided. This assumption was challenged during the consultation process, so the AUG Expert has revisited this.

Since the original estimate of 136GWh was made, the CDSP have provided the AUG Expert with updated data on the presence of volume converters and information on CFs which have been updated since 1 June 2018.

Based on the latest data, it has been found that 4,023 sites were assigned (prior to 1 June 2018) a standard CF where a site-specific CF should be used. The total AQ for these sites is 6,288GWh.

Looking at the CF updates from 1 June 2018 onwards, there have been 1,766 (43.9%) updated from the standard CF to a site-specific CF (only sites without volume conversion). Based on the AQs of these sites, the use of the site-specific CFs will reduce UIG by 46.76GWh/annum. Assuming that the corrected sites are representative, an estimate of the UIG resulting from the sites which are still using standard CFs can be made by scaling up. This suggests a remaining 59.77GWh of UIG.

As the process of updating CFs is ongoing, the level of UIG is likely to reduce further before the 2019/20 gas year for which the AUG factors apply. Over a period of 8.5months, 43.9% of sites have had an updated CF. It is reasonable to assume that of the remaining sites, a similar proportion will be updated over the 8.5months prior to the start of the gas year. Based on this assumption, an estimate of the permanent UIG for 2019/20 can be derived as 33.53GWh.

5.4.5.3 Incorrect Use of Site-Specific Correction Factors

The GTER mandates that all meters reasonably expected to consume in excess of 732,000 kWh (EUC 04 or higher) have a site-specific CF, but that all others use the standard CF value (1.02264). The latest asset data contains 40,689 meters with consumption below 732,000 kWh/annum (EUC 01-03) with non-standard CF. Of these, 46 have volume converters fitted.

Settlement calculations will use energy based on the non-standard CF values. This may lead to a more accurate energy consumption if a correctly calculated site-specific CF is being used. However, it may also be the case that the CF has been entered incorrectly and is wrong or that the site has undergone a reduction in consumption and is still using the site-specific factor despite falling below the 732,000 kWh/annum threshold. Whatever the reason, the energy calculation is not being carried out consistent with the GTER.

The AUG Expert has carried out analysis to assess the magnitude of this discrepancy. Excluding sites with anomalously low CF values (<0.97922 as these are treated separately), the AQ was scaled by the ratio of the actual CF to the standard CF. The difference was found to be 2.1 GWh/annum which is insignificant.

It should be noted that the AUG Expert's consumption calculations use the standard CF for all meters in 01B.

5.4.6 Accuracy of Volume Converters

The use of volume converters is covered by a number of standards. IGEM/GM5[40] covers design, installation, commissioning, validation and the ongoing monitoring, inspection and maintenance of such devices. If this standard is followed, then validation will ensure that the device is measuring correctly at the time of installation and regular inspections should be carried out. This standard specifies a required accuracy during validation and testing of +/-0.5% of reading for pressure and +/-2 °C for temperature. These accuracy requirements are somewhat higher than the device accuracy requirements (see below) to allow for in-situ measurement difficulties.

Volume conversion devices should comply with BS EN 12405[41]. This specifies a Maximum Permissible Error (MPE) of the device under operating conditions as follows.

- O.1% for temperature (~0.3 °C)
- 0.2% for pressure
- 0.5% overall device error (including calculation of Z)

Assuming that the appropriate standards are adhered to, there is no reason to assume a bias in volume conversion devices.

5.4.7 Accuracy of Calorific Value (CV) for Energy Conversion

The accuracy of CV has been investigated previously by Dave Lander Consulting. In 2012, a report was written for the Energy Market Issues for Biomethane Projects (EMIB) expert group regarding the accuracy of CV determination[25]. Among the conclusions, it was noted that

- "Current custom and practice is for Ofgem to require that (absolute) error in CV measured by CV determination equipment should not exceed 0.10 MJ/m^3. This requirement results in insignificant impact on domestic energy metering."
- "The dominant sources of bias and uncertainty in bias are associated with fixed factors for conversion of actual domestic metered volume to reference temperature and pressure."

The report then proposes a relaxation of the Maximum Permissible Error (MPE) of CV determination devices for biomethane flows into gas distribution systems as this would result in no material impact on the accuracy of FWACV.

Another report was written by Dave Lander Consulting for Ofgem in 2014 looking further at the accuracy of gas energy measurement[24]. This concludes that the uncertainty in the daily charging area CV is around +/-0.08 MJ/m^3, or 0.2%.

Based on this, the AUG Expert believes that any contribution to Unidentified Gas from CV inaccuracies will be small, and more importantly unbiased.

5.5 Changes to Product Class Populations

In order to calculate the UIG factors accurately for the forecast year, it is necessary to forecast both Unidentified Gas and the Product Class populations (number of sites and AQ) to this point in time. The methods used to calculate and forecast Unidentified Gas are documented in detail throughout Section 7 of this document. For the populations, the asset data provided by Xoserve is used.

This asset data has been provided for a number of points in time, from Nexus go-live to the most recent data set available, which is for February 2019. Each data set allows the population in terms of both the number of sites and the AQ to be calculated and split across EUCs and Product Classes for that point in time, and is based on the EUC and PC definition for each individual site. Data is also available for CSEPs (this was previously on an aggregate basis but is now on a site-by-site basis), and this allows the CSEP population to also be included in these calculations.

Instances of this asset data exist for a number of points in time, which allows a trend to be established and used as a basis for the forecast. The trends are calculated at the Product Class level, and the forecast Product Class totals split by EUC using the split from the most recent instance of the asset data. The following rules are applied during this process:

- Values are projected to the midpoint of the forecast year, in this case April 2020.
- Trends are not necessarily calculated across the whole of the time period for which instances of the asset data are available. The rates of change of Product Class populations have shown significant variation, due to:
    - Product Class take-up rates changing as the new regime becomes more established.
    - Compulsory assignment of non-mandatory DM sites to PC2. A similar situation will arise where sites with insufficient meter read submissions are assigned to PC4 if Mod 664 is implemented.

Therefore, the most appropriate training period is used to avoid any negative impact of these step changes and to ensure that the forecast method to April 2020 is reliable.

5.6 Reconciliation Analysis

Although the AUG Expert is not required to calculate total UIG, a value for UIF(f) (i.e. the figure at line in the sand) is required to calculate the UIG factors using the current methodology, and this value is included in the AUG Statement each year. There have been a number of questions regarding the value calculated by the AUG Expert, particularly with regard to whether it is consistent with the values currently coming from the daily UIG process.

The initial daily UIG values calculated under Nexus are estimates, based on the best available data on the day. Over time, as more information becomes available (e.g. consumer meter reads, LDZ offtake meter error estimates, updates to meter asset data etc.) corrections are made through reconciliation. The Total UIG calculated by the AUG Expert is an estimate of the permanent UIG remaining at line in the sand i.e. after all reconciliations. The line in the sand is reset each year to 1st April 3 years previously. There is therefore a period of 3-4 years over which reconciliations can occur (and Mod0429 allows certain adjustments after line in the sand).

In theory, over the long term the daily UIG and subsequent reconciliations should give an indication of the level of permanent Unidentified Gas. This process is still subject to a considerable degree of error, however, despite the fact that the reconciliations provide additional information and accuracy over and above the initial UIG estimate. In particular, reconciliations represent corrections applied in any given month. The actual reconciliation value however, may be for a period up to 4 years ago. Therefore, whilst it may be possible to make inferences about total permanent UIG based on the long-term average UIG and reconciliations, this analysis would need to cover a minimum of 4 years, and values for individual years and individual months should be treated with extreme caution. At present, the relatively small quantity of post-Nexus reconciliation data makes reliable analysis difficult.

The CDSP has analysed and presented results from analysis of post-Nexus reconciliation data [45]. This shows that for around a year, reconciliation reduced average UIG from 4.65% at allocation to 4% at reconciliation. After this point, the level of initial UIG dropped due to combination of data improvements and the introduction of ALP/DAF uplift factors in October 2018. Due to the combination of these (alongside other naturally-occurring variation in the process), from May to November 2018 the average initial UIG was -0.2%, which has reconciled up to approximately 1%. Both the 4% (up to April 2018) and the 1% (May 2018 onwards) are based on incomplete reconciliation and it is as yet unclear whether the two will meet in the middle when reconciliation is complete.

It is important to note that the split of reconciliations between pre-Nexus and post-Nexus is only approximate. Any offline corrections (LDZ offtake metering errors and DM errors) are also not included in the reconciliations. Given that this analysis only covers a period of 18 months post-Nexus, there will be further reconciliations over the next 3 years to be applied before final figures for final UIG can be assessed.

The AUG Expert's estimate of permanent Final Unidentified Gas is purposely based on historic data, and as such will include a large proportion of the reconciliations that will be applied. Each year, the AUG Expert makes a decision on which data to include for the calculation of total Unidentified Gas (and therefore Balancing Factor), with this decision being based on the reliability of the results produced by each year’s raw data. Unidentified Gas is a small difference between two very large numbers (i.e. metered input to the system and metered output), and hence it is very susceptible to changes in these numbers. The years used in the AUG Expert’s analysis have a success rate (i.e. successful calculation of consumption from meter reads) of approximately 90%. This is the reason that recent years are discounted: these have fewer meter reads available and hence a lower success rate, and this has been found to create unreliable output.

In order to maintain the integrity of the AUG Expert's estimate of permanent Final Unidentified Gas, a stringent error checking procedure is used over and above that used to class any given meter read as valid for billing purposes. As such, the ≈90% success rate referenced above applies to some years that are fully reconciled, with the difference due to the additional error checking. It can therefore be seen that the use of all "valid" meter reads used for reconciliation in the UIG calculation will lead to errors that the AUG Expert process removes, and any level of reconciliation below full reconciliation will also lead to potentially significant variation in the UIG estimate.

The UIG factors for 2019/20 are based on gas years 2011 to 2015 shown by the box in Figure 12 below. At the point of calculation, the line in the sand was 1 April 2015 i.e. mid-way through the 2014/15 gas year. Over half of the data used by the AUG Expert is therefore fully reconciled. The remainder will have included a significant amount of reconciliation. In addition to reconciliation resulting from meter reads, the data used by the AUG Expert will also include other corrections which may take longer to be identified, quantified and reconciled such as LDZ offtake metering errors and thefts.

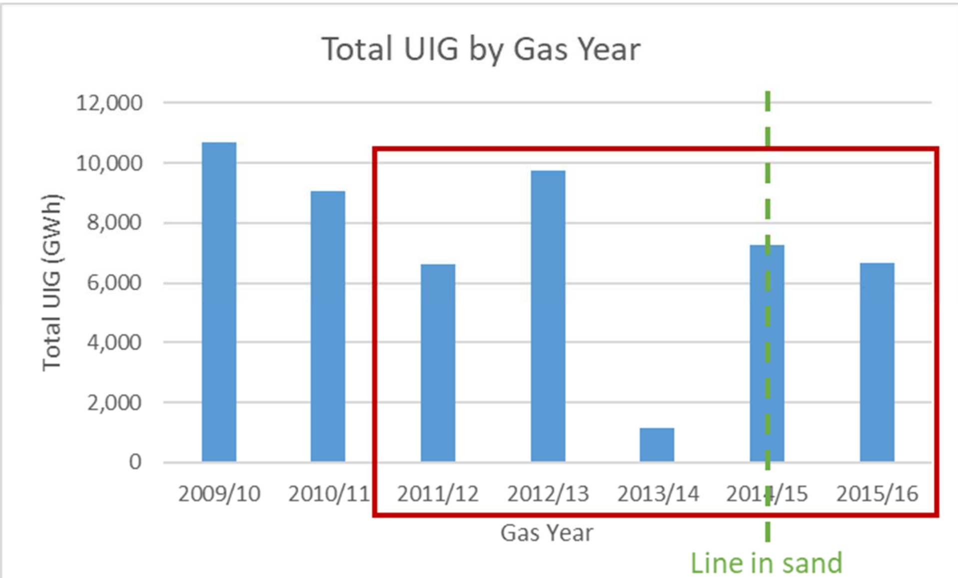

Figure 12: Historic Data Used in Total Unidentified Gas Estimation

The issues highlighted above mean that Unidentified Gas is a long-term concept. The AUGE's figures, despite being corrected to seasonal normal conditions and being subject to the rigorous validation procedures described above, still vary between 0.11% and 2.02% of throughput for individual years. The value only evens out across a number of years.

Therefore, whilst the CDSP analysis provides useful information, it is not a like-for-like comparison with the AUGE's UIG(f) figure for the following reasons:

- The analysis covers a time period that is different from the ones the AUG Expert uses since the reconciliation level (proportion of meter reads received) is too low.
- The meter reads used for reconciliation have been subjected to a lower level of validation than used by the AUG Expert.

- The effects of any meter errors that will be detected up to (and potentially beyond) line in the sand are not included.

As such, even when line-in-the-sand has been reached, there may not be a 1:1 relationship between UIG(f) and the AUGE's estimate of Unidentified Gas. Such an analysis still has the potential to provide useful information, however, and so both pre- and post-reconciliation UIG should continue to be monitored to see whether the best estimate of UIG(f) remains around its current level of around 1%, or whether it moves upwards towards the 4% suggested from the early post-Nexus months.

Figure 13 below shows the CDSP figures for UIG as a percentage of total throughput. The drop in both initial UIG and latest UIG around May 2018 can be clearly seen, along with the initial level around 4% for post-reconciliation UIG and the new level around 1%.

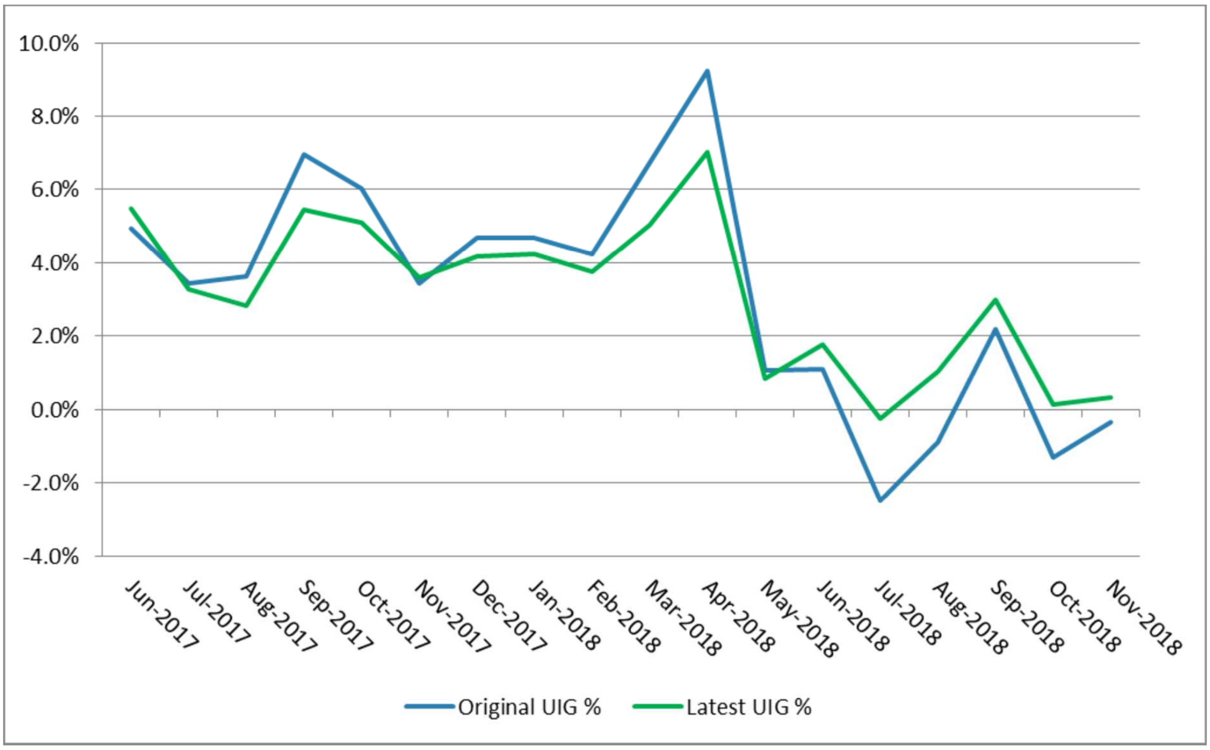

Figure 13: UIG as a % of Throughput (CDSP Analysis)

It is worth noting that recent levels of initial UIG have been much lower than those seen immediately post-Nexus. Over the period 1 June 2017 to 31 January 2018, average initial UIG was +4.6%. A year later, over the period 1 June 2018 to 31 January 2019, the average initial UIG was -0.4%. This figure will have been reduced due to the introduction of ALP/DAF scaling from 1 October 2018. Accounting for the effect of ALP/DAF scaling (applying a 3% uplift to UIG from 1 October 2018 onwards) still only increases initial UIG to 1.1% of throughput.

5.6.1 Uncertainty in AUG Expert estimate of Total Unidentified Gas

Given industry concerns about the discrepancies in the level of total Unidentified Gas, the AUG Expert has identified what it believes are the largest sources of uncertainty in their estimate. These are the choice of historic years used in the calculation and the consumption estimate for CSEPs. These are discussed in turn below.

5.6.1.1 Choice of years for Unidentified Gas calculation

The AUG Expert calculates total Unidentified Gas for a number of historic gas years. The most recent year is excluded due to the lower success rate of the consumption calculation (many sites with annual meter read frequency do not have acceptable reads). Each year, the AUG Expert makes a decision on which data to include for the calculation of total Unidentified Gas (and therefore Balancing Factor). The UIG factors for 2018/19 were based on gas years 2011 to 2015 shown by the box in Figure 12.

In choosing which years to use, the AUG Expert must consider the following:

- Relevance of data – more recent data is more likely to be representative of the current position. In particular, it will include UNC modifications introduced with a view to reducing UIG.
- Reconciliation – more recent data is subject to further update. Although a large number of reconciliations occur during the first year, this does not mean that most of the energy has been reconciled. For example, the AUG Expert's consumption calculation for gas year 2011/12 showed total Unidentified Gas of 1,324GWh when calculating factors for the 2015/16 gas year. When the calculation was repeated two years later, the total Unidentified Gas estimate had increased to 6,403GWh. A number of factors (e.g. data updates and methodology changes) will impact the total Unidentified Gas calculation but some of this change will be due to reconciliations being processed immediately prior to line in the sand
- Data quality – older data may be of lower data quality as data validation would be expected to have improved and there was a data cleansing exercise carried out prior to project Nexus.

Clearly, the choice of years to use in the calculation can have a significant impact on the calculation. The AUG Expert will continue to review the choice of data window each AUG year. As more data becomes available, patterns may become apparent that may change the way the historic years are chosen.

5.6.1.2 CSEP Consumption

All meter reads and volumes used to calculate total Unidentified Gas is currently from pre-Nexus. No information about individual meter points within CSEPs is available from this period, only AQs and meter point counts. If there is a significant bias in the CSEP AQ, this will have a knock-on effect in terms of estimated Unidentified Gas.

For gas years Post-Nexus, the AUG Expert will have access to individual CSEP meter point data. It is proposed that as part of the analysis for the next AUG year, an assessment is made of the bias in pre-Nexus AQs for CSEPs.

5.7 Unidentified Gas Resulting from Meter Exchange

As a result of the SMART meter rollout programme, a significant proportion of the gas meter population is being replaced each year. Figure 14 shows the number of meters being replaced each year based on asset data provided by the CDSP. From 2015 onwards, there is clearly an upward trend with about 10% of meters being replaced during 2017. Note, the value for 2018 is for a part year. Given the large number of meter replacements, there is significant potential for the creation of Unidentified Gas.

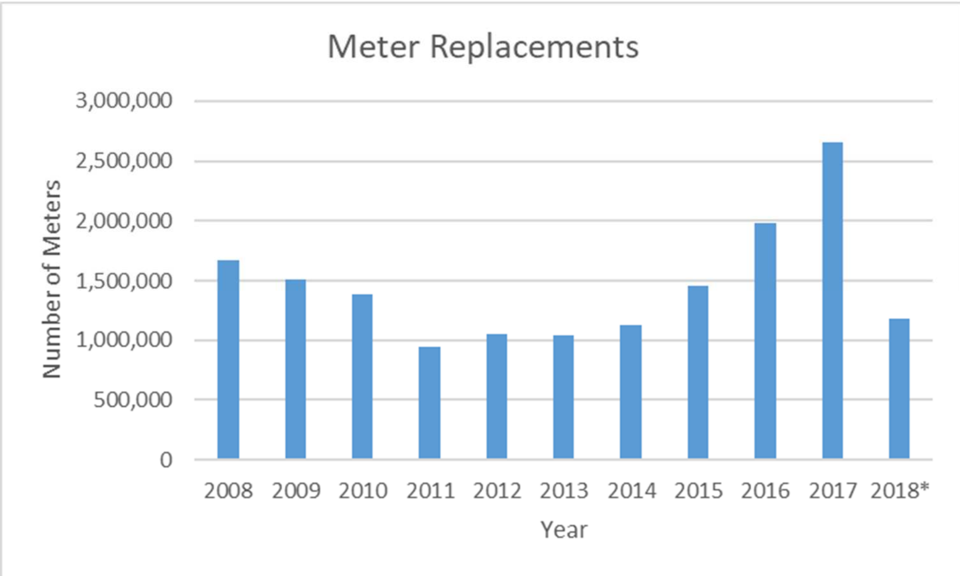

Figure 14: Number of Gas Meter Replacements by Year

A sample of meter replacements have been investigated in more detail to identify if there are any potential sources of Unidentified Gas. Examples of volumes being incorrectly calculated have been identified under the following circumstances

- Closing read but no opening read – when the new meter read is lower than the previous meter's closing read, a negative volume can be calculated
- An opening read but no closing read – the period between the last read of the old meter and the opening read of the new meter does not have a volume calculated
- Opening and closing reads with gap between – no volume calculated between reads. The example identified had ~3 months between the closing and opening reads
- Opening reads incorrectly flagged as final reads
- Opening and final reads out of sequence

These issues have been identified on a small sample of data and the AUG Expert is currently working with the CDSP to determine if these issues are present in the source data or if the issue is with the data provided to the AUG Expert. Further investigation is required.

5.8 Discrepancies between Converted and Unconverted Meter Reads

During Nexus implementation, the validation between the meter index and the unconverted converter index was suspended[33]. The AUG Expert is currently discussing how this could impact Unidentified Gas with the CDSP. Initial thoughts are that although removal of the validation could result in Unidentified Gas, this would be temporary.

5.9 Accuracy of NDM Algorithm

The accuracy of the NDM algorithm is being assessed by the UIG taskforce. Inaccuracies will result in allocation errors and therefore contribute to UIG. This UIG should be temporary as all meters will eventually be read and the energy reconciled. The accuracy of the NDM algorithm therefore has little impact on the AUG methodology or factors.

However, the algorithm is used to convert measured consumptions to Seasonal Normal (SN) values and to apportion consumption between gas years. Although this will not be optimal and Unidentified Gas may be assigned to the wrong gas year, the average over a number of years should be reasonable.

Although the AUG Expert's remit is to assess permanent Unidentified Gas and to calculate factors to apportion this, during analysis of the effect of pressure and temperature on volume conversion some findings were made which affect the accuracy of the NDM algorithm and the allocation process. These issues have been raised with the UIG taskforce for further investigation, but a brief summary is provided here for information.

