ALLOCATION OF UNIDENTIFIED GAS

Revised Allocation of Unidentified Gas Statement for 2018/19

Report No.: 1, Rev. 2.1
Date: 24th May 2018

Project name:	Allocation of Unidentified Gas	DNV GL - Software
Report title:	Revised Allocation of Unidentified Gas Statement for 2018/19	Software Consulting
Customer:		Holywell Park
Customer contact:	Fiona Cottam	Ashby Road
Date of issue:	24th May 2018	Loughborough
Project No.:	4011094	LE11 3GR
Organisation unit:	Software	Tel: +44 (0)1509 282000
Report No.:	1, Rev. 2.1	

Objective: This document is the revised AUG Statement for 2018/19 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 provides an estimate of the Final UIG (total quantity of permanent UG) and identifies each UG source (the proportionate contribution of each UG Source to the total is provided separately).

This document contains a revised estimate of the AUG weighting factors based on the latest available data and the documented methodology which includes any changes since the first draft AUG Statement.

Prepared by:	Verified by:	Approved by:
Tony Perchard Principal Consultant	Clive Whitehand Head of Section	Anthony Gilbert Industry Solutions 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, UIG

Rev. No.	Date	Reason for Issue	Prepared by	Verified by	Approved by
1	2018-02-01	First draft	Mark Lingham, Andy Gordon & Tony Perchard	Clive Whitehand	Anthony Gilbert
2	2018-04-30	Revised version including consultation feedback	Mark Lingham, Andy Gordon & Tony Perchard	Clive Whitehand	Anthony Gilbert
2.1	2018-05-24	Updated with text changes agreed at UNCC meeting on 17 May 2018	Tony Perchard	Clive Whitehand	Anthony Gilbert

EXECUTIVE SUMMARY

Project Nexus has now been implemented, replacing key IT systems for gas settlement and supply point administration in the gas industry ('UKLink') and changes 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 and allocation processes.

This document is the revised AUG Statement for 2018/19 and is an update to the first draft following the 42 day industry consultation period. It 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) Identification of each Unidentified Gas Source.

b) An estimate of the total quantity of permanent Unidentified Gas. The proportionate contribution of each Unidentified Gas Source to the total quantity of Unidentified Gas is provided separately in the calculation summary spreadsheet [33].

c) A revised estimate of the AUG weighting factors based on the latest available data and including updates to the methodology since the first draft AUG Statement.

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 (UG).

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 UG between the LSP and SSP market sectors: UG 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 UG equitably to the relevant gas market sectors. This approach involved an annual estimate of UG and a monthly transfer of costs between market sectors to address the misallocation of UG that occurs under the current 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 will be shared out to all live sites, on the basis of their recorded/estimated throughput for the day.

It should be noted that 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 AUGE 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 87% of sites with the remaining 13% 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 Final UIG (i.e. UIG at line in the sand, when all reconciliations have been processed). 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 by Xoserve, the Central Data Services Provider (CDSP) 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 UG and the data available/required from industry bodies to evaluate UG
- 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 UG 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 revised AUG Statement for 2018/19 and contains the following:

- A detailed description of the proposed methodology
- 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 UG and used to develop the methodology

This document will be published to the industry for approval.

The following are out of scope.

- The AUG Expert is not concerned with issues regarding the deeming algorithm or the 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 [1].

The AUG Expert will create the AUG Statement 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 UG where this is possible e.g. Shipperless sites. The AUG Expert has also developed a methodology to account for elements of UG which are Temporary in nature.

Additional data regarding theft of gas (the SPAA Schedule 33 data) was sought and obtained from 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.

The proposed methodology and assessment of what constitutes UG has been decided solely by the AUG Expert based on available information. Comments raised by shippers relating to the previous AUG Statement have been considered and additional analysis has been undertaken. Having considered all views, the final decisions are the AUG Expert’s own.

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 appropriate data supports 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.

Questions regarding the relative ease of theft and detection of theft from Smart Meters vs mechanical meters were submitted to the industry. Further communication will take place as and when necessary.

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

The CDSP has provided all the latest available data as requested by the AUG Expert. Further updates will be provided if available for use in the calculation of the final table of factors.

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 UG, both meter reads and metered volumes are provided. Pre-Nexus LSP metered volumes were subject to corrections, but the meter reads were 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. This decision will be reviewed when the training period contains post-Nexus data.

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 UG 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 in draft form (the "Draft 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 Draft and final AUG Statement to the Gas Transporters for publication.