Post Nexus, the NDM algorithm is used to estimate and allocate energy to all meter points without daily reads. It is also used as the basis for AQ calculation from a pair of meter reads. Each year, the algorithm is updated with new parameters based on the most recent sample data. The sample data is collected as daily volumes via data loggers. These volumes are then converted to energy prior to use in estimating the NDM algorithm parameters. This conversion uses CFs which are based on the assumption of average pressure of 1013.25 mbar and average temperature of 12.2 °C over a year.

As temperature varies day to day and has a distinct seasonal pattern, the energy calculated for the sample dataset will be biased. In winter, the energy calculated will be under-estimated and in summer it will be over-estimated. This will result in the ALP being flatter than it should be. The DAF will be affected in a similar way as the change in energy consumption between two days will be proportionately less than the change in volume between the days as a result of inaccurate volume conversion.

To assess the approximate size of this effect, the AUG Expert has used actual pressure and temperature data for EA LDZ provided by the CDSP to calculate the percentage error in the total CF as a result of using fixed factors (P=1013.25 mbar, T=12.2 °C) compared to the actual prevailing values over a period of five gas years (2012-2016). The percentage error in CF is shown in Figure 15. For this assessment, the temperature at the meter is taken to be 2°C higher than the average daily air temperature up to 18°C. Above 18°C, the daily average temperature is used without adjustment. This allows for the effect of meters located indoors.

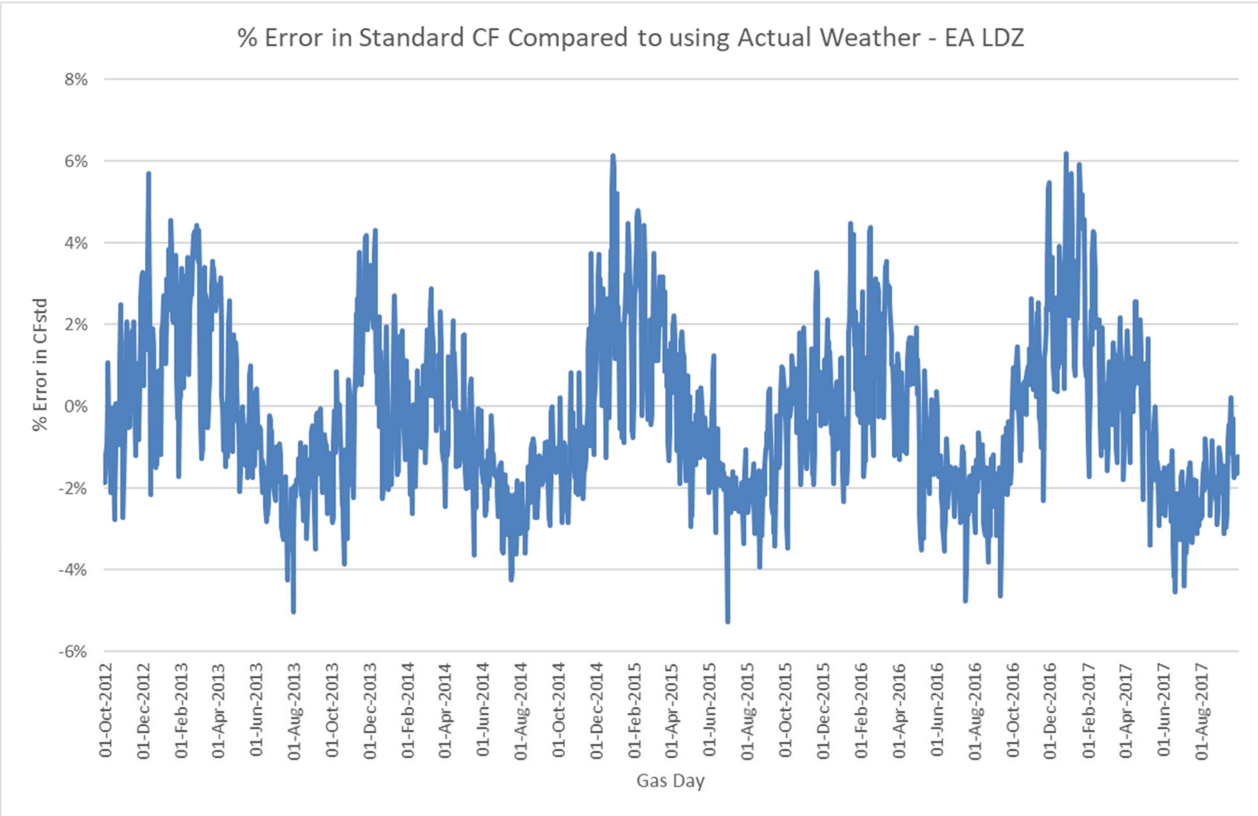

Figure 15: % Error in CF_{std} assuming T=12.2 °C and P=1013.25 mbar for EA LDZ

This clearly shows that any meters using the standard CF to calculate energy consumption will be subject to a significant error. This error is highly variable day to day and shows a seasonal pattern. It is also correlated to daily UIG. The average error in CF over the whole period (EA LDZ only) is 0.07% i.e. the overall contribution to permanent Unidentified Gas is relatively small as daily variations cancel out. However, there is a significant impact on daily UIG and any calculations of energy consumption using data covering periods significantly different to whole gas years. Under the current settlement regime, this level of volatility will remain within daily UIG.

5.10 AQ Accuracy

The issue of AQ accuracy is being assessed by the UIG taskforce. The use of inaccurate AQs will result in allocation errors and therefore contribute to UIG. This UIG should be temporary as all meters will eventually be read and the energy reconciled. From this perspective the AQ will have no impact on the AUG methodology or factors.

However, the AQ values are used in the consumption calculations carried out by the AUG Expert in two ways. Firstly, the AQ is used in the validation process to identify where a calculated consumption is erroneous. Given that this validation is in place to identify the large outliers, it is unlikely that AQ inaccuracies will have any material impact.

Secondly, and more importantly, the consumption calculation relies on AQs for the CSEP consumption estimates. Currently, all data used in the consumption calculations is from pre-Nexus and this could be a significant source of error in the estimate of total consumption. In the post-Nexus world, meter reads will be available for individual meters on CSEPs for the first time. Currently it is not possible to carry out a meaningful analysis due to lack of data. However, it is recommended that an assessment is made of CSEP consumption vs pre-Nexus AQs to determine the level of any bias which may exist once sufficient CSEP meter read data is available.

5.11 Accuracy of NDM Read Estimates

The impact of using estimated reads for NDM sites is being considered by the UIG workgroup in terms of impact on the daily UIG estimates. However, the consumption calculation methodology used by the AUG Expert also uses estimated meter reads. Inaccurate estimates could therefore create an error in the historic consumptions which in turn will affect the Total Unidentified Gas estimate and therefore the Balancing Factor (BF).

To assess the potential impact, the AUG Expert has reviewed the type of meter reads used in the consumption calculation for gas year 2019/20. As part of the calculation process, a record of the meter reads used in the calculation are stored. Table 6 is a high-level summary of the meter reads used in the consumption calculations by gas year. For each gas year it shows the total number of meter reads used, the number of these reads which are estimated and the percentage of reads which are estimated.

Across all gas years, the percentage of estimated reads is very small. Of the estimated reads, 98.6% are for meters in EUC 01B. As would be expected, the number of estimated reads is higher in the more recent years. There is no reason to suggest that the estimated reads will have a significant bias. We therefore conclude that the impact of including estimated reads in the consumption calculation is insignificant.

Gas Year	Number of Reads	Number of Estimated Reads	% of Estimated Reads
2011	40,843,119	8,227	0.02%
2012	40,804,781	12,010	0.03%
2013	40,896,074	13,138	0.03%
2014	40,229,396	38,320	0.10%
2015	36,677,781	34,333	0.09%

Table 6: Summary of Use of Estimated Reads

5.12 UNC/IGT Modifications

This section highlights UNC/IGT Modifications which do or could affect the level and/or split of Unidentified Gas. The AUG Expert monitors these modifications to ensure that the methodology takes these into account.

5.12.1 Mod0644 – Improvements to nomination and reconciliation through the introduction of new EUC bands and improvements for the ALP and DAF.

This modification has been withdrawn.

5.12.2 Mod0654S/IGT110 – Mandating the provision of NDM sample data

Increasing the quantity of training data used to calculate the factors used by the algorithm should improve the model fit and therefore improve its accuracy. In turn, this should reduce the volatility of UIG. The main impact of this modification will be on UIG. However, any changes to the algorithm will necessarily impact the calculation of UIG factors as the algorithm is used to adjust consumptions to Seasonal Normal (SN) conditions and to apportion consumption to gas years. Although this will affect the calculations (via the algorithm parameters), there will be no impact on the calculation methodology.

The AUG Expert also draws attention to the effects of inaccurate conversion from volume to energy (see Section 5.4). The NDM sample data is considered as 'actual' data i.e. it is an accurate measurement of energy. However, the sample data is daily and is used to estimate daily parameters for use by the algorithm. If the sample data is not accurately converted from volume to energy on a daily basis (i.e. not using static fixed factors which represent a yearly average but not daily conditions), this will introduce errors into the calculation of the Annual Load Profiles (ALPs) and Daily Adjustment Factors (DAFs). See Section 5.9 for more details.

The AUG Expert recommends that a review of the use of NDM sample data should be carried out to ensure that the issues relating to energy conversion are fully understood.

This modification will be implemented with effect from 1 March 2019.

5.12.3 Mod0658 – CDSP to identify and develop improvements to LDZ settlement processes

This modification is primarily aimed at UIG. Currently, the AUG methodology calculates factors to apportion based only on permanent Unidentified Gas, as would be calculated at line in the sand. The AUG Expert recognises that apportioning initial UIG based on the UIG factors for permanent Unidentified Gas is not ideal and may be creating cross-subsidies in the industry.

The AUG Expert will monitor the output from the UIG taskforce to ensure that any findings are considered where relevant to the AUG methodology.

5.12.4 Mod0659S – Improvements to the Composite Weather Variable

Improvements to the Composite Weather Variable (CWV) should help with the accuracy of initial allocation using the algorithm and therefore reduce UIG. The AUG Expert suggests that CWV analysis should look at the use of additional weather data locations in addition to the other weather related effects already being considered.

The AUG Expert also draws attention to the effects of inaccurate conversion from volume to energy (see Section 5.4). Any analysis looking at the relationship between weather and consumption should also consider the potential inaccuracies in the consumption (energy), especially as these inaccuracies have also been shown to be correlated to temperature.

Although changes to the definition of CWV will not directly affect the AUG methodology, there will be an effect on the calculations as CWV is used to adjust all consumptions to SN conditions prior to calculating the UIG factors. The AUG Expert will use the best available definition of CWV when calculating UIG factors.

Modification 0659S was withdrawn by the Proposer following approval of the required change via the Demand Estimation Methodology at the Demand Estimation Sub-Committee.

5.12.5 Mod0664 – Transfer of sites with low read submission performance from class 2 and 3 into class 4

If implemented, this modification could result in changes to the product class populations which cannot be predicted in advance by the AUG Expert. The UIG factors depend on the distribution of meters between product classes and the AUG methodology includes an approach to estimate the product class populations for the coming year (See Section 5.5 and Section 7.10). The AUG Expert will therefore monitor the progress of this modification and review the impact should it be implemented.

5.12.6 Mod0672 – Incentivise product class 4 read performance

If implemented, this modification would result in an improvement in the accuracy of allocation primarily due to the use of more up to date AQs. This should help reduce levels of UIG at the time of allocation and will also result in more timely reconciliation. However, this modification will have no impact on permanent Unidentified Gas (as it is assumed that all meters are currently reconciled before line in the sand) and will therefore not impact on the current AUG methodology.

6 DATA USED

This section describes the data requested, received and used to derive the methodology to calculate the UIG factors. The AUG Expert has taken care to ensure that all datasets include all components of NDM consumption, i.e. CSEPs and Scottish Independent Networks are included throughout.

Section 6.1 below gives a summary of the data items requested and their current status. The subsequent sections give more detail about the data items for each individual element of the analysis.

As part of the AUG Expert's quality control process, a number of standard data checks have been defined which are run prior to performing any consumption calculations. Any anomalous data is reported to the CDSP for further investigation. At the current time, not all issues identified have yet been resolved, so by necessity the values contained in this document are based on the best data available. There are also a number of checks during the calculation process to ensure that where data is unreliable it will not be used in the estimation of the UIG factors.

6.1 Summary

Analysis Area	**Required Dataset**	**Status**
Total Unidentified Gas Calculation (Consumption Method)	NDM allocations/throughput	Received
	NDM meter reads and volumes	Received but not usable (imperial and metric volume tagging issues)
	LDZ, DM and Unique Sites Metering Errors	Received but issues with data
	Meter Asset Information	Received but further updates to CFs expected
	Algorithm data (ALPs, DAFs, EWCFs) including CV	Received
	Mod0429 Claims Data	Received
	CSEP AQ data	Received
	Non-CSEP AQ data	Received
	MMSP details	Received
	Prime and Sub-Prime meter details	Received
	New and Lost Sites	Received
Unregistered and Shipperless Sites	Connection details for orphaned sites	Received
	Gas Safety Regulations visit data	Received
	Further investigation results for large/suspicious sites	Supplied on request
	Mod 0410A supporting data	Supplied on request

Analysis Area	Required Dataset	Status
	Shipperless sites supporting data	Supplied on request
	Snapshot files (including MPR details)	Ongoing
IGT CSEPs	Known CSEP data	Partially received
	Snapshot files	Not received
Meter Error	Meter capacity report	Received
Theft	Detected and alleged theft updated to end March 2018 - CDSP	Received
	SPAA Schedule 33 data – SPAA	Partially received
Volume Conversion Error	Information on presence of volume converters and meter locations, historic pressure & temperature data	Received but updates may be provided as there are outstanding queries
Product Classes	Meter point Product Class	Received

Table 7: Data Status Summary

6.2 Total Unidentified Gas Calculation (Consumption Method)

Data has been requested from the CDSP in the following formats. In all cases, data has been provided for gas years 2008/09 to 2016/17 with partial data for 2017/18.

- Allocation data on a day-by-day basis, split by End User Category (EUC). This data includes CSEP allocations.
- Meter read data on an MPRN-by-MPRN basis, with one record for each meter read. Therefore, the volume of data supplied for each MPRN is dependent on the meter read frequency for that meter.
- Aggregate meter error adjustments for LDZs, DMs and Unique Sites.
- Meter asset information on a MPRN-by-MPRN basis. This includes meter installation dates, metric/imperial flag, numbers of meter dials, meter index units and T&P correction factors. This information is used in several different parts of the consumption algorithm.
- NDM Deeming Algorithm factors and CVs for the analysis period.
- Aggregate MPRN count and AQ data by EUC for CSEPs. Meter read data is not available for these sites, but knowledge of the number and AQ of MPRNs allows them to be included in the total Unidentified Gas calculations when the sample consumption is scaled up to cover the full population.
- A history of AQ and EUC data for each MPRN so that calculated consumptions can be validated against AQs and failed meter points can be replaced with an appropriate EUC average.
- Details of all meter points which are or have been part of a Prime and Sub configuration during the analysis period. This includes re-confirmation data to track the potential disaggregation of prime and sub configurations.

- Lists of all new sites and lost sites during the analysis period, including start/end dates. These are used to accurately track the population over time and to ensure that each new or lost site is only included in calculations for the period for which it was active.

The provision of this data allows the consumption for each individual meter point, for each gas year of interest, to be calculated using the method described in Section 7.2. The exact format of the data provided is described in Appendix A.

6.3 Unregistered/Shipperless Sites

The following information is supplied by the CDSP for all Unregistered and Shipperless sites (data supplied on a site by site basis). The CDSP have created a regular report to ensure that new data is collated and sent to the AUG Expert every month. This report covers the following categories of Unregistered and Shipperless sites:

- **Shipper Activity**

  These are new sites created more than 12 months previously, that a Shipper has declared an interest in (such as by creating the MPRN), but are nevertheless not registered to any Shipper. This data is split into sites believed to have a meter and those believed to have no meter.

- **Orphaned**

  These are new sites created more than 12 months previously, that no Shipper is currently declaring an interest in. This data is split into sites believed to have a meter and those believed to have no meter.

- **Shipperless Sites PTS (Passed to Shipper)**

  These are sites where a meter is listed as having been removed and 12 months later the gas transporter visits the site to remove or make the service secure (the GSR visit), but finds a meter connected to the service and capable of flowing gas. If it is the same meter as supposedly removed 12 months previously it is passed to the Shipper concerned to resolve.

- **Shipperless Sites SSrP (Shipper Specific Report)**

  Similar to Shipperless (Passed to Shipper) sites, these are sites where the GSR visit finds a new meter fitted and capable of flowing gas, in which case it is reported to all Shippers.

- **No Activity**

  These are sites currently being processed. They will end up in one of the other categories.

- **Legitimately Unregistered**

  These are sites believed to have no meter and hence are not capable of flowing gas.

- **Unregistered <12 Months**

  These are new sites that have been in existence less than 12 months and are not registered with a Shipper. Action is not taken on such sites until they have been in existence for 12 months. At this point they will move to either the Shipper Activity or the Orphaned category.

For all of these Unregistered/Shipperless Unidentified Gas categories, the following information is supplied for each site:

- Dummy MPRN
- LDZ
- AQ
- Meter Point Status

In addition, the following data is supplied for individual Unidentified Gas categories:

- Meter Attached Y/N
    - Shipper Activity, Orphaned, No Activity, Legitimate
- Meter Point Effective Date
    - Shipper Activity, Orphaned, Unregistered <12 Months, No Activity, Legitimate
- Shipperless Date
    - Shipperless PTS, Shipperless SSrP
- Isolation Date
    - Shipperless PTS, Shipperless SSrP

In addition, the following information is supplied on an annual basis:

- A summary of the remaining Shipperless sites, i.e. those that have been recorded as Isolated for less than 12 months and are awaiting their GSR visit. These sites do not yet appear in the Shipperless PTS or Shipperless SSP lists because sites only qualify for these after the GSR visit has found a meter at the site. This data comes from GSR visit records.
- Connection details for Orphaned sites, including asset and Shipper meter reads and information on whether the confirming Shipper is the same as the Shipper whose Supplier requested asset installation. This data is used to determine the proportion of sites that have been flowing gas prior to becoming registered and the proportion of these that can be back-billed.
- Shipperless sites supporting data. This is used to ascertain the final outcome for each Shipperless site that has appeared in any snapshot but has subsequently been either disconnected or (re)confirmed. This is used to determine whether the Unidentified Gas arising from them is temporary or permanent under the terms of Mods 0424 [7] and 0425 [8].

6.4 IGT CSEP Setup and Registration Delays

Data for IGT CSEP setup and registration delays consists of two elements, as follows:

- **Unknown projects summary, including**
  - the number of unknown projects by LDZ
  - a count of supply points and aggregate AQ of unknown projects by LDZ

This data is supplied by the CDSP in monthly snapshot files on an ongoing basis. The last snapshot received was June 2017, and data from July 2017 to the present is outstanding. An alternative data source is required for post-Nexus information on unknown projects, and this is currently being investigated by the CDSP.

- **Known CSEP Data**

  This file contains data for registered sites on known CSEPs. For previous AUG years, equivalent data for Unregistered sites on known CSEPs was also supplied. The CDSP no longer manages the CSEP Rejection process, however, and so an alternative data source is also being sought for this data. In the meantime, the most recent data from the last AUG year has been used, which is for April 2018. Data is supplied on an annual basis and contains the following data fields:

  - LDZ
  - EUC
  - Number of supply points
  - Aggregate AQ

6.5 Meter Errors

Data for meter error calculations consists of meter capacity, AQ and LDZ for all commercial sites. This report is supplied on an annual basis, with the latest one having been received by the AUG Expert in November 2018. This data is used to identify sites that due to the combination of AQ and meter capacity are likely to be operating at either a high or low extreme of their range, where bias in the readings starts to occur.

6.6 Theft

Data supplied by the CDSP consists of all recorded detected and alleged thefts from 2008 to March 2018. For each theft, the following key data items are supplied:

- Dummy MPRN
- Theft start and end dates
- LDZ
- Meter AQ
- Estimate of energy value of theft (kWh)

In addition, SPAA have supplied the SPAA Schedule 33 data for 2015 to 2017 (summary level) plus line-by-line data from the TRAS Outcome files covering the following:

- Dummy MPRN (defined by the CDSP)
- Meter Serial Number
- EUC (derived by matching MPRN with asset data)
- Product Class (derived by matching MPRN with asset data)
- Meter Type (traditional/Smart/AMR – derived by CDSP)
- Meter installation date (derived by matching MPRN with asset data)
- Source of lead (MAM, MRA, GT, TRAS, own analysis, tip-off)
- Lead investigated? (Yes/No)
- Theft detected? (Yes/No)
- Assessed Losses

The intention is for this dataset to include records for all leads (including ETTOS and TRAS qualified outliers) regardless of whether they were investigated. At this point in time, however, operational issues with the data request and its authorisation have resulted in the data for the 2019/20 analysis being incomplete. Whilst all requested fields are present for the theft records supplied, only records involving actual investigations are available, and hence the following are missing:

- ETTOS leads that were not investigated
- TRAS Qualified Outliers that were not investigated

The loss of this data does not prevent the new theft analysis from being carried out, but it does limit its ability to remove some of the effects of Supplier theft detection targeting from the data.

6.7 Shrinkage

The data required to quantify CSEP Shrinkage is described below:

- GBNA network models (supplied by Cadent).
- Figures for the number of consumers in CSEPs (supplied by the CDSP as part of the IGT CSEPs data)
- National Leakage Test leakage rates (public domain)

6.8 Volume Conversion Errors

The data required to calculate volume conversion error due to atmospheric pressure variations is

- AQ of all meters with Volume Conversion equipment fitted
- Historic actual atmospheric pressure data at MSL for a sample of locations across GB

Currently, the volume conversion error due to temperature is assumed to be zero. In future, if an estimate is required, the AUG Expert would anticipate a need for the following additional information

- Meter location information for all meters
- Historic temperature data

The data required to calculate the volume conversion error due to the use of standard CF rather than site-specific CF is

- AQ of all meters in EUC 04 and above without volume conversion equipment fitted
- CF data for all meters in EUC 04 and above without volume conversion equipment fitted
- Rate of update of CFs from standard to site-specific

6.9 Product Classes

This data set is supplied in conjunction with the asset data used in the Consumption Method and consists of an additional field for each site specifying the Product Class. This data is therefore available for each instance of the asset data that has been supplied.

- An initial list of the Product Class of each meter point at Nexus go-live (1st June 2017)
- Periodic updates listing all product class changes including the date of the change. The latest dataset available is from February 2019.

7 METHODOLOGY

This section describes in detail the methodology for estimating each element of Unidentified Gas.

The first stage in the calculation process is to use the Consumption Method to estimate the total Unidentified Gas for each year in the training period. This process is very similar to that used by the AUGE previously [10] but has been updated to account for the change in definition of the AUG year to align with the gas year following implementation of Mod 0572 [12]. The method has also been updated to allow for the disaggregation of meter points in a prime/sub configuration.

All directly estimated Unidentified Gas categories are then calculated for the same period: this allows the amount of Temporary Unidentified Gas within the Consumption Method total for each year to be ascertained and allows the Balancing Factor (mostly undetected theft) to be calculated. All Unidentified Gas in the Balancing Factor is Permanent.

The data patterns observed in the training period for each Unidentified Gas component, including the Balancing Factor, are used to extrapolate to the forecast year (currently 2019/20) and provide the best estimate of each Permanent element of Unidentified Gas for this year. This is carried out individually for all 36 EUC/Product combinations for every Unidentified Gas category. Finally, these Unidentified Gas estimates are converted into factors by dividing the GWh Unidentified Gas estimates for the forecast year by the aggregate AQ for each EUC/Product combination, as per equation 3.3 in Section 4.3.5.

As given in equation 3.2 (Section 4.3.1), the Consumption Method can be stated in its simplest form as:

Total Unidentified Gas = NDM Allocation - Metered NDM Consumption

This calculation involves correcting the allocations to take account of meter errors (LDZ offtake and DM) and calculating the metered consumption using meter reads, metered volumes or an EUC average consumption for sites where no reliable metered data is available.

The Total Unidentified Gas calculated as above includes both Permanent and Temporary Unidentified Gas. Therefore, Temporary Unidentified Gas (calculated from the individual component parts of Unidentified Gas) must be subtracted from the initial Unidentified Gas total, and it is this amended figure that then goes forward into the remainder of the calculations.

7.1 Correcting the NDM Allocation

The NDM allocation is calculated as

$$Alloc_{NDM} = \text{Aggregate LDZ Load} - \text{DM Load} - \text{Shrinkage}$$

Any subsequently detected significant errors in these three components will constitute Temporary Unidentified Gas which has since been reconciled. Therefore, the allocations are corrected to remove this element.

Meter error adjustment data is received on an LDZ by LDZ basis split by billing month. The total value of the error is given, and this is split so that the correct proportion of each meter error can be assigned to each gas year in which the error occurred.

These errors affect the Aggregate LDZ Load and the DM Load, and have opposite effects on the allocation total, which is calculated at the gas year level of granularity. The result of applying corrections for the meter errors is as follows:

- LDZ meter under-reads *increase* the total NDM allocation
- LDZ meter over-reads *decrease* the total NDM allocation

- DM/Unique site meter under-reads *decrease* the total NDM allocation
- DM/Unique site meter over-reads *increase* the total NDM allocation

7.2 NDM Consumption Calculation

The consumption algorithm relies on a large quantity of data, summarised in Section 6.2. A full description of the raw data used to calculate consumption figures for each individual meter point is described in Appendix A. This raw data is then pre-processed to validate it and to derive additional information to help speed up the consumption calculation process, (details can be found in Appendix B). After the pre-processing the main algorithm is run to calculate consumption on a meter by meter basis. This calculation will not be successful in all cases, so a final step is required to scale up the consumption estimate to account for these 'failed' sites.

7.2.1 Algorithm

Figure 16 shows a flow chart of the process involved to calculate the consumption for a single meter and gas year with references to numbered steps, which are described in detail below.

Note, currently all gas years calculated are from the pre-Nexus period. Therefore, the methodology uses SSP/LSP market sector to determine the best method to use. This will be revisited when post-Nexus data is first used to determine what updates are needed.

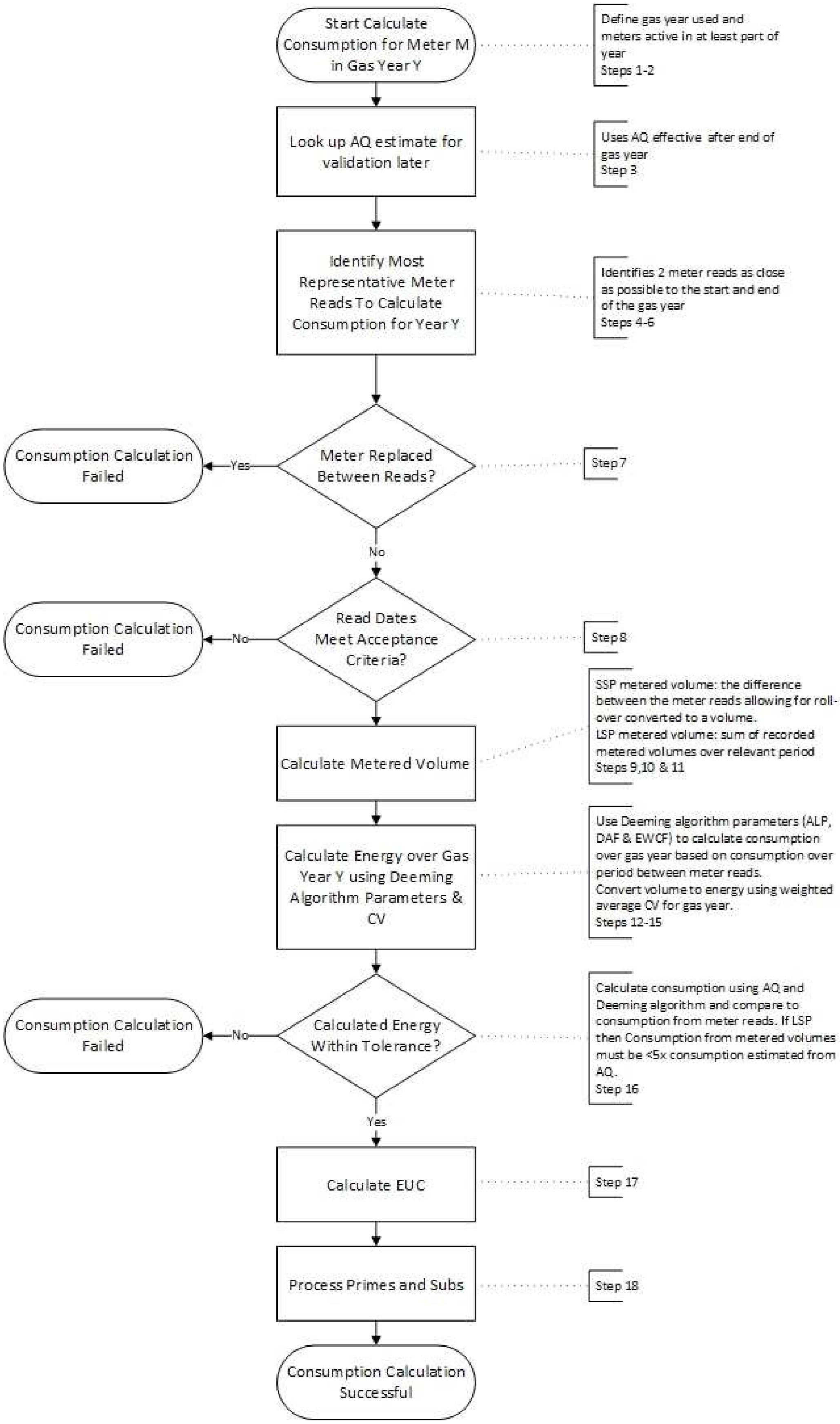

Figure 16: Consumption Algorithm Flow Chart

1. Given a gas year Y, define the start and end dates as 01 Oct Y and 30 Sep Y+1
2. Find all meter points that are not on a CSEP and were active and NDM in a least part of year Y.
3. Look up the first AQ estimate effective after the end of the gas year. If none exists after the end of the gas year use the latest value. From this record store
   i. The AQ value
   ii. The EUC provided by the CDSP
   iii. The pre-calculated consumption band derived by the AUG Expert from the AQ value.
   iv. Market sector (SSP/LSP) based on the EUC from the CDSP
4. For each meter point find the meter reading date and value for:
   - LB1 (Lower Bound 1) – the latest meter reading prior to the start of the gas year
   - LB2 (Lower Bound 2) – the earliest meter reading within the gas year
   - UB1 (Upper Bound 1) – the latest meter reading within the gas year
   - UB2 (Upper Bound 2) – the earliest meter reading after the end of the gas year

   For SSPs those readings which have been flagged as bad by the pre-processing are excluded.

   Where a meter point has changed between NDM and DM or vice versa try to select meter reads from the period when it was NDM.

   Note that for any given meter point, only a subset of this full set of reads may be available. At least one lower bound and one different upper bound meter read are needed. Possible scenarios are shown in Figure 17 below:

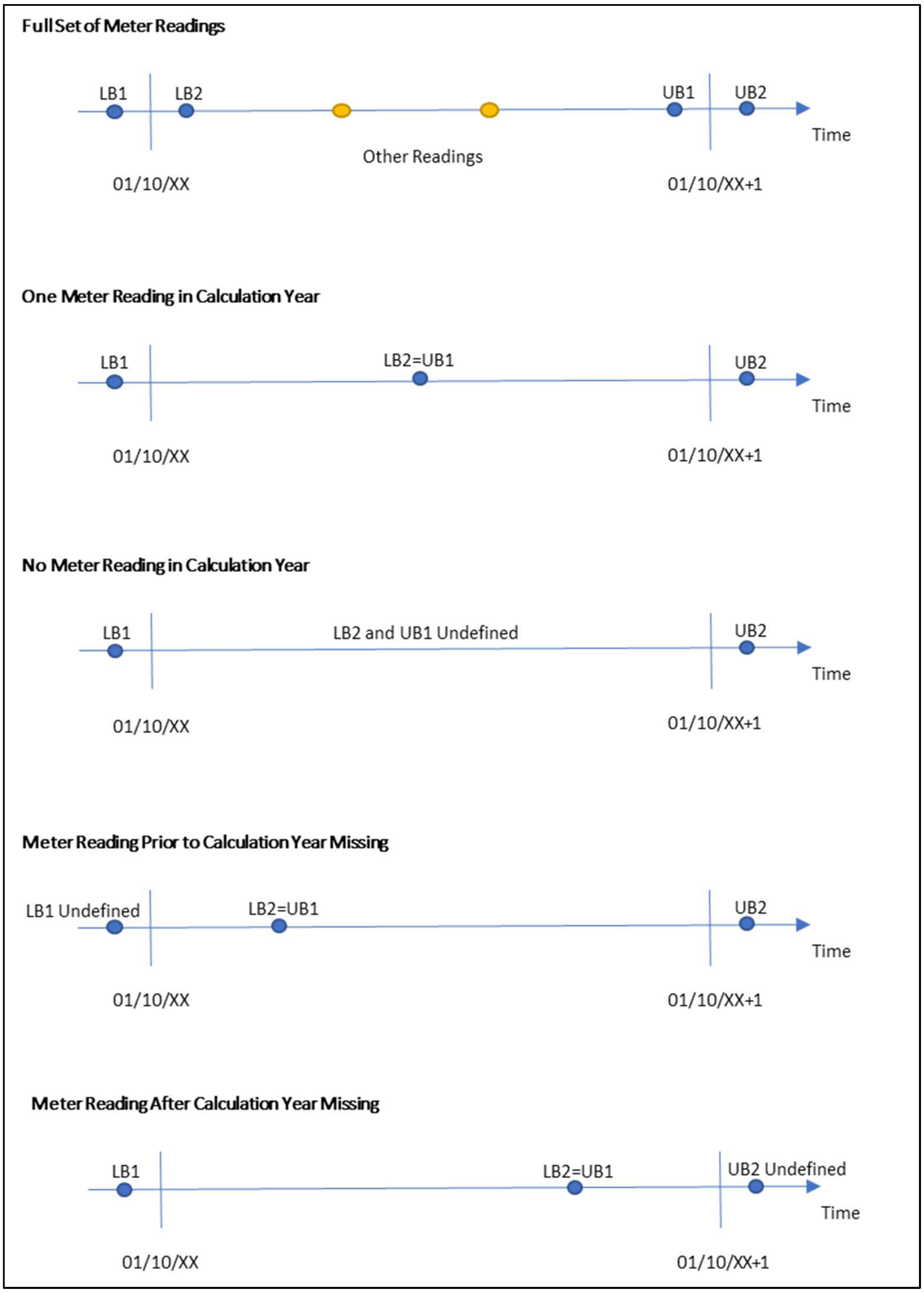

Figure 17: Meter Read Availability Scenarios

5. Set the start meter read date to LB1 unless

    A. the date of LB1 is more than 365 days from the start of the gas year, or

    B. the meter was replaced on or after LB1 and before LB2

   In which case set it equal to LB2.

6. Set the end meter read date to UB2 unless

A. the date of UB2 is more than 365 days from the end of the gas year, or

B. the meter was replaced after UB1 and on or before UB2

In which case set it equal to UB1.

7. If the meter is SSP and was replaced between LB2 and UB1 inclusive, then reject the meter point.

8. Check that:

A. The distance between the two chosen meter readings is at least 170 days

B. The overlap between the metering period and the gas year is at least 90 days

If this is true then proceed to calculating the metered volume, otherwise reject the meter point.

9. Apply either Rule A or Rule B depending on the market sector of the meter point:

A. If the site is SSP then calculate the volume consumed between the two chosen meter readings (mr_1, mr_2). If this gives a negative volume, then check if the meter index has rolled over (see subsection below).

B. Otherwise sum the metered volumes (mv_i) and volume corrections between the two chosen meter readings. If there are any negative volumes in the range, set the sum to -1.

If this step produces a positive volume then proceed to the next step, otherwise reject the meter point.

10. Calculate the fraction of the year that the meter point was active and NDM weighted by the WAALPs. Note, currently the pre-Nexus definition of WCF is used to calculate the WAALPs.

11. Calculate the volume taken over the gas year (or fraction of year calculated in the previous step) by multiplying the volume from Step 9 by

$$\frac{\sum_{\substack{FormulaYear \\ or\ Part\ Thereof}} WAALP^{v}}{\sum_{\substack{Metered \\ Period}} WAALP^{v}}$$

where $WAALP^{v}$ is the WAALP divided by the relevant CV value (i.e. a 'volume' WAALP rather than the usual energy WAALP).

12. Look up, in the meter asset information, whether the meter is/was metric or imperial and then apply either Rule A or Rule B to match the rule chosen in step 9.

A. If the meter point is SSP look up the read units (*U*).

- First choice is the units inferred from the meter read records.
- If this could not be calculated, then use the units provided by the CDSP.
- In the case where the read units from the CDSP are obviously wrong (i.e. are 0 or not a power of 10) use 1 for metric and 100 for imperial meters.

Combine this value with the default correction factor (*CF*) 1.022640 and relevant metric/imperial conversion factor to get a combined conversion factor.

B. Otherwise, if LSP look up the appropriate metric/imperial factor.

If no meter asset information can be found, reject the meter point.

13. Calculate the weighted average CV for the gas year, calculated as

$$\frac{\sum_{\substack{FormulaYear \\ or\ Part\ Thereof}} WAALP}{\sum_{\substack{FormulaYear \\ or\ Part\ Thereof}} WAALP^{V}}$$

14. Convert the gas year volume to energy in kWh by multiplying the output of Steps 11, 12 and 13 together. In summary, depending on the market sector of the meter point, this will be

$$Con = (mr_2 - mr_1) * U * CF * CV / 3.6\ (* 0.0283168466\ if\ imperial)\ \text{for SSP}$$

$$Con = \sum mv_i * CV\ / 3.6\ (* 0.0283168466\ \ if\ imperial\ )\ \text{for LSP}$$

15. Calculate an AQ from this consumption using the appropriate Cumulative Weather Adjusted Annual Load Profile (CWAALP)

$$AQ = Con * 365 / CWAALP$$

16. If a new AQ value has been calculated from the meter readings which is more than four times larger than the old AQ and the new AQ puts the site in the LSP market, then reject the meter point. If the old AQ is 1 then use four times the largest recorded AQ as the check instead.

17. If the consumption calculation was successful, calculate an EUC band based on the new AQ.

18. Carry out post-processing to avoid double counting of subs and deduct consumption. See subsection below for details.

7.2.1.1 Meter Index Rollover Check

Given two reads mr_1 and mr_2 where (mr_2 - mr_1) < 0 the following process is used:

1. Estimate the number of dials from mr_1

$$num\_dials = max(ceil(log_{10}(mr_1)), 4)$$

2. Determine the maximum possible meter read

$$max\_read = 10^{num\_dials}$$

3. Calculate the period between the two meter reads in years

$$num\_years = \sum_{mr_1(date)+1}^{mr_2(date)} ALP\ /\ 365$$

4. Assume meter index roll-over and re-calculate the volume

$$tmp_1 = max\_read - mr_1 + mr_2$$

5. Calculate the new volume as a fraction of the max read per year

$$tmp_2 = (tmp_1\ /\ max\_read)\ /\ num\_years$$

6. If $tmp_2 < 0.25$ then assume the meter index has rolled over and use tmp_1. Otherwise leave the calculated volume as negative and reject the meter point.

7.2.1.2 Prime and Sub Meter Post Processing

As the prime meter consumption is the difference between the total consumption (based on the prime meter reads) minus the sum of the sub-meter consumptions, issues can arise in cases where a full valid set of consumptions for all meters within a sub-prime configuration are unavailable. Note that the Consumption Methodology will not calculate consumption for a DM meter. There are four cases to consider:

1. If the prime meter is DM, no action is necessary as the methodology won't have calculated consumption for the prime meter (consumption not required for DM meters). Sub-meters will be calculated correctly based on available data.
2. If the prime meter is NDM and contains one or more DM sub-meters, then the prime meter consumption calculation is flagged as having failed so that an EUC average consumption is used (see 7.2.2).
3. If the consumption calculation fails for any of the sub-meters, then the prime meter calculation is flagged as having failed. An EUC average consumption is therefore used for the prime meter.
4. If the consumption calculation succeeds for the prime meter and all of its sub-meters, then the prime meter consumption is calculated by subtracting the sub-meter consumptions from the total prime meter consumption.

Prime and sub meter arrangements may be disaggregated so data was requested from the CDSP to track the reconfirmation of these meters. Using this information, the necessary correction is made only for the relevant period.

7.2.1.3 Interpolation of Failed Consumptions

In the case where the consumption calculation fails for a meter point in a specific year but is successful for the two years either side, the average of the two successfully calculated consumption values is used.

Two validation checks are carried out on this process:

- In line with the standard consumption calculation the interpolated value must be less than four times the prevailing AQ value. This is unlikely to be an issue given the consumptions either side must have passed this check.
- The consumption values either side must differ by less than 40%. Otherwise the meter point's behaviour is changing significantly, and a simplistic linear model is unlikely to be appropriate.

7.2.2 Aggregation and Scaling-Up

When applied to each meter point in any given LDZ, the algorithm outputs a set of consumptions that can be aggregated to EUC level. The aggregated data for each EUC is also naturally split into the following categories by the algorithm:

- Meters for which a consumption could be calculated

- Meters for which the algorithm failed (failed to calculate consumption or calculated consumption failed validation)
- Meters in CSEPs (for which Pre Nexus meter reads are not available)

The sum of these three categories across all EUCs gives the total NDM population of the LDZ.

Where a consumption value was successfully calculated the EUC is based on this consumption, otherwise it is calculated by the AUG Expert based on the AQ.

Therefore, for each EUC band the following can be calculated:

1. The number of meter points with a successfully calculated consumption.
2. The number of meter points without a calculated consumption (i.e. calculation failed and interpolation not possible).
3. The average consumption for those meter points with a calculated consumption greater than zero.

The values for 3) are then used to estimate the consumption for meter points in 2). This involves a number of subtleties:

- In 3) attention is restricted to consuming meters only, in order to account for potential differences in the proportion of non-consuming meters within and outside the sample.
- Meter points where the consumption calculation fails are classified as consuming/non-consuming based on AQ, as this is the only reliable data available for such meters. It is recognised that due to changing circumstances for each meter, those with an AQ of 1 for Year X are not necessarily non-consuming during Year X. Likewise, those with an AQ greater than 1 for Year X are not necessarily consuming in Year X. Therefore, two figures have been calculated using available information (i.e. meters within the sample):
    - the proportion of meters with AQ = 1 for Year X that are consuming in Year X = A
    - the proportion of meters with AQ > 1 for Year X that are consuming in Year X = B
- The consumption for the non-calculated meter points is then calculated as

  Consumption = A x (meters with AQ = 1) x "AQ=1" average consumption

  + B x (meters with AQ > 1) x EUC average consumption

  Where:

  - *"AQ=1" average consumption* is the average consumption of meter points where the AQ=1, but our consumption estimate is greater than zero. This can arise when an AQ review produces AQ=1 yet the period of consumption being validated is actually non-zero.
  - *EUC average consumption* is the average consumption for successfully calculated meters in the corresponding EUC Band. The 01B EUC average excludes meters where AQ=1.
- CSEPs are treated differently to failed meters. This is because meter points are assigned to EUC band based on their maximum potential AQ which may not be the same as their current AQ. It is not appropriate to estimate their consumption using the number of meter points in each EUC band multiplied by the EUC band average consumption. The approach used is described in detail in Section 7.2.2.1.

- Where the sample size for a particular EUC for a given LDZ and gas year is less than 30 the national average is used in place of the LDZ average.
- Failed meter points which were only active for part of the year are assigned an average demand scaled based on the sum of WAALPs for that part of the year.

Figure 18 below summarises the process for obtaining a consumption value for each type of meter point.

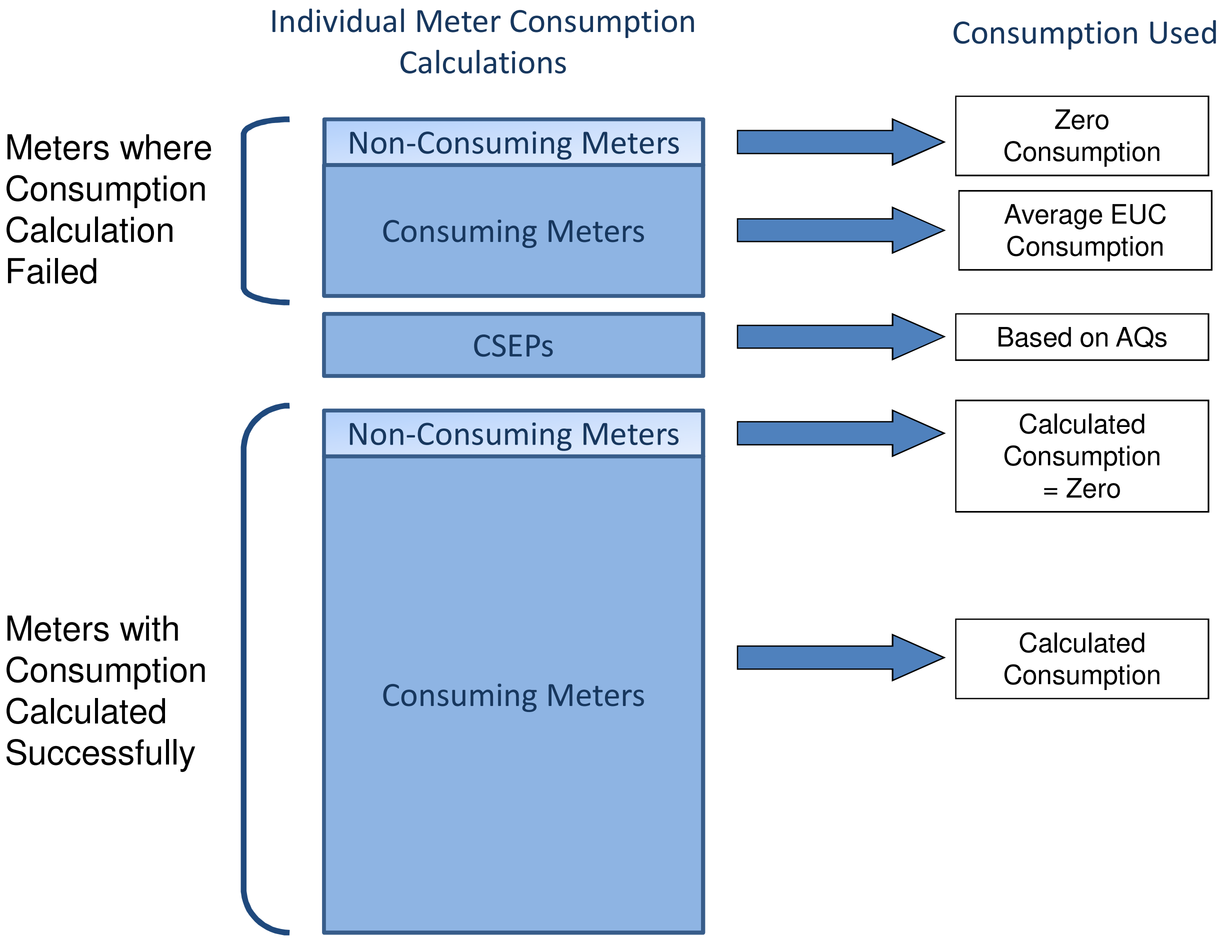

Figure 18: Consumption Method used for each type of Meter Point

Unidentified Gas for the LDZ for the gas year in question is then calculated by summing the metered NDM consumptions across all EUCs and subtracting these from the total combined allocations for the same period.

It is important to note that at this stage these figures still include Temporary Unidentified Gas. Therefore, whilst giving an indication of the order of magnitude of the Permanent Unidentified Gas total for that historic year, this is simply a step in the calculation process and not an estimate of the final value. The method for calculating the remaining Temporary elements is defined in detail in the relevant subsections below.

7.2.2.1 CSEP Consumption Calculation

The following steps are used in the calculation of CSEP consumption for gas year Y. The steps are carried out separately for each LDZ and EUC for each gas year. All references to AQ mean the CSEP AQ for the corresponding LDZ and EUC unless stated otherwise.

CSEP AQ and supply point count are provided by the CDSP from invoicing files. The data therefore aligns to the start of each gas year (1 Oct). However, the AUG Expert has identified significant anomalies in the data for gas years 2014 and 2015. To overcome this issue, the AUG Expert has used the CSEP data from 2013 for both the 2013 and 2014 gas years. The 2016 CSEP data has been used for both the 2015 and 2016 gas years.

1. If the number of meter points in the CSEP is less than 100 in year Y then the CSEP consumption estimate for year Y is the aggregate AQ for year Y

$$C_Y = AQ_Y$$

   Otherwise,

2. Calculate the average AQ for new meters

   For EUC01B before gas year 2012,

$$NMAQ = AQ_{2007} / N_{2007}$$

   For EUC01B for gas years 2012 and 2013,

$$NMAQ = AQ_{2012} / N_{2012}$$

   For all other EUCs and EUC01B from gas year 2014 onwards,

$$NMAQ = AQ_{Y+1} / N_{Y+1}$$

   Prior to 2012, the CSEP NExA table had not been updated since 2007. A new table was calculated for 2012 and this is now updated annually following IGT 053 which was implemented from 1 October 2014 [13]. This removes the additional bias which previously existed for 01B meters as a result of using default AQ values.

3. Calculate the average AQ for lost meters

$$LMAQ = AQ_Y / N_Y$$

4. Estimate the aggregate AQ for year Y by adjusting Y+1 for meter changes

$$AQ'_Y = AQ_{Y+1} - Max(0, (N_{Y+1} - N_Y) * NMAQ) - Min(0, (N_{Y+1} - N_Y) * LMAQ)$$

5. Adjust the aggregate AQ to allow for the recalculation success rate

$$AQ''_Y = AQ_Y + (AQ'_Y - AQ_Y) / RR$$

   where

   AQ_Y is aggregate AQ in year Y. Note that for 2009, AQ_Y is adjusted to the new SNCWV definition.

   AQ'_Y is the aggregate AQ in year Y+1 adjusted to allow for the different number of meters to year Y (as in step 4 above)

   RR is the recalculation success rate expressed as a fraction

6. Estimate consumption for year Y by adding effect of new/lost meters

$$C_Y = AQ''_Y + Max(0, (N_{Y+1} - N_Y) * MAQ_{Y+1}) * YFrac + Min(0, (N_{Y+1} - N_Y) * MAQ_Y) * (1\text{-}YFrac)$$

where

C_Y is the final estimate of aggregate CSEP consumption for year Y

AQ''_Y is the estimate of aggregate AQ for year Y calculated from year Y+1 and adjusted for the recalculation success rate

N_Y is the number of meters in year Y

MAQ_Y is the average AQ per meter in year Y and is calculated as AQ_Y / N_Y

YFrac is an estimate of the proportion of a year's consumption which new meters contribute. It is assumed that new and lost meters will be consuming on average for half of the year so a default factor of 0.5 is used.

If the consumption estimate is negative, then use the AQ as the best estimate of consumption i.e. $C_Y = AQ_Y$.

7.3 Unregistered and Shipperless Sites

The magnitude of every Unregistered and Shipperless category of Unidentified Gas is affected by a number of Mods introduced between 2013 and 2014 to address and reduce these specific areas of Unidentified Gas. The Mods in question are as follows:

- Mod 0410A [9] applies to Shipper Activity, Orphaned and Unregistered <12 Months, and any site created on or after 01/09/2013 is subject to the terms of this Mod.
- Mod 0424 [7] applies to Shipperless PTS, and any site with an isolation date on or after 01/04/2013 is subject to the terms of this Mod. This results in all Unidentified Gas from post-Mod sites being Temporary rather than Permanent.
- Mod 0425 [8] applies to Shipperless PTS, and any site with an isolation date on or after 01/04/2014 is subject to the terms of this Mod. In theory, this should have resulted in all Unidentified Gas from post-Mod sites being temporary rather than Permanent, but the process required to implement this was not available in the pre-Nexus system. Therefore, all Unidentified Gas from post-Mod sites is only Temporary after the Nexus go-live date.

Note that Mods 0424 and 0425 apply to the PTS and SSrP elements of the "Awaiting GSR Visit" category in addition to their main categories.

A full description of the calculation method for all Unregistered and Shipperless Unidentified Gas categories is given below. Raw data for all of these except "Awaiting GSR Visit" is contained in snapshot files supplied by the CDSP every month. These are described in Section 6.3 above.

The following files also contain data that is used in the calculation process and are supplied on an annual basis.

- **Connection Details for Orphaned Sites**
  This dataset includes asset and Shipper meter reads and information on whether the confirming Shipper is the same as the asset Shipper. This data is used to determine the proportion of sites that have been flowing gas prior to becoming registered and the proportion of these that can be back-billed. Backbilling can only occur if the confirming Shipper is the same as the Shipper that carried out site works. The Connection Details dataset is split into two categories (pre-Mod 0410A sites and post-Mod 0410A sites) and different flow proportions and backbilling proportions are calculated for each. This is necessary because the terms of Mod 0410A affect how Unregistered sites are processed, and this leads to different conditions for pre-Mod and post-Mod sites.
- **Gas Safety Regulations Visit Details**
  The gas safety visit data is used to estimate the number and AQ of sites that have been recorded as isolated for less than 12 months and hence have not yet had their GSR visit and do not yet appear in the snapshots as Shipperless PTS or Shipperless SSrP, but are nevertheless still consuming Shipperless gas.
- **Shipperless Sites Supporting Data**
  This dataset contains the confirmation date of each Shipperless site that has appeared in any snapshot but has subsequently been (re)confirmed. It is used to ascertain the final outcome for each of the sites, i.e. whether it was (re)confirmed or whether it was disconnected. This is used to determine the proportion of Shipperless sites that have a meter and are *capable* of flowing gas that actually *are* flowing gas.

Further details of these data files are also given in Section 6.3 above.

The step by step calculation process for Shipperless and Unregistered Unidentified Gas is as follows:

1. Each MPRN in the snapshot files is assessed and flagged for further investigation by the CDSP if any of the conditions specified below are satisfied.
   - If a graph of AQs sorted by descending magnitude contains a “shoulder” point (i.e. a distinct change in gradient), any points to the left of the shoulder are flagged.
   - Any site with an AQ more than 100 times the average AQ for EUCs 02B-09B is flagged.
   - Any site with an AQ greater than 58.6 GWh is flagged.

   The resultant list of flagged sites is sent to the CDSP.
2. The CDSP respond with details where any of the flagged sites have been confirmed on their system, and the confirmed AQ of each such site is provided. Any differences between the queried AQs and the confirmed AQs are applied to the snapshot files. Sites where the CDSP have no further information are left as is.
3. All sites with a listed AQ above the VLDMC threshold (1.465 TWh) have their AQs replaced with the average EUC 02B-09B AQ. VLDMCs cannot be Unregistered or Shipperless due to the greater scrutiny the network code requires on such sites, and hence any AQ above this threshold in the Unregistered or Shipperless lists must be erroneous (e.g. MPRN or phone number accidentally entered in AQ field).
4. Before the analysis is run, the following coefficients are also calculated using the latest available data:

- Fraction of opening meter reads with gas flow for Unregistered sites (for the Permanent/Temporary split for Unregistered Unidentified Gas categories, with different fractions for pre-01/09/2013 and post-01/09/2013 sites).
- Fraction of Unregistered Unidentified Gas not backbilled (for the Permanent/Temporary split for Unregistered Unidentified Gas categories, with different fractions for pre-01/09/2013 and post-01/09/2013 sites).
- Proportion of Shipperless sites being disconnected rather than re-registered.

5. "Fraction of opening meter reads with gas flow" is calculated using the "Connection Details for Orphaned Sites" spreadsheet. This file contains a list of Orphaned meters and includes both their asset meter reading and their opening Shipper meter reading. The number of meters with gas flow (i.e. those where the reading has changed) is expressed as a proportion of the total number of meters in the sample. The dataset is split into two sections - pre-Mod 0410A sites and post-Mod 0410A sites – and separate factors are calculated for each in order to account for changes introduced in the Mod. The calculated proportions are applied to the AQs of each pre-Mod 0410A and post-Mod 0410A site with a meter in the snapshots, to give an estimate of the consumption from sites that are actually flowing gas in the Unregistered Unidentified Gas calculations.
6. "Fraction of Unregistered Unidentified Gas not backbilled" is also calculated using the "Connection Details for Orphaned Sites" spreadsheet. In addition to the meter readings, this file contains a flag that indicates whether the asset Shipper is the same as the confirming Shipper. This flag is used to calculate the proportion of sites with gas flow (as calculated in Step 5 above) that also have a different Shipper. This is the proportion of Unregistered sites that cannot be backbilled and hence contribute permanent Unidentified Gas. As for "Fraction of opening meter reads with gas flow", separate factors are calculated for pre-Mod 0410A and post-Mod 0410A sites.
7. The proportion of Shipperless sites that are disconnected rather than reconfirmed is calculated using information from the "Shipperless Sites Supporting Data" spreadsheet. Any site that disappears from the Shipperless lists without appearing in the "Confirmed" list in the supporting data has been disconnected and this is used to calculate the proportion that are disconnected rather than being reconfirmed. This figure is used as the best estimate of the proportion of sites capable of flowing gas that actually *are* flowing gas (i.e. it is assumed that if a site is flowing gas it is reconfirmed, and if it is not it is disconnected).
8. The raw Shipperless/Unregistered Unidentified Gas calculations are now carried out. This is carried out in a series of spreadsheets, with a set of spreadsheets for each individual Unidentified Gas category. Each of these sheets contains a full history of all available snapshot data for the relevant Unidentified Gas category, which at present runs from September 2011 to September 2018. The availability of snapshot data over such a period of time means that trends can be identified within each Unidentified Gas category and extrapolated to cover training years and the forecast year as necessary.
9. Every Shipperless and Unregistered category is affected by at least one Mod, and the implementation of each of these Mods therefore affects the magnitude and trend of the Unidentified Gas category it refers to. In addition, once the Mod has been implemented this creates a division of each category into pre-Mod and post-Mod sites: only post-Mod sites are affected by the Mod, and hence pre- and post- sites will now behave in different ways and exhibit different trends over time. Hence, they must be analysed separately. In addition, the rules for Unidentified Gas being temporary or permanent can be different before and after the Mod and this must also be taken account of. This

creates a set of four sub-categories of Unidentified Gas for each main Unidentified Gas category as follows:

- Pre-Mod permanent
- Pre-Mod temporary
- Post-Mod permanent
- Post-Mod temporary

The trends for each of these must be assessed individually. For each one, the standard 36-way split also exists (9 EUCs and 4 Product Classes), and it can also be seen from preliminary analysis of the data that there is an LDZ-by-LDZ effect where different LDZs can show different patterns in Unidentified Gas consumption. This creates a total of 1872 individual sub-categories for every main Unidentified Gas category (pre- or post-Mod/permanent or temporary/LDZ/EUC/Product).

The trend for each of these must be individually assessed in order to allow accurate calculation of Unidentified Gas for any time period covered by the snapshots, and accurate extrapolation to any required years that fall outside this range (e.g. the forecast year).

Finally, the introduction of each Mod and the reaction of the industry to it necessarily affects these trends, which will therefore not remain constant over the entire time period covered by the snapshots. Therefore, for each main Unidentified Gas category, the snapshot period is split into a small number of separate sub-periods where the trend is consistent: these differ for each Unidentified Gas category due to the timing and nature of each Mod. Individual trends are therefore calculated for each sub-period and apply to specific time periods only.

10. The Unidentified Gas estimate for each year of the training period and for the forecast period is calculated as follows:

- Where the training year lies within the snapshot period, the fitted trend for that year (calculated using the appropriate sub-period) is used.

- For the forecast year and where the training year lies outside the snapshot period, the appropriate fitted trend is extrapolated to cover the year in question. These extrapolated values are used for the calculation.

- Each point in the trend is the expected AQ of Unidentified Gas from that source at that particular point in time. Therefore, the best estimate of the Unidentified Gas consumed in Year Y is the average of the monthly AQ figures across the whole of the year.

Figure 19 below shows an example of the piecewise fit approach for one Unidentified Gas sub-category (permanent Pre-Mod Orphaned sites, Product 4, EUC 01B, EA LDZ).

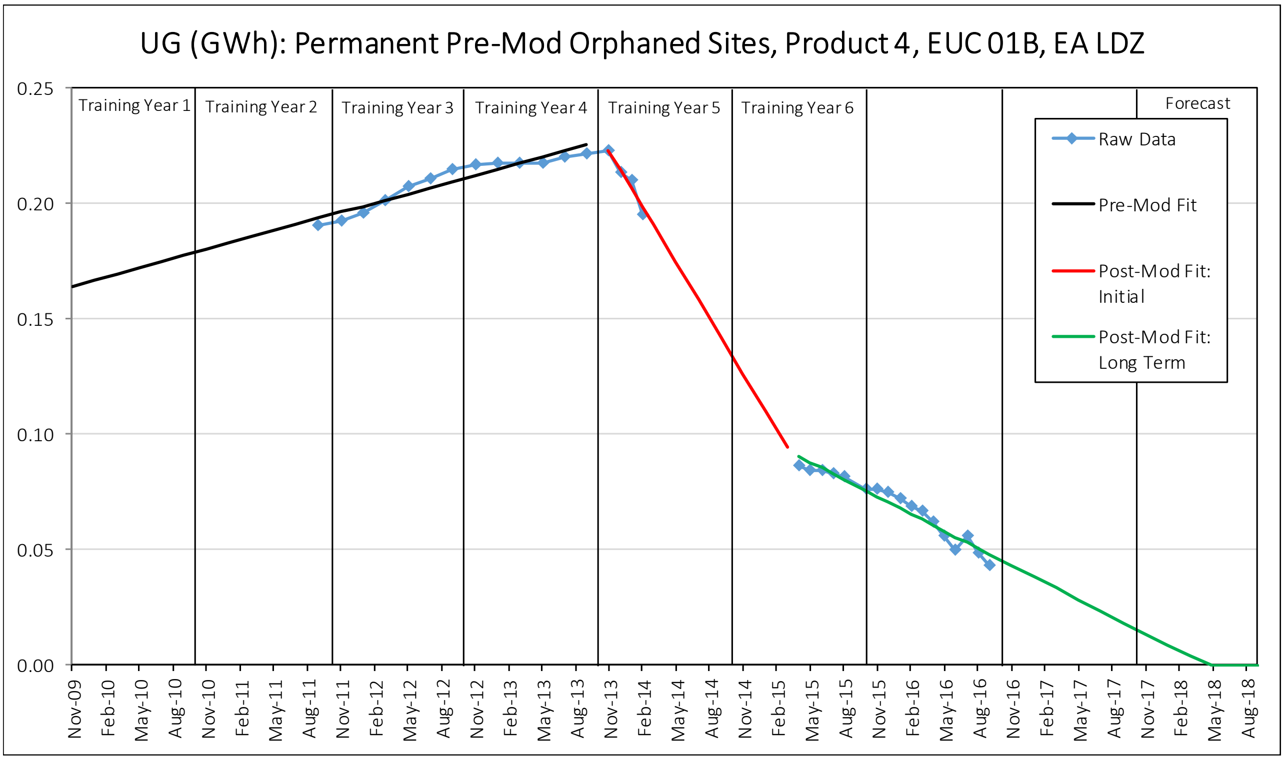

Figure 19: Piecewise Trend for Unregistered Unidentified Gas, Training and Forecast

11. Unregistered sites may or may not have a meter fitted. Where no meter is present, it is assumed that consumption will be zero. For meters in the Shipper Activity and Orphaned categories, the snapshot files contain data split into meter points with and without a meter present. Consumption for these categories is therefore calculated as described above only for meter points where a meter is actually known to be present. For the Unregistered <12 Months category, it is not recorded whether a meter is present or not. For these sites, it is therefore assumed that the fraction of meter points where a meter is present is the same as that found across the other two Unregistered categories.

12. The Unidentified Gas estimate for each type of Unregistered site is adjusted to account for the proportion of such sites with meters that actually flow gas whilst Unregistered, as described in Step 5 above. This adjustment is applied to the annual Unidentified Gas values calculated from the piecewise trends described in Step 10.

13. For Shipperless sites, the snapshots contain all sites found to be *capable* of flowing gas at the GSR visit. From these, the number *actually* flowing gas must be estimated. This is carried out using data in the "Shipperless Sites Supporting Data" file supplied by the CDSP. This contains the confirmation details of all sites that have appeared in the Shipperless report but have subsequently been confirmed and is used to determine the proportion of sites that were reconfirmed rather than being disconnected. The following assumptions are then made:

    - If the site was disconnected it was not flowing gas
    - If the site was reconfirmed it was flowing gas

    The reconfirmation percentage is therefore applied to the Shipperless Unidentified Gas totals (calculated as described in Step 10) to give the best estimate of the amount of Shipperless Unidentified Gas actually consumed.

14. For each Unregistered category, factors are used to convert from Requested AQ to Confirmed AQ and then from Confirmed AQ to AQ Following Review, as follows:

    Requested AQ → Factor 1 → Confirmed AQ → Factor 2 → AQ Following Review

    The "AQ Following Review" figure is regarded as a reliable indicator of the annual consumption at the site.

15. The Unidentified Gas estimates produced for each Shipperless/Unregistered category are multiplied by the appropriate combination of these factors. This is carried out as follows:

    - Shipperless sites (PTS, SSrP): no adjustment required for AQ bias
    - Unregistered (Orphaned, Shipper Activity and Unregistered <12 Months): adjust using composite *Factor1(n) x Factor2*, where *n* represents the Unidentified Gas category in question.

16. The resultant values are now split by Product Class using the principles defined in Section 4.3.3 above. For all Shipperless and Unregistered Unidentified Gas categories the *forecast year* population split by Product Class is also applied to the training period to remove the nuisance factor of changes in the underlying population and allow consistent trends to be created for extrapolation to the forecast year.

17. The above process allows Unidentified Gas estimates for all Unregistered and Shipperless Unidentified Gas categories apart from "Awaiting GSR Visit" to be calculated. At the time of calculation these are split by:

    - Unidentified Gas Category
    - Permanent and temporary
    - Pre-Mod and Post-Mod
    - Year
    - EUC
    - Product Class

    For use in further Unidentified Gas calculations, the estimates must be split by permanent / temporary, Year, EUC and Product Class and so figures are aggregated across categories and pre- or post-Mod status. Tables of figures split this way feed into the final UIG factor calculations.

18. "Awaiting GSR Visit" Unidentified Gas is calculated using Gas Safety Regulations visit data supplied by the CDSP. This file contains the details of each Shipperless site that has crossed the 12-month threshold during a period of a year and has subsequently been visited and found to be capable of flowing gas. The actual sites listed in this file by definition appear in the summarised data in the snapshot files because they have been Shipperless for more than 12 months. If it is assumed that sites become Shipperless at a steady rate, however, it can be assumed that the number and AQ of sites crossing the 12-month threshold in Year Y is a good approximation of the number and AQ that will cross in Year Y+1. At the end of Year Y these sites will have been (recorded as) Isolated for less than 12 months and hence make up the "Awaiting GSR Visit" Unidentified Gas category for this year. Therefore, in order to estimate the Unidentified Gas from this category, the AQs for each site from the GSR visit data are analysed and split by:

    - Permanent and temporary
    - Year
    - EUC
    - Product Class

The transient nature of this Unidentified Gas category means that there is no requirement to define sites as pre-Mod or post-Mod in this case. Whilst all sites in this category will go on to become either Shipperless PTS or Shipperless SSrP sites and hence are affected by Mods 0424 and 0425 (subject to the delay described above), all sites move on to their destination categories in a year and hence no legacy pre-Mod sites remain.

The permanent/temporary split for each year is defined by the implementation date of the relevant Mod, as follows:

- For the PTS element, all Unidentified Gas from this source is permanent up to the 2011/12 gas year, and all temporary from 2014/15 onwards. If a steady influx of sites is assumed for the remaining years, this results in 75% of Unidentified Gas being permanent for 2012/13 and 25% permanent for 2013/14.
- For the SSrP element, all Unidentified Gas from the pre-Nexus period is permanent due to the delay in implementation of back-billing. Following Project Nexus go-live, any site without a record of the installation date of the new meter still contributes permanent Unidentified Gas until its GSR visit regardless of the Isolation date. It is assumed that new meters are installed at a steady rate for these sites and as such the average period of the year for which a meter was present and capable of flowing gas is 6 months (i.e. 50% of the year). The following calculation steps are applied:
    - Multiply the Shipperless SSrP Unidentified Gas element by the "% of Year with Meter" factor. This gives the total (permanent plus Temporary) Unidentified Gas for the Shipperless SSrP element of "Awaiting GSR Visit".
    - Multiply this total by the "% of Meters without Install Date" factor. This gives the Permanent element of the Unidentified Gas total from this source. The remainder of the Unidentified Gas for the Shipperless SSrP element of "Awaiting GSR Visit" is Temporary.

Finally, the reconfirmation percentage (as described in Step 13 above) is applied to convert from the AQ of sites *capable* of flowing gas to the AQ of sites *actually* flowing gas and the data is split by Product Class (as per Step 16 above). This gives the final total Unidentified Gas estimate for this category, split by permanent/temporary, Year, EUC and Product Class.

The Unidentified Gas from this category is now taken through into the final UIG factor calculations.

7.4 IGT CSEPs

Connected System Exit Points (CSEPs) are typically small networks owned by Independent Gas Transporters (IGTs) that connect to the GTs' systems. They are often new housing estates, where the gas network for the estate has been built and is owned by an IGT. CSEPs can potentially contribute to Unidentified Gas where either sites within them or entire IGT networks are not recognised by the CDSP's system and are thus consuming gas in an unrecorded manner.

7.4.1 Overview

Unidentified Gas from CSEPs arises from two sources: Unknown Projects and Unregistered sites on known CSEPs.

Unknown Projects are CSEPs that are known to exist but for various reasons are not on the CDSP's systems. Regular meetings are held between the IGTs and the CDSP in order to resolve these issues and reduce the number of Unknown Projects.

Unregistered sites on known CSEPs lie in CSEPs that are on the CDSP systems, and the CDSP are notified of such Unregistered sites on them.

For both these sources of Unidentified Gas from CSEPs, in the pre-Nexus market the Unidentified Gas from LSP sites was backbilled and therefore temporary, whilst Unidentified Gas from SSP sites on CSEPs was not backbilled and therefore permanent. Under Project Nexus, all sites within CSEPs are individually recorded on the CDSP's systems and can hence be backbilled. Therefore, all CSEP Unidentified Gas from this source is temporary under Nexus.

It is necessary to calculate both the permanent and temporary elements of IGT CSEPs Unidentified Gas for the training period, whilst only permanent is required for the forecast year (although as described above, this is now zero). This allows temporary Unidentified Gas to be removed from the raw Unidentified Gas total (permanent plus temporary) calculated using the Consumption Method for the training period, leaving the final total as permanent Unidentified Gas only. In previous analysis years, the permanent element of Unidentified Gas was then extrapolated to the forecast year to facilitate the calculation of the UIG factors, but this is no longer required.

7.4.2 Data

Unknown Projects data is supplied by the CDSP in monthly snapshot files. These contain data for all Unknown Projects, split by LDZ and by the year in which the CSEP first came to the attention of the CDSP. For each LDZ, the total number of projects, the total number of supply points within them, and the sum of their AQ is given. Note that no split between market sectors is given.

Unregistered sites on known CSEPs data is supplied in a file provided on an annual basis. This file contains data for all known CSEPs, summarised to the LDZ and EUC level. For each EUC, the count of supply points within CSEPs and their aggregate AQ is given. This is provided in separate tables for registered sites and Unregistered sites.

Data for registered sites is used to calculate the average CSEP throughput percentages by EUC for each LDZ, and this is used to split the Unknown Projects data by market sector.

Data for Unregistered sites is used to directly calculate the Unidentified Gas from this source for each LDZ.

It should be noted that the supply point count data for Unregistered sites is actually the number of times the CDSP have been notified that the supply point is Unregistered rather than the number of sites that are actually Unregistered. The CDSP are often notified about the same site on multiple occasions, and this artificially inflates the supply point count figures in this dataset. Therefore, further analysis is carried out on this data in order to estimate the actual number of Unregistered supply points. This is described in more detail below.

7.4.3 Process

Processing is carried out in spreadsheets, which are supplied to the industry to allow auditing of the AUG Expert's calculations to take place. The following process steps are performed:

1. Data for *registered sites on known CSEPs* and *Unregistered sites on known CSEPs* is imported into the calculation spreadsheets. The average AQ per site for each LDZ/EUC combination for registered sites is calculated. As noted above, for the Unregistered data the number of notifications is recorded rather than the number of sites, and hence the actual number of sites must be estimated.
2. The average AQ per site derived from the registered sites is used to estimate the number of Unregistered sites in each EUC using the aggregate AQ for each EUC in the Unregistered dataset. This gives an estimated number of Unregistered sites in each EUC under the assumption that each site has the average AQ for that EUC.
    - If this calculated figure is lower than the number of notifications, it is used as the best estimate of the number of Unregistered sites in that EUC.
    - If the number of notifications is lower, this is used as the best estimate of the number of Unregistered sites in that EUC.
3. The total site count and aggregate AQ by EUC and by LDZ is calculated for registered sites. These figures are then used to calculate the percentage split of CSEP site count by EUC and the percentage split of CSEP AQ by EUC. This split is used in the calculations for Unknown Projects, described below.
4. The total site count and AQ by EUC and by LDZ is calculated for Unregistered sites. The AQ figures produced are split by Product Class using the principles defined in Section 4.3.3 and are used directly in the Unidentified Gas figures: these represent the estimated annual contribution to Unidentified Gas from Unregistered sites in known CSEPs based on current conditions.
5. When each new monthly snapshot file becomes available, data for Unknown Projects is updated. An example of the snapshot table format is shown in Table 8 below. In these tables, the "Year" field refers to the year in which the CSEP came to the attention of the CDSP. For each LDZ the total number of Unknown Projects, their aggregate AQ and the total number of supply points within them is given. Each snapshot represents the situation at the point in time when it was produced.

YEAR	LDZ	Count of Unknown Projects	Sum Of AQ	Count of Supply Points
2016	EA	39	3,483,428	242
2016	EM	34	3,310,074	283
2016	LC	2	656,051	71
2016	NE	9	713,548	38
2016	NO	7	449,652	27
2016	NT	28	2,808,802	136
2016	NW	23	2,579,977	110
2016	SC	71	82,397,218	1,450
2016	SE	40	19,677,626	316
2016	SO	47	11,738,973	651
2016	SW	33	2,548,838	213
2016	WM	23	2,948,638	201
2016	WN	4	292,555	11
2016	WS	12	810,594	94
		372	134,415,974	3843

Table 8: Unknown Projects Snapshot

6. The total number and composition of Unknown Projects by LDZ is calculated by summing across all years. The total Unknown Projects supply point count and AQ for each LDZ is split by EUC using the percentages calculated from known CSEPs, described in Step 3 above.

7. In some cases, there may be additional Unknown Projects where the LDZ is unknown. These are assumed to have average composition by EUC (in terms of supply point count and AQ), with this composition again calculated from registered sites on known CSEPs.

8. The total AQ by EUC across all LDZs plus Unknown LDZ is calculated. These figures are split by Product Class using the standard rules and represent the best estimate of annual consumption in Unknown Projects at the time the snapshot was produced. As for the Shipperless and Unregistered categories, the *forecast year* population split by Product Class is also applied across the training period to allow consistent trends to be created for extrapolation to the forecast year.

9. The total IGT CSEPs Unidentified Gas is calculated for each LDZ as the sum of Unknown Projects Unidentified Gas for the LDZ (from Step 8 above) and the Unregistered Sites on Known CSEPs Unidentified Gas for the LDZ (from Step 4 above). Any Unknown Projects Unidentified Gas from Unknown LDZ is smeared across all LDZs.

10. The above process gives, for each point in time, an estimate of what annual Unidentified Gas from IGT CSEPs would be if conditions for the full year remained as they were in the snapshot. The estimates from successive snapshots show any trend that exists, which then requires extrapolation to the year for which Unidentified Gas is being forecast. An example of this trend across a number of snapshots is shown in Figure 20 below.

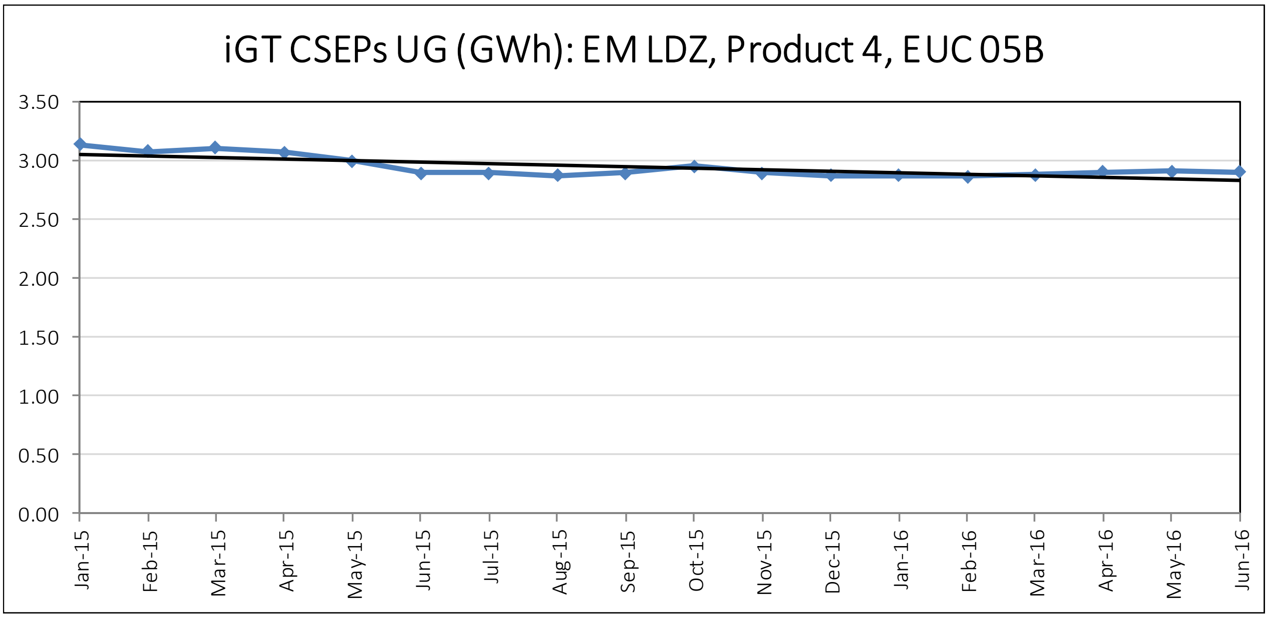

Figure 20: IGT CSEPs Unidentified Gas by Snapshot

11. The above process leads to the creation of a total of 468 individual trend lines for Unidentified Gas from IGT CSEPs. The identified snapshot-to-snapshot trend for each LDZ/Product Class/EUC combination is used to extrapolate either forwards or backwards to each time period of interest. The Unidentified Gas for each year used in the analysis is calculated using the fitted values for each snapshot date that falls within it. The time periods in question cover the Unidentified Gas forecast year and the historic Unidentified Gas training years, with values for each calculated using the fitted trend lines for each LDZ. The forecast year table is used directly in the final UIG factor calculations, whilst those for historic years are used in the calculation of total Unidentified Gas and the Balancing Factor, which are based on data from the training years (currently 2011/12 to 2015/16).
12. For the pre-Nexus period (i.e. the training period), the Unidentified Gas calculated in this way from LSP sites is temporary. The Unidentified Gas calculated in this way from SSP sites is permanent. Under Nexus, the CDSP have confirmed that all Unidentified Gas from this source will be reconcilable once the issue for the CSEP in question has been resolved. Therefore, for the forecast period all IGT CSEPs Unidentified Gas is temporary and does not form part of the permanent Unidentified Gas calculations.

7.5 Consumer Metering Errors

The effects of LDZ metering errors and known DM/Unique Site supply point errors are discussed in Section 7.1 above. In addition, undetected errors in all consumer supply point meters can cause gas to be burnt in an unrecorded or inaccurately recorded manner and hence have the potential to contribute to Unidentified Gas. An assessment of this area of metering error has therefore been carried out by the DNV GL Metering Team, and the conclusions drawn are presented here:

- Very little work has been done in the field of accurately assessing meter drift over time. Information is available about calibration curves taken at a particular point in time for certain meters, but there has never been any dedicated work looking at how these change over time. Therefore, conclusions drawn in this area are largely based on anecdotal evidence and/or extrapolation.

- Smaller sites (e.g. EUCs 01B and 02B) typically have diaphragm meters. The rubber diaphragm is known to warp over time, which causes drift in meter readings. Available evidence suggests that drift is equally likely to be up or down, which would result in a net bias of zero across each population. In the absence of any evidence to the contrary, this is therefore the assumption made throughout the Unidentified Gas calculations.
- In order for a more detailed analysis of such meter drift to be carried out, a large amount of data would have to be collected via a national meter survey. To carry out such a survey would be a significant undertaking as it would require a random sample of a sufficient size to cover many classes of meter (e.g. age of meter, type, model, level of consumption, capacity etc.), as well as co-operation of the customers and the physical testing of the meter itself with properly calibrated equipment. If such a survey was commissioned and carried out, the results could be used in future analyses of meter error. In the meantime, however, the evidence available leads to the assumption of a net zero drift over the population being used.
- Larger sites and offtakes generally have rotary/turbine meters that are constructed of metal and are unlikely to warp over time. These drift less than diaphragm meters, and again are equally likely to drift up or down, resulting in a net bias of zero across the population.
- Where large errors requiring an ad-hoc adjustment are found, these affect the Unidentified Gas calculations directly as described in Section 7.1. Data regarding such adjustments is supplied to the AUG Expert by the CDSP on a regular basis and is used to adjust the initial Unidentified Gas estimates.
- Calibration curves for both diaphragm and rotary/turbine meters follow a similar pattern. Such a curve for a NDM LSP Rotary Positive Displacement (RPD) meter is shown in Figure 21 below.

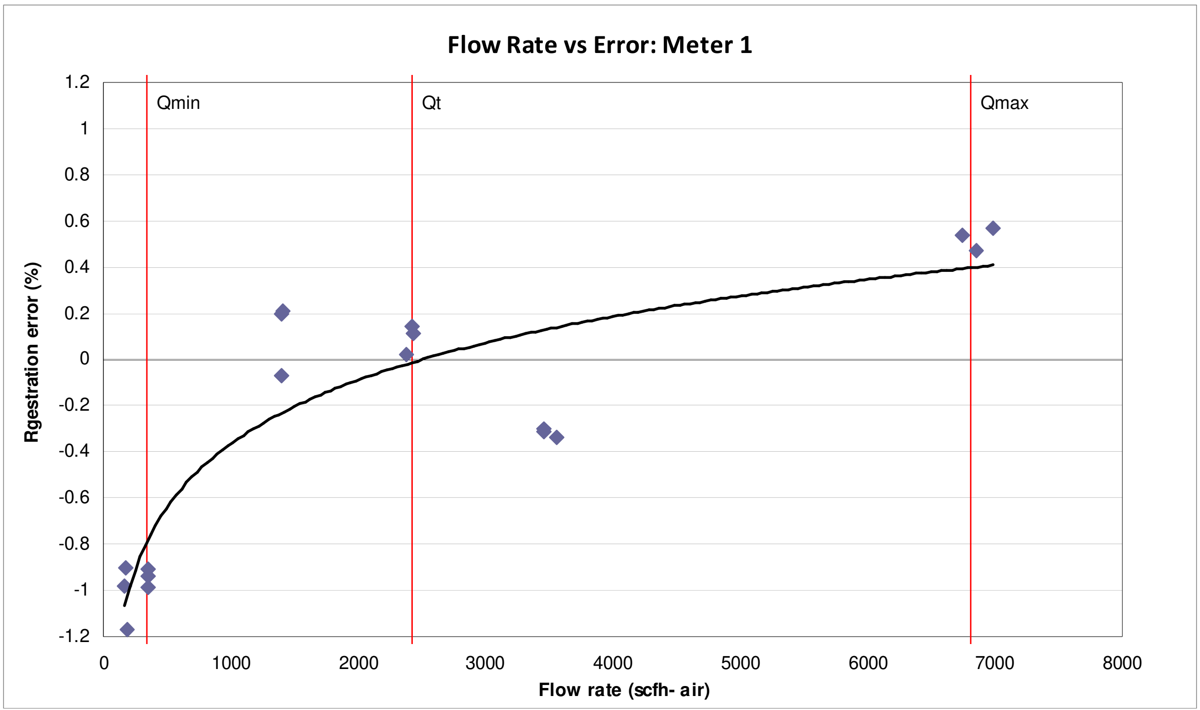

Figure 21: Typical Calibration Curve for an RPD Meter

Data for this graph was provided by the DNV GL metering team and comes from laboratory testing of a typical RPD meter. All identifying information has been removed for confidentiality purposes.

- The prominent features of this calibration curve are a consistent under-read of 1%-1.5% when operating at or below Q_{min}, unbiased readings around Q_t, and a consistent over-read at or close to Q_{max}.
- Meters are designed to operate at or around Q_t, ensuring that unbiased readings are obtained. This is not always the case, however, and circumstances may arise that cause some meters to operate close to Q_{min} or Q_{max}:
  - Loads at a particular site can drop over time, either due to changes in gas usage or because of economic conditions. This can lead meters to operate consistently close to Q_{min}.
  - Where businesses expand their operations without informing their gas supplier, the meter may no longer be appropriate for the load, causing it to run at or above Q_{max}.

Based on the above conclusions, an assessment of likely meter operating zones was carried out. Available data is limited to the meter capacity and AQ of each LSP site (EUCs 02B and above), and this requires the AQ to be used to estimate average hourly load, which can then be compared to meter capacity. This translation from annual load to hourly load necessarily introduces uncertainty into the analysis, but the comparison of average hourly load and meter capacity allows those meters that are likely to be operating at their extremes to be identified.

As stated in Section 4.3.3 above, for this Unidentified Gas category the latest available Product Class data can be used to directly look up the Product for each such site. The EUC of each site can be derived from its AQ and so the full 36-way split by EUC/Product Class can therefore be achieved.

- Sites with an average hourly flow of less than 1% of meter capacity are considered to be likely to be operating at or around Q_{min} when gas is flowing. These are assumed to be operating with an average under-read of 1.5%.
- Sites with an average hourly flow of more than 95% of meter capacity are considered to be likely to be operating at or around Q_{max} when gas is flowing. These are assumed to be operating with an average over-read of 0.5%.

The effects of under-reads and over-reads work in different directions, and the difference between them represents the net over- or under-read in the population.

- A net under-read results in Permanent Unidentified Gas equal to the value of the under-read.
- A net over-read results in the raw estimate of Unidentified Gas being over-stated, and it is therefore adjusted down by the value of the over-read.

The Meter Capacity data file is supplied annually and hence there is insufficient data to allow any trends can be calculated over time. Therefore, in this case the fixed calculated Unidentified Gas energy values are applied to all of the training years and the forecast year.

7.6 Detected Theft

Detected theft that occurs within the back-billing window (3-4 years) is a temporary source of Unidentified Gas and it is therefore quantified and subtracted from the total Unidentified Gas for each part of the training period based on year of occurrence. For AUG years prior to 2017/18, the theft calculations were based entirely on data provided by the CDSP. For more recent years, the SPAA Schedule 33 data was made available via SPAA. The detected theft calculations are therefore now based on a combination of these two data sources.

The SPAA data is reported at an aggregate level only and hence the detailed elements of the analysis still take place using the record-by-record data supplied by the CDSP. These theft values are used directly in the Unidentified Gas calculation for the training period and are split into the relevant year of occurrence to ensure each theft applies to the correct gas year(s) – i.e. the year(s) in which it was active. The output from this detailed analysis is then scaled to equal the total number of thefts reported in the SPAA data, as it is this report that contains the most reliable aggregate figure for the total number of thefts.

The detected theft analysis is carried out in Excel, where for each theft the proportion of it that falls into each training year it is estimated to have been active is first calculated. These figures are used to calculate total detected theft by gas year for the training period as follows:

1. For each theft record and for each training year:
    a. Calculate the number of days in each year that the theft was active. There are four scenarios:
        i. The theft starts and ends within the gas year: the full duration of the theft occurs in a single year and values for all other years are zero for this theft.
        ii. The period of theft spans the gas year in question: duration of the theft in this year is all days in the year.
        iii. The theft is estimated to start in the gas year in question and end in a subsequent gas year: in this case the difference between the end of the gas year and theft start date is used.
        iv. The theft is estimated to end in the gas year in question and had started in a previous gas year: in this case the difference between the start of the gas year and theft end date is used.

        For all years for which the theft is not active, the value is zero.
    b. Calculate the amount of theft estimated to have occurred in the year in question by splitting the total estimated energy value of the theft according to the number of days it was active in each year.
2. Aggregate all thefts by gas year for the training period. Dummy MPRNs that are consistent with all of the other datasets are supplied as part of the CDSP data and so the Product Class for the site associated with each theft can be directly queried. The ultimate use of this data is as an aggregate adjustment to the output of the Consumption Method, however, and so neither a split by EUC nor Product Class is used in subsequent calculations.
3. Scale the theft totals for each year by the ratio of the number of thefts reported in the SPAA Schedule 33 reports to the number of thefts reported in the CDSP data for the equivalent time period:

    $$Adjusted\ Theft_{YEAR\ Y} = Raw\ Theft_{YEAR\ Y} * SPAA/CDSP$$

4. The annual totals calculated in this way are then amended to account for the number of thefts for each training year that have not yet been detected but will be detected in the future up to the point where they can no longer be back-billed (i.e. those that are detected within the settlement period and are therefore temporary). This is up to year 5 prior to 2014 and year 4 afterwards following implementation of Mod 0398 [14] in April 2014. Thefts can often run for years before they are detected, and so in many cases it can be several years before a theft active in Year Y is detected. Analysis of thefts and their time to detection has been carried out using the data supplied by the CDSP. Figure 22 below shows the percentage of thefts detected (i.e. as a percentage of the total thefts that will eventually be detected up to year 8) by year based on all years in the CDSP data.

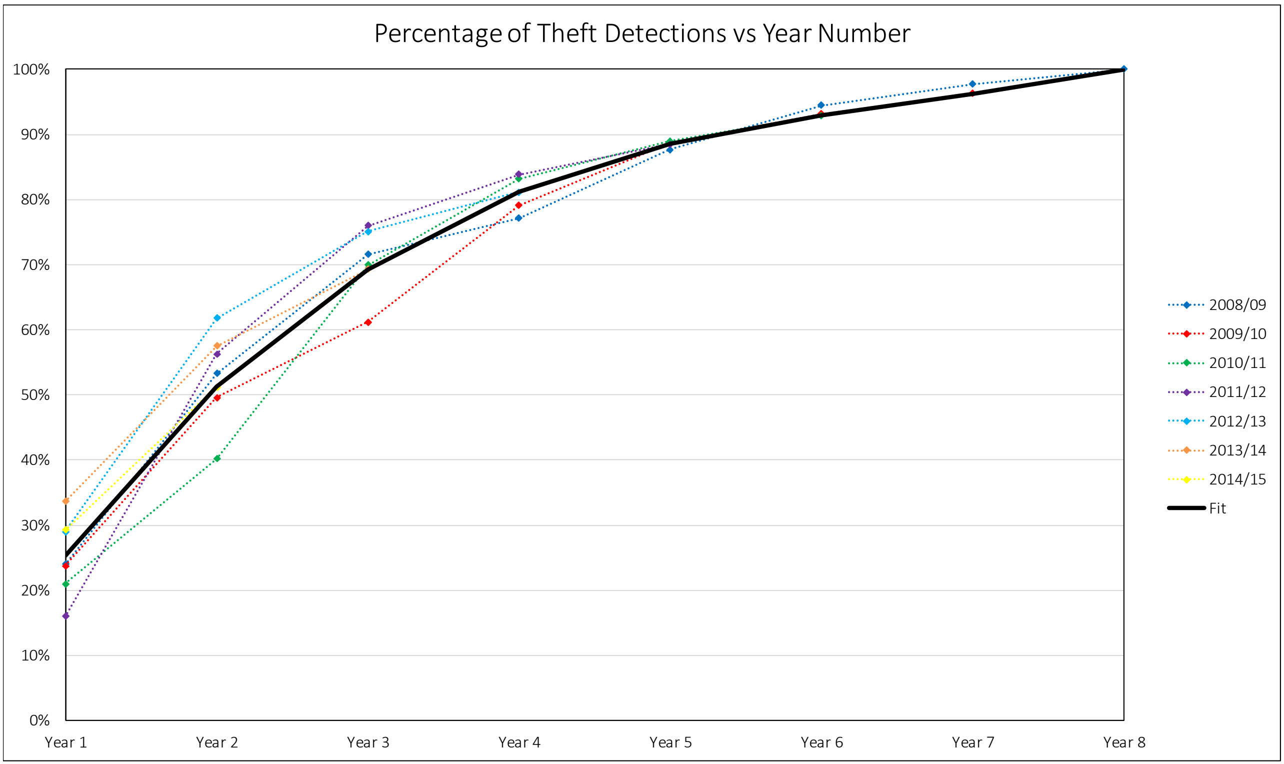

Figure 22: Percentage of Theft Detections by Year Number

The dotted lines show figures for thefts active in individual years, whilst the black line shows the overall fit. This shows that only around 25% of thefts that are going to be detected are detected in the year they become active, and that it takes six years before a detection rate of 90%+ (of the thefts that will eventually be detected) is reached. At Year 8 it can be regarded that the number of ongoing thefts that will be detected but have not been yet is non-zero but negligible.

Given that this full 8 year cycle has not completed for the majority of training years, the detected theft figures (which can now be regarded as representing thefts that have been detected so far) must be amended to incorporate an additional estimate of further thefts that will be detected. This is done using the values from the above graph, where detected thefts for active years less than 8 years ago are scaled up based on the estimated percentage to the level of theft which would be expected in total. Thefts detected up to and including year 4 contribute temporary Unidentified Gas, whilst the remainder are permanent.

5. The estimated annual temporary theft totals for the training period are then applied as corrections to the raw Unidentified Gas estimates from the Consumption Method. The permanent element becomes part of the Balancing Factor.

7.7 CSEP Shrinkage

The AUGE Framework has recently been updated to reflect the fact that Shrinkage Error (i.e. any bias in the Shrinkage and Leakage Model) is considered to be out of scope for the Unidentified Gas analysis. The facility to include an estimate of Shrinkage Error has therefore been removed from the calculation process.

CSEP Shrinkage must still be included in the Unidentified Gas calculations, however. It is acknowledged by the industry that CSEP Shrinkage is not accounted for anywhere in the settlement process and hence ends up in UIG. It is therefore taken account of in the Unidentified Gas calculations and hence the factors that are applied to the daily UIG figure. A method has been developed for estimating CSEP Shrinkage as accurately as possible, and this is documented below. If, in the future, CSEP shrinkage is estimated by the GTs (but still not accounted for elsewhere in settlement) then their figure will be used for Unidentified Gas and will replace the AUGE's own CSEP Shrinkage estimate.

The CSEP Shrinkage estimation method is based on the following principles:

- CSEP Shrinkage arises entirely from leakage.
- Leakage rates from the National Leakage Tests (as used in the Shrinkage and Leakage Model) are available and in the public domain.
- It can be assumed that all mains within CSEPs are PE.
- The calculation of leakage for CSEPs using these leakage rates requires data on the aggregate length of mains in CSEPs. Diameter is not important due to the assumption of all PE, because the leakage rates for all diameters of PE are the same. This information is not available, but it can be estimated as follows:
    - The network models used for validation and planning using the Graphical Based Network Analysis (GBNA) software contain all the information required to calculate both the number of demands and the total mains length in any given network or area of a network as defined by the user. Cadent have given the AUGE permission to use data from these models as part of the Unidentified Gas analysis.
    - GBNA allows the selection of given areas of any network by polygon, with results returned for that polygon only. Therefore, a sample of network areas with a composition similar to that of a typical CSEP can be identified.
    - For each such polygon, the total mains length and the number of customers can be calculated from the network model. This allows an average "mains length per customer" figure to be calculated across the whole sample.
    - The total number of sites in CSEPs is known from data provided to the AUGE by the CDSP. This can therefore be used to calculate an estimate of total CSEP mains length by applying the number of sites to the factor described above.
    - It is assumed that this mains length consists entirely of PE and hence the PE leakage rate from the National Leakage Tests is applied to the estimated CSEP mains length. This is the best estimate of CSEP Shrinkage.
    - The energy value of CSEP Shrinkage is converted to a percentage by dividing it by the total LDZ Shrinkage estimate from the Gas Transporters [19] for the equivalent period. The percentage of the Shrinkage total that arises from CSEPs is assumed to remain constant across the training and forecast period.

This approach involves the application of an LP leakage rate to all CSEP mains. Whilst it is known that not all CSEPs are low pressure, LP leakage rates are the only ones that currently exist, and these are used for both the LP and MP/IP pressure tiers in the Shrinkage and Leakage Model. The approach to CSEP Shrinkage is therefore consistent with the accepted approach to the calculation of overall LDZ Shrinkage.

7.8 Volume Conversion Error

For each year of the training period, the energy affected by the volume conversion issues (EAV) depends on the prevalence of volume conversion devices and is calculated as

$$EAV = Cons_{SN} * (1 - x)$$

Where:

EAV = Energy Affected by Volume conversion issues

$Cons_{SN}$ = Seasonal Normal Consumption calculated by the AUG Expert

x = Proportion of NDM AQ from meters with volume conversion devices (9.26%), as described in Section 5.4.2.

The total Unidentified Gas from assuming that average atmospheric pressure = 1013.25 mbar (UIG_p) is then calculated as described in Section 7.8.1 below and the total Unidentified Gas from assuming that average temperature = 12.2 °C (UIG_t) is calculated as described in Section 7.8.2 below.

The total Unidentified Gas from volume conversion error (UIG_p + UIG_t) is then split by EUC and Product Class based on the total prevailing AQ excluding those meters with volume conversion devices. This gives the split in Table 9.

	PC1	PC2	PC3	PC4
01B	0.00%	0.00%	0.28%	77.78%
02B	0.00%	0.00%	0.77%	5.83%
03B	0.00%	0.00%	0.97%	3.88%
04B	0.00%	0.00%	0.62%	4.59%
05B	0.00%	0.00%	0.27%	2.43%
06B	0.00%	0.00%	0.19%	1.28%
07B	0.00%	0.00%	0.08%	0.61%
08B	0.00%	0.00%	0.04%	0.35%
09B	0.01%	0.00%	0.00%	0.00%

Table 9: Volume Conversion (Pressure & Temperature) Error Split

7.8.1 Average Atmospheric Pressure Assumption

The error in the standard correction factor due to the average pressure assumption is estimated as

$$CF_{err}^{std} = \left[\frac{(P_{av} - 1013.25)}{1013.25}\right] * 1.0098/1.02264$$

Where:

CF^{std}_{err} = Error in the standard Correction Factor expressed as a ratio i.e. $(CF\text{-}CF_{std})/CF_{std}$

P_{av} = Average actual pressure at MSL over all LDZs for the relevant gas year. As pressure data is only available from 2012 onwards, the average value across all gas years (2012-2017) is used for gas years before 2012.

The Unidentified Gas as a result of this pressure correction is then given by

$$UIG_{std} = EAV * CF_{err}^{std}$$

Where:

UIG_{std} = the Unidentified Gas assuming all meters use the standard Correction Factor

EAV = Energy Affected by Volume conversion issues

CF^{std}_{err} = Error in the standard Correction Factor expressed as a ratio

The estimate of UIG_{std} is a first order approximation assuming all meters without volume conversion operate at or near 21 mbar pressure. In reality, there are meters which operate at higher pressures. These meters are less sensitive to the changes in atmospheric pressure so UIG_{std} will be an over-estimate.

A further adjustment is therefore calculated for all meters with CF>1.03193 (roughly equates to meter pressures above 30 mbar). This affects only 1,690 meters. For each of these meters individually, the error in the CF is calculated as follows

$$CF_{err}^{act} = \left[\frac{(P_{av} - 1013.25)}{1013.25}\right] * 1.0098/CF$$

Where:

CF^{act}_{err} = Error in the actual Correction Factor expressed as a ratio

P_{av} = Average actual pressure at MSL over all LDZs for the relevant gas year. As pressure data is only available from 2012 onwards, the average value across all gas years (2012-2017) is used for gas years before 2012

CF is the actual CF for that meter.

The correction for each meter is then calculated as

$$UIG_{hp} = AQ * \left[1 - \frac{(1 + CF_{err}^{act})}{(1 + CF_{err}^{std})}\right]$$

Where:

UIG_{hp} = a correction to allow for meters operating at higher set pressures

AQ = Prevailing AQ of the meter

CF^{std}_{err} = Error in the standard Correction Factor expressed as a ratio

CF^{act}_{err} = Error in the actual Correction Factor expressed as a ratio

The total Unidentified Gas as a result of this pressure correction is then given by

$$UIG_p = UIG_{std} - UIG_{hp}$$

Where:

UIG_p = the total Unidentified Gas as a result of assuming average pressure is 1013.25 mbar

UIG_{std} = the Unidentified Gas as a result of assuming average pressure is 1013.25 mbar calculated based on

the error in standard CF

UIG_{hp} = a correction to allow for meters operating at higher set pressures. This value reduces

the Unidentified Gas total as meters operating at higher set pressures are less sensitive to changes in

atmospheric pressure.

The value of UIG_{hp} is actually negligible (~0.2 GWh on average between 2011 and 2015).

7.8.2 Average Temperature Assumption

The error in the correction factor due to the average temperature assumption is estimated as

$$CF_{err} = \frac{(273.15 + 12.2)}{(273.15 + T_{av})} - 1$$

Where:

CF_{err} = Error in the Correction Factor expressed as a ratio i.e. $(CF\text{-}CF_{std})/CF_{std}$

T_{av} = Average actual temperature in °C over all LDZs for the relevant gas year.

The Unidentified Gas as a result of this temperature correction is then given by

$$UIG_t = EAV * CF_{err}$$

Where:

UIG_t = the Unidentified Gas as a result of assuming average T is 12.2 °C

EAV = Energy Affected by Volume conversion issues

CF_{err} = Error in the Correction Factor expressed as a ratio i.e. $(CF\text{-}CF_{std})/CF_{std}$

Although a methodology to handle a bias in the assumed average temperature has been developed, it is currently assumed that the temperature bias and therefore UIG_t are zero.

7.8.3 Use of Standard Correction Factor

The permanent UIG resulting from using a standard CF where a site-specific CF should be used is calculated as shown in Section 5.4.5.2. This 33.53GWh of permanent UIG needs to be split by EUC and product class. This split is calculated based on relative throughput (AQ) for affected meters. Affected meters are those in EUC 04 and above with standard CF and without volume converters fitted. The resulting split is shown in Table 10.

	PC1	PC2	PC3	PC4
01B	0.00%	0.00%	0.00%	0.00%
02B	0.00%	0.00%	0.00%	0.00%
03B	0.00%	0.00%	0.00%	0.00%
04B	0.00%	0.00%	5.66%	52.95%
05B	0.00%	0.00%	2.86%	17.15%
06B	0.00%	0.00%	1.26%	10.89%
07B	0.00%	0.00%	1.60%	5.17%
08B	0.00%	0.00%	0.00%	2.44%
09B	0.00%	0.00%	0.00%	0.00%

Table 10: UIG from Standard CF Split

7.9 Balancing Factor

All of the analysis described above allows the total permanent Unidentified Gas to be calculated for each year of the training period. The permanent elements of the directly calculated Unidentified Gas categories (IGT CSEPs, Shipperless/Unregistered, consumer meter errors, CSEP Shrinkage and volume conversion error) can be subtracted from this to give the Balancing Factor for each year.

At present there are five years in the training period, which allows the pattern in the Balancing Factor to be assessed and the appropriate method selected to extrapolate it to the forecast year. This is discussed further in Section 7.11 below. In addition to extrapolating the Balancing Factor as accurately as possible, it must also be split by EUC and Product Class like all categories of Unidentified Gas. In order to do this, rules must be applied.

It is assumed that the Balancing Factor is composed mainly of undetected theft, and this provides the basis for how it is split. The following sections describe the logic and calculation steps for this process.

7.9.1 Background

The purpose of this analysis is to use detected theft records to create a set of factors that can be used to split the Balancing Factor between the 36 combinations of EUC/Product Class. This approach is based on the assumption that is made throughout the Unidentified Gas analysis: that the Balancing Factor is composed mostly of undetected theft and should therefore be split based on our best estimate of the relative incidence of such theft.

The issue with using detected theft records to split undetected theft is that detected theft patterns are not necessarily consistent with wider (undetected) theft. Theft will only be detected where it is looked for, and so detected theft rates are heavily influenced by the detection activity that each Supplier chooses to carry out. There is a lot of detected theft information available and it is a potentially useful resource for deriving information about undetected theft, but the challenge of this process is to remove the bias caused by the targeting of the Suppliers' detection activities to produce unbiased factors that will be indicative of overall theft.

The process of detecting theft has three stages, as follows:

- Lead
  Suspicious meter read pattern or tip-off. Meter reads can be identified as suspect via the Supplier's own analysis, TRAS outlier analysis, or a notification from Meter Asset Manager (MAM), Meter Reading Agent (MRA) or a Gas Transporter (GT).
- Investigation
  The Supplier decides which leads, if any, to investigate further. Each Supplier uses their own criteria and investigation rates vary widely – from investigating nothing to investigating everything. On average, around 35% of leads are investigated (2016 SPAA Theft of Gas report).
- Detection
  The proportion of investigations that lead to a detected theft again varies between Suppliers, from 0% to 40% (2016 SPAA Theft of Gas report). On average, approximately 20% of investigations result in a detection of theft.

7.9.2 Principles

In order to produce objective factors that can be applied to undetected theft (and hence the Balancing Factor), the bias needs to be removed from two of the above stages:

1. The bias caused by what detection activity (or lack of it) the Supplier chooses to carry out.
2. The bias caused by the Suppliers' selection of which leads to investigate.

At this point in the development of the methodology, the relationship between investigation and detection (as recorded in the raw data) can be assumed to be constant and representative for any given EUC/Product Class category, although Investigation → Detection rates will vary from category to category. This assumption of single Investigation → Detection rates for each category may not be strictly true, because if the Suppliers' selection criteria are effective there is likely to be a "diminishing returns" effect – i.e. the highest ranked (and therefore first) investigations are more likely to relate to a real theft, whilst the more investigations are carried out the lower the quality of the lead being followed and the lower the probability of it relating to an actual theft. The data currently available does not allow this effect to the quantified, however, and very detailed data from Suppliers covering the ranking of leads and the sequence of investigations would be required in order to do this.

Therefore, the Investigation → Detection relationships are assumed to be constant. This assumption will remain valid whatever "diminishing returns" effect really does exist as long as Supplier detection behaviour remains consistent (i.e. they maintain a similar strategy over time).

As with all elements of the Unidentified Gas analysis, the undetected theft factors must be split by EUC and Product Class as described above, and in this case by Meter Type (traditional/Smart/AMR) as well. The need to split by Meter Type only creates 3 more categories because the remaining 33 all have mandatory requirements for the meter: Smart Meter or AMR for anything except Product Class 4, and AMR for EUC 04B and above. This leaves just Product Class 4 EUCs 01B-03B that can take either a Smart or traditional meter.

The general principle is therefore to base our method on existing leads and investigations (split by EUC/Product Class/Meter Type), but to adjust the numbers of these to what they would have been if the investigations had been carried out equitably based only on:

1. Population
2. Propensity to suspicious meter reads

Once these figures have been derived, category-by-category Investigation → Detection rates can be used to convert to unbiased detection numbers. The relative size of these adjusted detection figures provides the basis for undetected theft factors which can then be used to split the Balancing Factor.

7.9.3 Data

The data specified below represents a subset of the contents of the TRAS Outcome files that are provided by Suppliers. This data has been formally requested by the AUGE, and this request was approved by the SPAA Change Board on 29 November 2018.

The removal of bias from the "leads" data and the production of category-by-category Investigation → Detection rates both require record by record theft data from the TRAS Outcome files. This must include **all** leads (not just those that were investigated or led to a detection) for the full analysis described here to be carried out. As described in previous sections the ETTOS leads that were not investigated and TRAS Qualified Outliers that were not investigated were not included in this year's dataset, however, and therefore the parts of the analysis described below that depend on these particular datasets have been excluded. This means that it has not been possible to remove all of the bias caused by Suppliers' differing theft detection strategies in the calculations for the 2019/20 UIG factors.

The minimum information each record must contain is as follows. Items in italics are provided by the CDSP based on the data supplied by SPAA.

- (Dummy) MPRN
  Real MPRN to be supplied to the CDSP, who will convert to dummy MPRNs consistent with other datasets.
- Meter Serial Number
  Provided to CDSP only, not to the AUGE.
- *EUC*
  To be provided by the CDSP, queried using MPRN.
- *Product Class*
  To be provided by the CDSP, queried using MPRN.

- *Meter Type (traditional/Smart/AMR)*
  To be provided by the CDSP. This will be queried from the asset data using MPRN, and also calculated using the Meter Serial Number based rule set provided by the AUGE. If either source returns AMR or Smart, this is the assigned value.

- *Meter installation date*
  To be provided by the CDSP, queried using MPRN.

- Source of lead (MAM, MRA, GT, TRAS, own analysis, tip-off)
  From TRAS Outcome files.

- Lead investigated? (Yes/No)
  From TRAS Outcome files.

- Theft detected? (Yes/No)
  From TRAS Outcome files

- Assessed Losses
  From TRAS Outcome files.

Data is therefore initially supplied to the CDSP, and then from the CDSP to the AUGE once it has been anonymised and the additional fields added.

7.9.4 Calculation

The first step of the analysis is to produce a set of unbiased leads, split by the 39 population categories (i.e. the original 36-way EUC/Product Class split plus 3 for meter type). Leads from the following sources can be considered to be free from Supplier bias:

- TRAS
- MRA
- Tip-off

Leads from all of these sources either come from the whole population where any given theft is equally likely to be flagged (MRA and tip-off) or are the result of dedicated analysis that is applied to the whole population without Supplier specific targeting (TRAS). Leads from other areas (own analysis, MAM, GT) *may* have inherent targeting that could skew the number of leads coming from these sources and create bias across the population categories. These are therefore discounted in the "unbiased leads" calculation.

The unbiased leads derived in this way (as a subset of the overall leads) and split by EUC/Product Class/Meter Type therefore reflects a combination of both the population of each category and the propensity to suspicious meter reads. At this point the option exists to scale the "unbiased leads" total to the overall leads total, which would result in estimates for each population category of what the number of leads would have been without any targeting. The final output from the overall detected theft analysis is a set of factors that are used to split the Balancing Factor, which therefore operate on the basis of their relative level rather than their absolute level. As such, scaling the leads in this way does not result in any tangible difference in the output, but it nevertheless ensures that the leads total remains the same - this may aid industry parties in understanding the process. This is therefore strictly speaking an optional step, but one which is applied for this reason.

These “unbiased leads” figures must now be converted first into “unbiased investigations” and from there into “unbiased detections”.

Whilst the method for the Investigation → Detection step has already been defined in Section 7.9.2 above, and appropriate rates for each population category can be calculated from the record-by-record theft data, the category-by-category Lead → Investigation step has not yet been defined.

Whilst we now have unbiased figures for leads, as described above, we cannot use category-by-category Lead → Investigation rates calculated from the raw data because the decision to investigate certain leads but not others still lies with the Supplier and hence may still contain an element of targeting. If such Lead → Investigation rates were calculated from the raw data (and aggregated across Suppliers), any targeting effect would manifest itself as a deviation from uniform values: if there was no targeting, all the rates would be the same.

At this point an assumption needs to be made that the quality of leads (which can be regarded as the likelihood of any given lead meriting further investigation) does not vary between population categories: so, for example, leads from Population Category A will have a similar probability of meriting further investigation as leads from Population Category B. This is a reasonable assumption because whilst the data granularity (i.e. meter read frequency) will vary between population categories, this will affect only the speed with which a lead can be identified rather than the quality of the lead itself.

Therefore, based on this assumption, any deviation from uniformity in the Lead → Investigation rates across population categories reflects different Supplier behaviour in following up these leads. If, for example, a large domestic Supplier rarely acts on any leads but a large Supplier for small commercials follows up almost all leads, this will knock on into differences between the Lead → Investigation rates for EUCs 01B-03B (all relevant Product Classes).

With no differences in Supplier behaviour there would be a constant Lead → Investigation rate across all population categories, and therefore this needs to be the basis for the step from unbiased leads to unbiased investigations in the theft analysis. A single rate calculated as the ratio of all leads to all investigations (both aggregated across all categories and all Suppliers) should therefore be used as a constant value to convert unbiased leads to unbiased investigations.

As described in Section 7.9.2 above, category-by-category Investigation → Detection rates derived directly from the raw data can now be applied to the unbiased investigations figures. These rates can be calculated from the raw data without the need for further manipulation because it is the identification of leads and the decision to investigate that are affected by the different theft regimes of different Suppliers – once an investigation is under way, its likelihood of resulting in a detection of theft is unaffected by the decision process that led to the investigation.

The result of this stage of the calculation is a set of estimated unbiased theft detections for the time period covered by the raw theft dataset, split by the 39 population categories (36 EUC/Product Class plus 3 for meter type). In the final steps of the process these must first be converted to kWh of theft rather than the number of thefts, and then projected forward to the forecast year. At this point, the dual figures for Product Class 4 EUCs 01B-03B are combined into single figures for each, and these final figures are converted to factors. This final output is used to split the Balancing Factor. This process is carried out as follows:

1. For each EUC/Product Class/Meter Type category, calculate the average kWh stolen per theft. These figures will reflect not only the higher quantities of gas consumed by larger sites, but also any effects caused by meter read frequency and data granularity affecting theft duration.
2. Multiply the number of thefts by the average kWh to give the unbiased total stolen energy from detected thefts.
3. Calculate the change in population for each population category from the theft dataset year to the forecast year, as a percentage Pn% for each category.
4. Scale each unbiased stolen energy figure by each Pn – this is the best estimate of unbiased total stolen energy from detected thefts for the forecast year.
5. Add the individual component unbiased stolen energy figures for PC4 01B, PC4 02B and PC4 03B to give single estimates for each of these EUC/Product Class categories.
6. Convert these raw figures to proportions for each EUC/Product Class category.
7. Apply these proportions to split the Balancing Factor estimate for the forecast year.

7.9.5 Smart Meter Theft Adjustment

The above sections describe the proposed theft calculation method in full. In addition to this, the following extension to the methodology is proposed, for implementation when sufficient data exists to implement it.

The Smart Meter population is young, and existing theft work shows that there is an approximate lead time of 8 years until all thefts that are going to be detected have been detected. This timescale may be reduced for Smart Meters due to the more detailed information that comes from them, but this is yet to be proven.

This phenomenon will not affect the Investigation → Detection rates for Smart Meter population categories, but it will affect the number of leads, i.e. where “young” thefts have not yet produced enough suspicious meter readings to generate a lead, and so they will not yet be investigated and detected. Therefore, an adjustment for this effect will be possible when there is enough data to support it, which would use the meter installation date record for each detected theft in the training data year with the following logic applied:

- For Smart Meters up to 1 year old, only P_1% that will generate a lead have yet done so.
- For Smart Meters 1-2 years old, P_2% that will generate a lead have done so.
- ⋮
- For Smart Meters 7-8 years old, P_8% that will generate a lead have done so.
- For Smart Meters over 8 years old, all that will generate a lead will have done so.

where $P_8 > P_7 > … > P_1$

The installation date field in the detected theft data allows the total number of Smart Meter leads from the raw data (for any given population category) to be further stratified by meter age. The factors P_1, P_2, …, P_n as defined above will then be applied to these stratified figures to scale the number of (raw, untargeted) leads to what they would have been if the population was mature, i.e. an estimate of the actual number of Smart Meter sites with suspicious meter reads rather than just those it has been possible to identify at this early stage. Raw Smart Meter leads from all sources will first be scaled in this manner to give a set of revised targeted leads before the processes detailed above are applied to remove bias and output the sets of unbiased leads, investigations and detections.

This calculation will only be possible when sufficient data exists to support it and allow theft detection rate curves (similar to those used in the existing Detected Theft calculation for the whole population) to be generated specifically for Smart Meters. In the data currently supplied, there are 305 confirmed thefts from Smart Meters out of a total of 9000 confirmed thefts. This is not currently sufficient to implement the Smart Meter adjustment, but this will be assessed again during the calculation of the 2020/21 factors.

If the reasonable assumption is made that propensity to steal is steady over time, the Smart Meter population is the only one that requires this pre-adjustment. The AMR and traditional meter populations are mature and so in these cases the issue will not occur.

7.9.6 Process Summary

Figure 23 below shows a simplified graphical representation of the theft analysis process steps. The start point for the process is the dataset containing all leads, with the final output being the split of the Balancing Factor in line with proportions of unbiased theft (from each EUC/Product Class category) for the forecast year. The additional potential steps of the Smart Meter population adjustment, which will be carried out when sufficient data is available, are shown in grey.

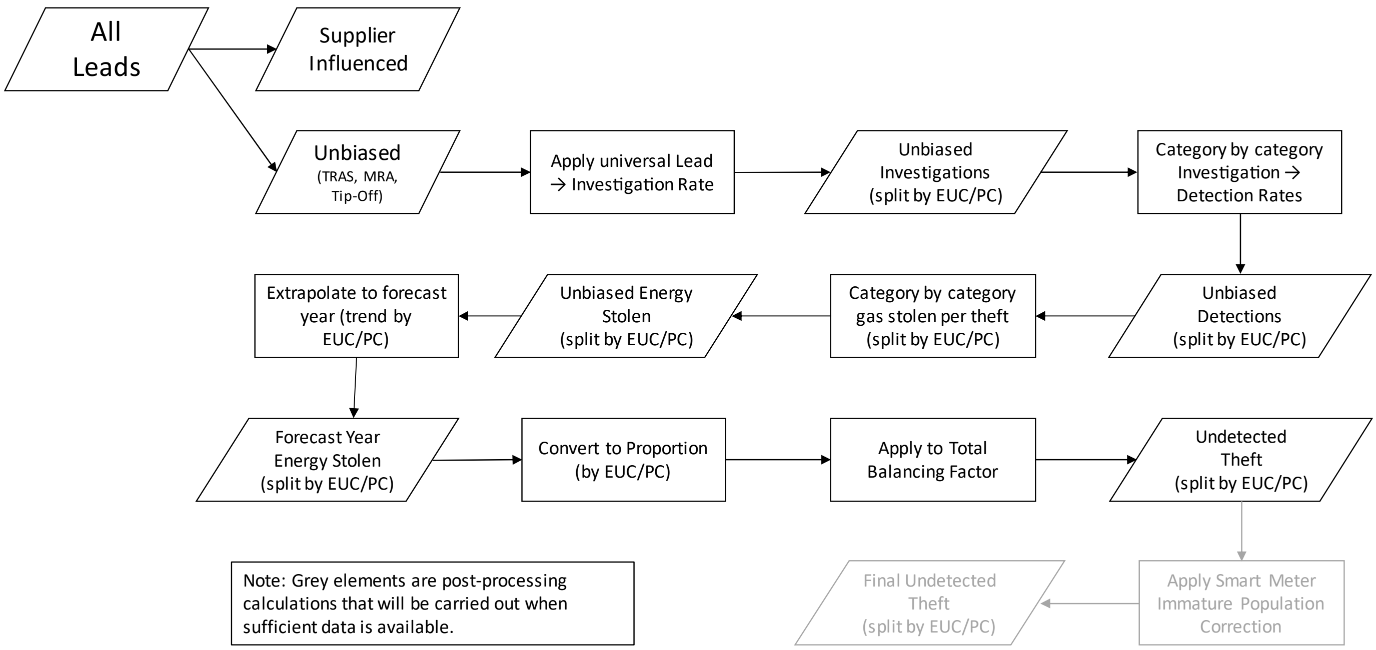

Figure 23: Theft Analysis Process Steps

7.9.7 Balancing Factor Split

Based on the methods described above, the split of the Balancing Factor between EUCs and Product Classes is as shown in Table 11 below.

	01B	02B	03B	04B	05B	06B	07B	08B	09B
Product 1	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Product 2	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Product 3	0.01%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Product 4	95.20%	4.70%	0.07%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table 11: Balancing Factor Split

7.10 Product Class Population and Aggregate AQ

The total population of meter points (either in terms of meter point count or aggregate AQ), split by EUC and Product Class, is used in a number of parts of the Unidentified Gas calculations. In all cases the figures used are those for April 2020, which are based on a number of instances of the asset data provided by the CDSP with figures then extrapolated to the midpoint of the forecast year. This method is described in Section 5.5 above. The areas these figures are used in are as follows:

- The principles for splitting IGT CSEPs and Shipperless/Unregistered Unidentified Gas between EUCs and Product Classes are based on meter point counts.
- The split of the Balancing Factor is based on the projection of unbiased theft to the forecast year, with the projection based on population change.
- The aggregate AQ by category is used as the denominator in the UIG factor calculations.

The instances of the asset data available begin at Nexus go-live in June 2017. The most recent data available is for February 2019. They allow trends to be established for each EUC and Product Class combination for extrapolation to the forecast year, and they also allow step changes to be detected and discounted from trends (such as the compulsory transfer of non-mandatory DM sites from PC1 to PC2).

CSEPs are included in all populations. The current AUG year is the first for which these populations have been available on a site-by-site basis for direct inclusion in the figures: in previous years, an estimation process had to be applied to aggregate data (at the EUC level) for CSEPs.

Table 12 and Table 13 below show both the count of meters and aggregate AQ split by Product Class and EUC for Jun 2017 (actual data, Nexus go-live) and the forecast year (extrapolated data).

Number of Sites

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	21	15	18	44	59	203	189	207	267	1,023
Product 2	0	1	0	0	2	0	2	0	0	5
Product 3	53,920	12	1	0	1	0	0	0	0	53,934
Product 4	23,625,140	192,795	45,803	19,449	4,818	1,478	505	187	6	23,890,181
Total	23,679,081	192,824	45,822	19,493	4,881	1,681	696	394	273	23,945,143

Aggregate AQ (GWh)

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	0	3	9	65	251	2,108	3,895	8,899	39,309	54,539
Product 2	0	0	0	0	11	0	48	0	0	59
Product 3	771	1	1	0	3	0	0	0	0	776
Product 4	315,171	26,569	20,772	23,314	16,387	13,333	10,146	7,435	436	433,563
Total	315,943	26,573	20,782	23,379	16,652	15,440	14,089	16,334	39,745	488,937.0

Table 12: Meter Point Population and Aggregate AQ, June 2017

Number of Sites

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	0	0	0	0	0	0	0	0	309	309
Product 2	14	10	14	34	53	175	178	201	0	679
Product 3	62,824	20,108	8,997	2,246	419	125	44	20	0	94,783
Product 4	23,764,072	179,074	35,672	16,777	4,312	1,402	503	190	0	24,002,002
Total	23,826,910	199,192	44,683	19,057	4,784	1,702	725	411	309	24,097,773

Aggregate AQ (GWh)

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	0	0	0	0	0	0	0	0	39,672	39,672
Product 2	0	2	8	48	225	1,834	3,699	8,427	0	14,244
Product 3	1,130	3,147	3,985	2,616	1,402	1,121	881	789	0	15,070
Product 4	318,855	23,932	16,045	20,198	14,608	12,535	10,201	7,496	0	423,870
Total	319,985	27,081	20,038	22,862	16,235	15,490	14,781	16,712	39,672	492,855

Table 13: Projected Meter Point Population and Aggregate AQ, April 2020

7.11 Extrapolation to Forecast Period

The training period for the Unidentified Gas calculations consists of a maximum of five years' data, running from 2011/12 to 2015/16. Data for the Consumption Method for calculating total Unidentified Gas is available for this length of time. Training data for individual directly-calculated elements of Unidentified Gas is limited by the time period for which snapshots are available: for these, the snapshot files typically do not start as early as 2011/12, but run beyond 2015/16, with Shipperless/Unregistered snapshots available until September 2018.

The analysis of Unidentified Gas is carried out on an LDZ by LDZ basis, due to the fact that it is dependent on identifying and extrapolating trends in each Unidentified Gas component and these trends differ across LDZs. The final UIG factors are based on the combined national total Unidentified Gas, however, and so figures are aggregated across all LDZs once they have been calculated to achieve this.

The total Unidentified Gas is first calculated for the training period using the Consumption Method, which is described in detail in Section 7.2, and the permanent Unidentified Gas total for each year is obtained by subtracting temporary Unidentified Gas from it. For the forecast year, the total permanent Unidentified Gas is obtained by extrapolating permanent Unidentified Gas from each category (both the Balancing Factor and the directly calculated elements) forward to the relevant year and summing the resultant values. These Unidentified Gas energy values are then converted into factors using the method described in Section 7.12 below. There is no requirement for an additional independent estimation of the permanent Unidentified Gas total for the forecast year.

The method of extrapolation differs depending on the Unidentified Gas category. This section contains a description of the process for each category to clarify how permanent Unidentified Gas is calculated for the forecast year.

7.11.1 Unregistered and Shipperless

The principles of extrapolating this element of Unidentified Gas are based around creating piecewise trends for each Unregistered/Shipperless category and LDZ/EUC/Product Class combination and using these to make estimates for the training years and the forecast year. Piecewise fits are required because of the effects of Mods 0410A, 0424 and 0425, all of which affect Unregistered or Shipperless Unidentified Gas. The implementation of these Mods during the training period results in the Unidentified Gas pattern changing at specific points (e.g. the implementation dates). As many individual time segments as necessary are used to accurately model the Unidentified Gas pattern for each combination. Many segments may be required for an individual sub-category, as described and illustrated in Section 7.3.

The fit for the most recent snapshots is used to extrapolate forward to the forecast year in each case with the maximum degree of accuracy. In order to capture all Unregistered/Shipperless categories, LDZs, EUCs and Product Classes, a total of 1872 piecewise trends are required and hence it is not possible to present them in this document. Excel files containing all of this information are available to code parties on request.

The only exception to the above method is for the "Awaiting GSR Visit" category, which uses different input data. For this category, it is assumed that for the forecast year sites will continue to become Shipperless at the present rate. Within this context, the effects of Mods 0424 and 0425 are then modelled as follows:

- For the PTS element of "Awaiting GSR Visit" Unidentified Gas, all Unidentified Gas from this source is temporary from 2014/15 onwards.
- For the SSrP element of "Awaiting GSR Visit" Unidentified Gas, following Nexus go-live, any site without a record of the installation date of the new meter will contribute permanent Unidentified Gas until its GSR visit regardless of its Isolation date. It is assumed that new meters are installed at a steady rate for these sites and as such the average period of the year for which a meter was present and capable of flowing gas is 6 months (i.e. 50% of the year). The following calculation steps are applied:

- Multiply the Shipperless SSrP Unidentified Gas element by the "% of Year with Meter" factor. This gives the total (permanent plus temporary) Unidentified Gas for the Shipperless SSrP element of "Awaiting GSR Visit".
- Multiply this total by the "% of Meters without Install Date" factor. This gives the Permanent element of the Unidentified Gas total from this source. The remainder of the Unidentified Gas for the Shipperless SSrP element of "Awaiting GSR Visit" is temporary.

7.11.2 IGT CSEPs

The principle of defining trends for the training period, as described above for Unregistered/Shipperless sites, is also used for IGT CSEPs. In this case, 468 individual trends are required to cover each LDZ/Product Class/EUC combination. Whilst this approach provides a method of extrapolating to the forecast period, which was used in the Unidentified Gas analyses for previous years, this is not necessary for this Unidentified Gas category post-Nexus and the trends are only used for the training period. This is due to the fact that only permanent Unidentified Gas is calculated for the forecast period and under Nexus all IGT CSEPS Unidentified Gas is temporary.

For the training period, only Unidentified Gas arising from EUC 01B is permanent, whilst the remainder is temporary. It is not possible to present all of the trends in this document, but Excel files containing this information are available on request.

7.11.3 Consumer Metering Errors

For previous AUG years, this process has been assumed to be steady, and the figures calculated from the latest dataset have been applied across all training and forecast years. A number of instances of this dataset are now available, however, which allows a trend to be established for each EUC/Product Class combination. These trends are used to extrapolate to April 2020, this midpoint of the forecast year, and these figures are used in the forecast Unidentified Gas calculations.

7.11.4 Detected Theft

Detected theft is mostly temporary in nature and hence is used only in the calculations for the training period. Any permanent Unidentified Gas arising from theft detected after the line-in-the-sand is taken forward to the forecast year via the Balancing Factor.

7.11.5 CSEP Shrinkage

Shrinkage estimates for the forecast year are available in the annual statements from the Gas Transporters [19], and these are used as the basis for the forecast year calculations. The best estimate for the percentage CSEP Shrinkage is applied to these and then split by EUC/Product as described in Section 7.7 above.

For the CSEP Shrinkage calculation, for the training period only a single total for Unidentified Gas from this source is required. For the forecast year a split by EUC/Product is required, and this is done on the basis of CSEP throughput (as opposed to overall throughput). The latest available figures for CSEP throughput split by EUC and Product (as described in Section 7.4.3 above) are used for this calculation.

7.11.6 Volume Conversion Error

The volume conversion error is calculated individually for each gas year of the training period. The average from the five years of the training period (2011/12 to 2015/16) has been used as the value to take forward to the forecast year. This value is 75 GWh.

For the forecast year a split by EUC/Product is required. This is based on the aggregate AQ of meters without volume conversion by EUC and product Class. The PC1 contribution is zero as it has been assumed that all PC1 meters have volume conversion devices.

7.11.7 Balancing Factor

The Balancing Factor is calculated individually for each gas year of the training period by subtracting temporary Unidentified Gas plus all directly estimated permanent Unidentified Gas from the total Unidentified Gas derived using the Consumption Method. However, it should be noted that the partitioning into gas years is arbitrary and relies on the allocation algorithm to assign consumption to individual gas years. It is therefore important to use multiple gas years in the calculation of the Balancing Factor to correct for any misallocation of consumption between gas years. The average from the five years of the training period (2011/12 to 2015/16) has been used as the value to take forward to the forecast year. This value is 5,393 GWh.

7.11.8 Total Permanent Unidentified Gas

The total permanent Unidentified Gas for the forecast year, split by EUC/Product Class, is the sum of the Balancing Factor and each directly calculated category of Unidentified Gas. Each of these has been estimated for the forecast year as described above.

The forecast of total permanent Unidentified Gas for 2019/20 is 5,958 GWh.

7.12 UIG Factor Calculation

The final step in the calculation process is the production of the UIG factors. These are a fundamental link between the population of the EUC/Product Class combination and the Unidentified Gas from it.

They are therefore calculated using the detailed estimates of the value of Unidentified Gas (for the year in which the factors will be in force) described in Section 7.11 above. Once the Unidentified Gas for each EUC/Product combination for the forecast year has been estimated, this is converted into a factor by dividing it by the aggregate AQ for that EUC/Product (i.e. the best estimate of the AQ for that EUC/Product combination for the forecast year). This AQ data has been supplied to the AUG Expert by the CDSP.

Therefore, each factor is calculated as follows:

$$\textit{UIG Factor}_{PRODUCT,EUC} = \textit{Unidentified Gas (GWh)}_{PRODUCT,EUC} \,/\, \textit{Aggregate AQ (TWh)}_{PRODUCT,EUC} \times 100$$

The Unidentified Gas term and the Aggregate AQ term have different units in this calculation (GWh and TWh respectively) and the raw output is also multiplied by a factor of 100. These steps are to ensure that the resulting factors give sufficient precision when expressed to 2 decimal places, which is a requirement for the output.

The final step of the factor calculation is to apply a smoothing process as described in Section 4.3.5 above. This is achieved using individual cubic fits for each Product Class and ensures a smooth transition of the factors from EUC to EUC, removing any random variation that has arisen from the statistical modelling process. The final factors presented in Section 8 below represent the output from the smoothing calculations.

8 UNIDENTIFIED GAS FACTORS

The final Unidentified Gas factors, calculated as described in this document, to apply to the 2019/20 gas year are as follows:

Supply Meter Point Classification	Class 1	Class 2	Class 3	Class 4
EUC Band 1	0.20	4.07	24.23	163.68
EUC Band 2	0.20	4.07	15.33	110.79
EUC Band 3	0.20	4.07	10.20	17.92
EUC Band 4	0.20	3.89	7.71	12.51
EUC Band 5	0.20	3.50	6.75	7.87
EUC Band 6	0.20	2.86	6.20	4.31
EUC Band 7	0.20	1.96	4.93	2.14
EUC Band 8	0.20	0.78	1.82	1.70
EUC Band 9	0.20	0.20	0.20	0.20

Table 14: Final AUG Table

These factors are based on updated consumption estimates calculated in 2018 using the most recent meter asset data and meter read data. They also reflect the change in the consumption calculation to use interpolated values where possible (See Section 5.3). It should however be noted that the consumptions were calculated over the same gas years as for the last AUG year (2011 to 2015) as the AUG Expert felt that the additional year would still be subject to a significant number of reconciliations. There are also a significant number of missing meter reads. This issue has been highlighted and the CDSP is looking to provide this additional data for use in the next AUG year.

The most striking feature of these figures is that the factors for EUCs 01B and 02B in Product Class 4 are much higher than those for any other category. In comparison to the factors calculated for the first draft of this AUG Statement:

- The factor for Product Class 4 EUC 03B has dropped from being at a similar level to EUCs 01B and 02B, and is now of a similar magnitude to the remainder of the factors.
- The factors for Product Class 3 have dropped considerably.
- The factors for Product Classes 1 and 2 remain at a similar level to the first draft.

The reason for all of these changes is the implementation of the new theft method, which provides a much more accurate estimate of where theft really takes place. This has shown that the vast majority of detected thefts come from Product Class 4 EUC 01B, and that even when the effects of Supplier theft detection strategy are removed – providing the best estimate of the breakdown of undetected theft – this remains the case. The following statistics support this in particular:

1. Of 9000 detected thefts, 8841 came from PC4 EUC 01B. 157 were from PC4 EUCs 02B-03B, whilst only 2 were from PC3. There were no thefts at all from PC1 or PC2 sites.
2. 78% of all thefts were from ETM devices. These meters are overwhelmingly in PC4 EUC 01B.

3. 307 thefts were from Smart Meter or AMR. 302 of these were EUC 01B, and 305 were PC4 despite having the technology to be in a different Product Class.

Table 15 below shows the relative contribution to the permanent UIG(f) total for the forecast year from each Unidentified Gas source.

Unidentified Gas Source	Contribution (%)
CSEP Shrinkage	0.24%
Shipperless/Unregistered	1.52%
Consumer Meter Errors	0.47%
Volume Conversion	1.89%
Total Directly Measured	*4.12%*
Balancing Factor	95.88%

Table 15: Contribution to Unidentified Gas by Source

9 CONSULTATION QUESTIONS AND ANSWERS

This section captures a history of the questions raised by industry bodies during consultation periods and the AUG Expert's responses. These currently relate to the 2017/18, 2018/19 and 2019/20 AUG Statements. The questions have been assessed against the AUGE Guidelines [1] and responses provided as appropriate. All questions and answers have also been published on the Joint Office website.

Due to the in-depth nature of the questions raised and the detailed responses required, it is not appropriate to publish full transcripts in this document. Instead, this section contains a summary of the organisations that provided questions. The questions themselves and their associated responses can be found in the following external documents:

- "AUGS 2017-18 BG Response - DNVGL Comments" [20]
- "AUGS 2017-18 ICoSS Response - DNVGL Comments" [21]
- "AUGS 2017-18 E.ON Response - DNVGL Comments" [22]
- "AUGS 2018-19 British Gas - DNV GL Comments" [46]
- "AUGS 2018-19 Corona Response - DNV GL Comments" [47]
- "AUGS 2018-19 Eon Response - DNV GL Comments" [48]
- "AUGS 2018-19 ICOSS Response - DNV GL Comments" [49]
- "AUGS 2018-19 General Query - DNV GL Comments" [50]
- "AUGS 2019-20 British Gas - DNV GL Comments" [52]
- "AUGS 2019-20 ICoSS - DNV GL Comments" [53]
- "AUGS 2019-20 Scottish Power - DNV GL Comments" [54]
- "AUGS 2019-20 Shipper1 - DNV GL Comments" [55]

Note that all responses contained in these documents relate to the Unidentified Gas calculations at the time they were written, rather than reflecting the process as it currently stands. Therefore, wherever information differs between the responses and the latest AUGS, this is because the Unidentified Gas analysis has evolved and information in the response documents has been superseded. The information supplied in the latest version of the AUGS is always the most up-to-date.

Organisation Name	Date of Communication
British Gas	15/03/2017
ICoSS	15/03/2017
E.ON	15/03/2017
British Gas	09/04/2018
Corona	09/04/2018
E.ON	09/04/2018
ICoSS	09/04/2018
Undisclosed	18/04/2018
British Gas	21/01/2019

Organisation Name	Date of Communication
ICoSS	22/01/2019
Scottish Power	04/02/2019
Undisclosed	21/01/2019

Table 16: Responses to the Proposed 2017/18, 2018/19 and 2019/20 AUG Statements

10 CONTACT DETAILS

Questions can be raised with the AUG Expert at AUGE.software@dnvgl.com

11 REFERENCES

[1] Framework for the Appointment of an Allocation of Unidentified Gas Expert V7.0, 20th April 2017

[2] Uniform Network Code – Transportation Principal Document Version 5.04, 7th June 2017

[3] Mod 0432: Project Nexus – Gas Demand Estimation, Allocation, Settlement and Reconciliation reform, 17 January 2014

[4] Mod 0473/0473A: Project Nexus – Allocation of Unidentified Gas, 20 November 2014

[5] Standard Conditions of Gas Supply Licence (Gas Act 1986)

[6] Review of Shrinkage Calculations in Shrinkage Leakage Model, DNV GL, June 2017

[7] Mod 0424: Re-establishment of Supply Meter Points – prospective measures to address Shipperless sites, 20 December 2012

[8] Mod 0425V: Re-establishment of Supply Meter Points – Shipperless sites, 7 January 2014

[9] Mod 0410/0410A: Responsibility for gas off-taken at Unregistered Sites following New Network Connections, 12 July 2013

[10] 2011 Allocation of Unidentified Gas Statement for 2012/13

[11] Mod 0594R: Meter Reading Submission for Advanced & Smart Metering, 15 September 2016

[12] Mod 0572: Amendment to the definition of AUG Year within UNC TPD Section E

[13] IGT 053: Introduction of annual updates to the AQ values within the CSEP NExA table, 1 October 2014

[14] Mod 0398: Limitation on Retrospective Invoicing and Invoice Correction (3 to 4 year solution), 1 April 2014

[15] Joint Gas Distribution Network (GDN) Response to the Energy UK Gas Retail Group Study into the effect of shrinkage on domestic customers, November 2016

[16] Energy UK Gas Retail Group Study into the effect of shrinkage on domestic customers, October 2015

[17] Shrinkage Leakage Model Review, December 2017

[18] Department for Business, Energy & Industrial Strategy: Smart Meters – Quarterly Report to end September 2018, 29 November 2018

[19] LDZ Shrinkage Quantity Proposals, 2009/10 to 2017/18, http://www.gasgovernance.co.uk/Shrinkage/Shrinkage-Quantity-Proposals

[20] AUGS 2017-18 BG Response - DNVGL Comments, April 2017

[21] AUGS 2017-18 ICoSS Response - DNVGL Comments, April 2017

[22] AUGS 2017-18 E.ON Response - DNVGL Comments, April 2017

[23] Allocation of Unidentified Gas Statement for 2017/18, DNV GL, 28/04/2017

[24] Gas Energy Measurement in Consumer Billing, D.F. Lander, 23 July 2014

[25] Accuracy of CV Determination Systems for Calculation of FWACV, D.F. Lander, 22 May 2012

[26] The Gas (Calculation of Thermal Energy) Regulations 1996

[27] Code of Practice for Gas Meter Asset Managers, Edition 10, 3 Nov 2017

[28] https://www.getthedata.com/open-postcode-elevation

[29] Correction Factor Application, UK Link Programme Change Request Form, UKLP CRDBI065, 9 Jan 2015, http://www.gasgovernance.co.uk/sites/default/files/ggf/UKLP065%20-%20Correction%20Factor%20application%20CR.pdf

[30] Written evidence submitted by Ray Cope to the Energy and Climate Change Committee, January 2013, https://publications.parliament.uk/pa/cm201314/cmselect/cmenergy/108/108vw02.htm

[31] Modification NC0396, Proposal to allow for RbD Adjustments to be processed consequent to the recalculation of LDZ Specific Shrinkage Factors for revised temperature data for the period from the implementation of RbD to the start of current Supply Year 1999/2000, Transco, 4 August 2000, http://www.gasgovernance.co.uk/index.php/NCMP/400

[32] Final Proposed LDZ Shrinkage Factors North East and Northern LDZ Gas Year 2005/06, Northern gas Networks, 30 August 2005

[33] Suspension of the Validation between Meter Index and Unconverted Converter Index, DSC Change Proposal XRN4621, 23 February 2018

[34] Gas energy measurement, A consultation document, Ofgem, November 2000

[35] Summary Report for Domestic Gas Temperature Survey for Shippers (2000), BG Technology

[36] The derivation of LDZ gas temperatures for the period 1996-2000, L M Wallis & D Gregory, Advantica R5181, February 2001

[37] Examination of environmental factors affecting gas metering accuracy, L M Wallis, BG Technology, 1998

[38] The domestic meter temperature survey, BG Technology, 6 July 2000

[39] Framework for the Appointment of an Allocation of Unidentified Gas Expert, V8.1, 8 June 2018

[40] Electronic gas volume conversion systems, IGEM/GM/5 Edition 3

[41] Gas meters – Conversion Devices, Part 1: Volume conversion, BS EN 12405-1:2005+A2:2010

[42] Misreading of Conventional Diaphragm Domestic Gas Meter, Rene Hermkens, KIWA Technology, 2012 Kuala Lumpur World Gas Conference

[43] Framework for the Appointment of an Allocation of Unidentified Gas Expert V8.1, 8 June 2018

[44] Open letter: Gas energy measurement in consumer billing, Ofgem, 21 August 2014

[45] UIG Task Force Update V1.1, Xoserve, 13 Nov 2018, https://www.xoserve.com/wp-content/uploads/UIG_Nov_Sian_Update-v1.1.pdf

[46] AUGS 2018-19 British Gas - DNV GL Comments, April 2018

[47] AUGS 2018-19 Corona Response - DNV GL Comments, April 2018

[48] AUGS 2018-19 Eon Response - DNV GL Comments, April 2018

[49] AUGS 2018-19 ICOSS Response - DNV GL Comments, April 2018

[50] AUGS 2018-19 General Query - DNV GL Comments, April 2018

[51] UNC642 and 642A: 'Changes to settlement regime to address Unidentified Gas issues'; and UNC643: 'Changes to settlement regime to address Unidentified Gas issues including retrospective correction'. OFGEM, 7th June 2018

[52] AUGS 2019-20 British Gas - DNV GL Comments, March 2019

[53] AUGS 2019-20 ICOSS Response - DNV GL Comments, March 2019

[54] AUGS 2019-20 Scottish Power Response - DNV GL Comments, March 2019

[55] AUGS 2019-20 Shipper1 Response - DNV GL Comments, March 2019

[56] Gesche M. Huebner, Megan McMichael, David Shipworth, Michelle Shipworth, Mathieu Durand-Daubin & Alex J. Summerfield (2015) The shape of warmth: temperature profiles in living rooms, Building Research & Information, 43:2, 185-196, DOI:10.1080/09613218.2014.922339

[57] UK shallow ground temperatures for ground coupled heat exchangers, Jon Busby, British Geological Survey

GLOSSARY

AGI	Above Ground Installation
ALP	Annual Load Profile (deeming algorithm parameter)
AMR	Automated Meter Reading
AQ	Annual Quantity. An estimate of annual consumption under seasonal normal conditions
AUG	Allocation of Unidentified Gas
AUGE	Allocation of Unidentified Gas Expert
AUGS	Allocation of Unidentified Gas Statement
Balancing Factor	An aggregate of the combined Unidentified Gas of various items calculated by subtraction. This includes theft, errors in the shrinkage estimate, open bypass valves, meters "Passing Unregistered Gas", unknown sites, and additional common cause variation.
CMA	Competition & Markets Authority
Consumption Method	Unidentified Gas methodology using meter reads and metered volumes
CSEP	Connected System Exit Point
CV	Calorific Value
CWAALP	Cumulative Weather Adjusted Annual Load Profile
CWV	Composite Weather Variable
DAF	Daily Adjustment Factor (deeming algorithm parameter)
DM	Daily Metered
DME	Daily Metered Elective. A site below the DM mandatory threshold of 58,600,000 kWh which the shipper has elected to be DM. The meter read equipment is provided by the shipper.
DMM	Daily Metered Mandatory. A site with an AQ above the DM mandatory threshold of 58,600,000 kWh.
DMTS	Domestic Meter Temperature Survey
DMV	Daily Metered Voluntary. A site below the DM mandatory threshold of 58,600,000 kWh which is voluntarily DM. The meter read equipment is provided by the transporter.
ECV	Emergency Control Valve
EMIB	Energy Market Issues for Biomethane Projects
EUC	End User Category
EWCF	Estimated Weather Correction Factor (deeming algorithm output - alternative to WCF based on CWV rather than demand)

Found Meter	A meter being supplied by a Shipper but for which the CDSP have no record
GFD	Gas Flow Day
GSR	Gas Safety Regulations
IGT	Independent Gas Transporter
Isolated Meter	A meter that has been disabled (through capping or clamping) and hence is no longer capable of flowing gas, and this information has been conveyed to the CDSP and recorded on their system.
LDZ	Local Distribution Zone
LSP	Larger Supply Point
MAM	Meter Asset Manager
Model Error	The statistical error associated with any modelling or estimation process. It an inherent part of any statistical model and does not imply that the model itself is inadequate or incorrect.
MPRN	Meter Point Reference Number
MSL	Mean Sea Level
NDM	Non-Daily Metered
ODR	OFGEM Data Request
OUG	Own Use Gas
PC	Product Class
PSND	Pseudo Seasonal Normal Demand, calculated using AQ values rather than being based on historic metered demands
PTS	Passed To Shipper
RbD	Reconciliation by Difference
RPD Meter	Rotary Positive Displacement meter
SAP	System Average Price
SF	Scaling Factor (deeming algorithm output)
SNCWV	Seasonal Normal Composite Weather Variable
SND	Seasonal Normal Demand
SPC	Supply Point Component
SSP	Smaller Supply Point
SSrP	Shipper Specific rePort
TPD	Transportation Principal Document (of UNC)
UIP	Utility Infrastructure Provider
UNC	Uniform Network Code

UIG	UNC definition of total Unidentified Gas calculated on a daily basis as part of the settlement process
WAALP	Weather Adjusted Annual Load Profile
WCF	Weather Correction Factor (deeming algorithm output)
WSENS	Weather Sensitivity (deeming algorithm parameter used in EWCF definition, reflecting the sensitivity of an EUC to difference in CWV from seasonal normal)

APPENDIX A
Raw Data

This appendix describes the raw data provided by the CDSP for the Consumption Method.

ALLOCATIONS

This data contains all allocations including CSEPs from 01/10/2008 onwards.

Name	Description
GAS_DAY	Date - Gas day for which allocation applies
LDZ	Char[2] - LDZ identifier e.g. EA
EUC	Char[11] - Full EUC Code e.g. WM:E0708W02
ALLOCATED_ENERGY	Number - Final allocated energy value (kWh). Includes CSEPs

ANNUAL_QUANTITY

This data includes all meter points active at any point from 01/10/2008 onwards, not just those currently live. It includes all within gas year updates, appeals etc.

Name	Description
MPR_ID	Number – Unique dummy ID for meter point which is used consistently throughout the data
AQ_EFFECTIVE_DATE	Date - Date on which AQ becomes effective
EUC	Char[11] - Full EUC Code e.g. WM:E0708W02
AQ	Number - Annual Quantity to apply from effective date (kWh)
SITE_TYPE_FLAG	Char[1] - Indicator ="N" for NDM meter point, "D" for DM meter point or "U" for Unique site
CLASS	Number – Meter point product class

CSEPS

This data contains information for gas year 2008 onwards.

Name	Description
GAS_YEAR	Date - Gas year for which CSEP AQ/Numbers apply
EUC*	Char[11] - Full EUC Code e.g. WM:E0708W02
TOTAL_AQ	Number - Aggregate CSEP AQ at start of gas year

COUNT_OF_SUPPLY_POINTS	Number - Count of supply points at start of gas year

* Note that the EUC classification for CSEPs is based on a nominal maximum AQ

The CDSP also provide the success rate for the AQ calculation process. This is used in the process to adjust CSEP consumption for AQ bias.

FACTORS

This data is provided from 01/10/2008

Name	Description
LDZ	Char[2] - LDZ identifier e.g. EA
EUC	Char[11] - Full EUC Code e.g. WM:E0708W02
GAS_DAY	Date - Gas day for which factors applies
ALP	Number - Annual Load Profile (6 d.p.)
DAF	Number - Daily Adjustment Factor (6 d.p.)
EWCF	Number - Estimated Weather Correction Factor (8 d.p.)
CV	Number - Calorific Value (1 d.p.)

METER_ERRORS

All meter error adjustments from 01/04/2008 onwards. In addition to these resolved meter errors, any open errors will be taken from the JoT website. Where appropriate, the AUG Expert will clarify these with the CDSP.

Name	Description
BILLING MONTH	Month and year where billing correction was applied for the given meter error
LDZ	Char[2] - LDZ identifier e.g. EA
AGGREGATE ENERGY	Total correction for period of error in kWh
REASON	Reason for adjustment
ADJUSTMENT	Value of adjustment in kWh over billing period. Positive value represents an over-read

METER_READS

This data includes all meter reads from 01/10/2008 onwards. Multiple records for a meter point with the same date are filtered by the CDSP using the following methodology.

Where there is an A (Actual) Read Type and an E (Estimate) Read Type the CDSP remove the E and retain the A Read. Where there are Read Types of R (Replacement) the CDSP retain this read and remove the original read type that it replaced. Where there are multiple R Reads they are ranked by number e.g. R01 and R02 and the highest number is the latest replacement read that is retained.

Name	**Description**
MPR_ID	Number - Unique dummy ID for meter point which is used consistently throughout the data
START_READ_DATE	Date – Start date of metered period
METER_READ_DATE	Date - Date of meter read
IMP_IND	Char[1] - Indicator ="Y" for imperial meter read, else "N"
METER_READ_VAL	Number - Value of meter read
METERED_VOL	Number - volume of gas since previous meter read in units appropriate for meter (imperial or metric)
ROUND_THE_CLOCK_IND	Number – Number of times the meter index has passed zero since the last read.
AQ	Number – Prevailing Annual Quantity at time of meter read (kWh)
METER_READ_FREQ	Char[1] - Indicator for frequency of meter reads (A-Annual, 6-6 monthly, M-monthly)
SSP_LSP	Char[3] - "SSP" or "LSP"
EUC	Char[11] - Full EUC Code e.g. WM:E0708W02
READ_TYPE_CODE	Char[4] - Code for type of meter read

METER_INFO

This data includes all available meter asset data.

Name	**Description**
MPR_ID	Number - Unique ID for meter used across ALL data
LDZ	Char[2] - LDZ identifier e.g. EA
NUM_DIALS	Number - Number of meter dials
IMP_IND	Char[1] - Indicator ="Y" for imperial meter read, else "N"
METER_FITTED_DATE	Date - Date meter was fitted
UNITS	Number - Multiplier for meter read units (1, 10, 100 etc)

CORRECTION_FACTOR	Number – Volume correction factor
STATUS_UPDATE_DATE	Date – Date of record
METER_STATUS	Char[27] – e.g. Live, Removed
READ_FACTOR	Number
METER_MECHANISM	Char[30]
CSEP_FLAG	Char[1] - Indicator ="Y" if meter point is on a CSEP, else "N"

NEW_LOST_SITES

This data contains all meter points with a first confirmation date or an end date from 01/10/2008 onwards.

Name	**Description**
MPR_ID	Number - Unique dummy ID for meter point which is used consistently throughout the data
START_DATE	Date - First confirmation date for meter point
END_DATE	Date - Date meter point was excluded from allocations process

PRIMES

This data includes details of all prime meter points active at any time since 01/10/2008.

Name	**Description**
MPR_ID	Number - Unique dummy ID for meter point which is used consistently throughout the data
LDZ	Char[2] - LDZ identifier e.g. EA

SUBS

This data includes details of all sub-prime meter points active at any time since 01/10/2008.

Name	Description
MPR_ID	Number - Unique dummy ID for meter point which is used consistently throughout the data
PRIME_MPR_ID	Number – Dummy ID for the prime meter

A list of re-confirmation dates has also been provided for meter points which were previously in a prime/sub configuration but are no longer.

APPENDIX B
Data Pre-processing

This appendix describes the various processes used to derive information from the raw data prior to carrying out the Consumption Method calculation.

LDZ Movers

Some meters have records associated with more than one LDZ. These meters are identified and their MPR_ID is recorded in a separate table which also includes the latest LDZ they are recorded against and any previous LDZs. Any data associated with a meter which is not for the latest LDZ is then copied to the appropriate LDZ. This is required as processing is done on an LDZ by LDZ basis with data held in separate LDZ specific tables. This copying of data ensures that when the meter consumption is calculated in the latest LDZ, all of the relevant data is present. To ensure no double counting occurs, the SITE_LIST tables are populated to ensure that each meter is only present in one LDZ (its latest LDZ).

Identification of Bad Meter Reads

The algorithm for flagging bad reads is as follows:

Given sequential meter reads mr_1, mr_2, mr_3 and mr_4 calculate:

$$con_1 = mr_2 - mr_1$$

$$con_2 = mr_3 - mr_2$$

$$con_3 = mr_4 - mr_3$$

If any of these are negative, we check for meter index rollover

If the meter was replaced, we leave the consumption null

Then if the meter was not replaced during the period we check

- If ($con_3 > 0$) and ($con_2 < 0$) and ($con_1 > 0$) then we have a bad reading
- If $con_1 > \text{abs}(con_2)$ then mr_2 is bad
- Else if $con_3 > \text{abs}(con_2)$ then mr_3 is bad

Identification of Meter Read Gaps and Overlaps

For each meter read the number of days between the start read date for a record and the meter read date from previous record is calculated.

- A positive value indicates a gap
- A negative value indicates an overlap

Note, start read dates are only available for the records provided from 2013 onwards.

For the records with a gap/overlap the difference between the volume consumed between that record and the previous read (calculated using the meter read values and relevant asset data) and the recorded metered volume is calculated and stored.

NDM DM Changes

Using the site type flag from the AQ records meter points are identified which are recorded as both DM and NDM. The start and end date (where relevant) of the meter point's status as NDM are then recorded.

APPENDIX C
Example Temperature Data

The AUG Expert has identified a number of sources of relevant temperature data which are provided below for information. The wide range of values highlights the uncertainty in the temperature.

The Domestic Meter Temperature Survey (DMTS)

Although the AUG Expert does not have access to the results of the DMTS (data or reports), a summary of the findings are given as part of a later study into gas temperatures[36]. A summary of the DMTS temperatures are given in Table 17 and Table 18. Table 17 shows the temperatures split between internal and external meters over 2 years by LDZ. This highlights the large difference in temperature between internal and external meters (5-6°C). The location of the meter therefore has a significant impact on the accuracy of the conversion from volume to energy.

Data from the CDSP shows 38.8% of meters are located internally and a further 9.77% are in sheltered locations (e.g. garage).

LDZ	1998/1999		1999/2000	
	Internal T	External T	Internal T	External T
EA	15.50	8.90	15.50	9.00
EM	13.80	8.90	13.40	8.70
NE	14.70	8.80	14.00	9.00
NO	14.00	8.30	13.40	8.30
NT	16.70	9.70	16.30	9.80
NW	13.30	8.80	12.94	8.70
SC	16.90	7.70	16.60	7.70
SE	17.10	9.90	15.90	9.80
SO	16.00	9.60	15.30	9.60
SW	15.30	9.50	14.10	9.40
WM	13.40	9.00	12.60	9.00
WN	15.30	9.00	11.60	9.10
WS	14.90	9.90	14.00	9.70
Average	**15.15**	**9.08**	**14.28**	**9.06**

Table 17: Summary of Measured Temperatures from DMTS

Table 18is a summary of the average flow-weighted temperatures from the DMTS. However, there is no information on how these values were calculated and a footnote states that the DMTS start and end dates do not represent a whole year[36].

LDZ	Average Flow-Weighted T
EA	10.40
EM	10.60
NE	10.80
NO	9.60
NT	13.80
NW	10.70
SC	10.70
SE	12.60
SO	12.00
SW	10.20
WM	10.40
WN	9.40
WS	10.20
Average	**10.88**

Table 18: Flow-Weighted Temperatures from DMTS

Temperatures used in LDZ Shrinkage

The study into gas temperatures carried out in 2002[36] also provides the temperatures assumed when temperature corrections were being applied through LDZ Shrinkage. These are provided in Table 19.

LDZ	1999/2000	1998/1999	1997/1998	1996/1997
EA	11.50	11.50	11.40	11.20
EM	12.80	13.10	13.10	12.90
NE	11.30	11.40	11.20	11.10
NO	11.00	10.90	10.70	10.20
NT	15.70	15.90	16.10	16.10
NW	11.40	11.40	11.30	10.90
SC	10.80	10.80	10.60	10.20
SE	12.90	12.80	12.70	12.20
SO	11.80	11.80	11.90	11.80
SW	12.30	12.30	12.60	12.50
WM	11.20	11.20	11.10	10.80
WN	11.10	11.00	10.80	10.10
WS	13.20	13.50	13.40	13.20
Average	**12.08**	**12.12**	**12.07**	**11.78**

Table 19: Temperatures used by Transco for LDZ Shrinkage Determination

Seasonal Normal Effective Temperatures (SNET)

Pseudo SNET data for gas year 2016/17 was provided by the CDSP. This is based on aggregate NDM demand as well as weather. It is a specific feature of the current (and previous) CWV definitions intended to beneficially alter the underlying shape of the CWV profile throughout the year and thereby help mitigate seasonal modelling bias. Table 20 is a summary of average pseudo SNET by LDZ. This data is not flow-weighted and represents an approximation to air temperature rather than gas temperature in the meter. It therefore does not include any adjustment to allow for the internally located meters.

LDZ	Average SNET
SC	9.30
NO	8.91
NW	10.50
NE	10.25
EM	10.25
WM	10.38
WN	10.50
WS	11.03
EA	11.86
NT	11.86
SE	11.86
SO	11.98
SW	11.07
Average	10.75

Table 20: Seasonal Normal Effective Temperature Summary

Actual Daily Average Air Temperatures

Actual bi-hourly temperature data for gas years 2011 to 2016 was provided by the CDSP. Table 21 is a summary of average (un-weighted) temperature by LDZ and gas year. This data is not flow-weighted and represents air temperature rather than gas temperature in the meter. It therefore does not include any adjustment to allow for the internally located meters.

LDZ	2011	2012	2013	2014	2015	2016
EA	11.62	10.65	12.29	11.50	12.43	11.85
EM	10.04	9.00	10.59	9.75	10.60	10.26
NE	10.04	9.00	10.59	9.75	10.60	10.26
NO	8.61	7.91	9.33	8.45	9.02	9.05
NT	11.62	10.65	12.29	11.50	12.43	11.85
NW	10.32	9.22	10.67	9.73	10.12	9.59
SC	9.21	8.28	9.67	8.58	9.31	9.29
SE	11.62	10.65	12.29	11.50	12.43	11.85
SO	11.89	10.88	12.49	11.71	12.59	11.99
SW	10.92	9.99	11.50	10.67	11.68	11.07
WM	10.23	9.00	10.56	9.77	10.80	10.38
WN	10.32	9.22	10.67	9.73	10.12	9.59
WS	11.04	10.10	11.62	10.75	11.63	10.97
Average	10.57	9.58	11.12	10.26	11.06	10.62

Table 21: Actual Average LDZ Temperature Summary

About DNV GL

Driven by our purpose of safeguarding life, property and the environment, DNV GL enables organizations to advance the safety and sustainability of their business. We provide classification and technical assurance along with software and independent expert advisory services to the maritime, oil & gas and energy industries. We also provide certification services to customers across a wide range of industries. Operating in more than 100 countries, our professionals are dedicated to helping our customers make the world safer, smarter and greener.