This revised AUG Statement for 2018/19 is provided to the Gas Transporters for publication by 30th April 2018. This document will become the final AUG statement following the May UNCC meeting where the committee can vote to approve this document as it stands or unanimously agree any changes required.

The Committee's final determination in this process shall be binding on Users.

This section is not relevant to AUG Expert.

The AUG Expert will undertake to ensure that all data that is provided to it by all 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.

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.

3 TERMINOLOGY

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

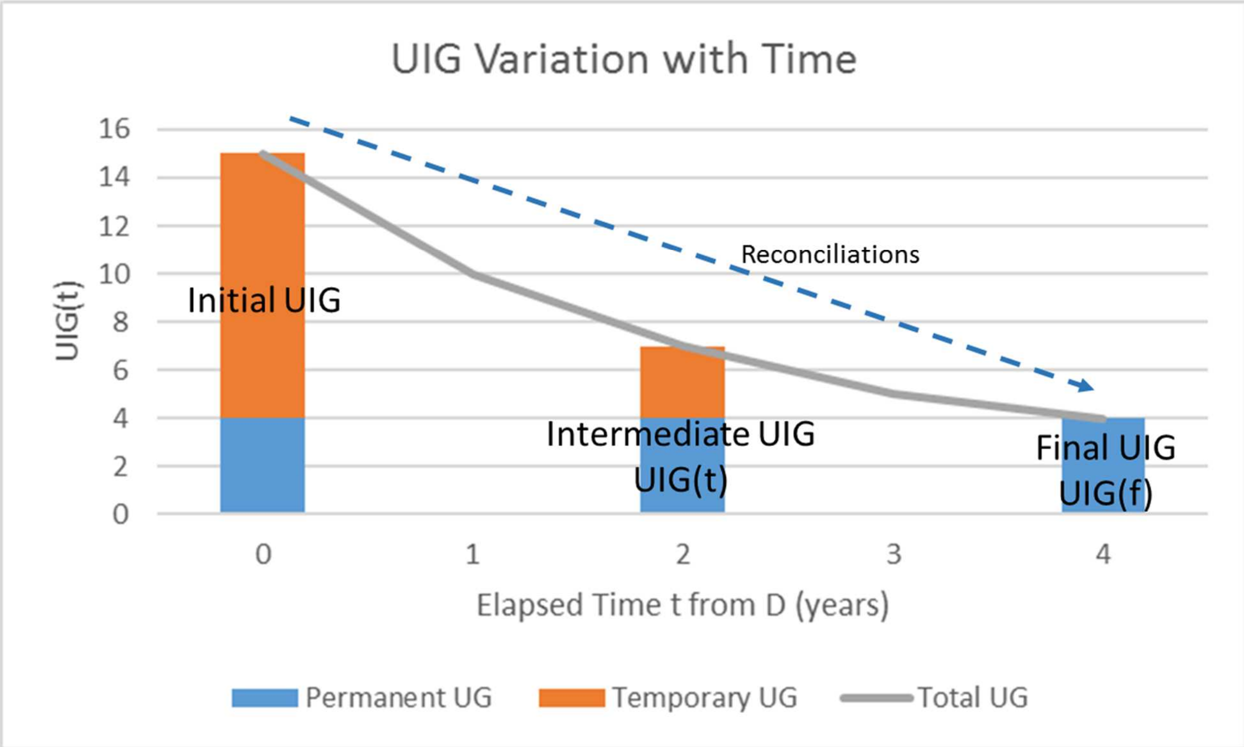

Figure 1: Unidentified Gas Definitions

Unidentified Gas (UG) refers to gas which is consumed in an unrecorded manner. Estimates of UG for a given gas day (D) are initially made on the day as part of the settlement process. This "Initial" value of UG 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 UG as it is corrected for over time through reconciliation. There are further sources of UG 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 UG.

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 UG.

Throughout this document, Unidentified Gas (UG) will refer to the general concept of unidentified gas, 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 UG presented here a more detailed discussion is given in Section 7.

Under the old settlement regime, an independent forecast of permanent UG split by market sector was made to allow correct allocation of UG 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 2018/19 analysis, for which actual Product Class data is available. For completeness, and to provide a link with the historic UG 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 were 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 were 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 were 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 the table 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 are 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**

UG 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

UG is currently unknown and hence must be estimated.

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

Total UG = Aggregate LDZ Load – DM Load – NDM Load– Shrinkage (3.1a)

This can be reformulated for the post-Nexus regime as:

Total UG = Aggregate LDZ Load – Product 1 to 4 Load – Shrinkage (3.1b)

4.2 Permanent and Temporary Unidentified Gas

Unidentified gas can be divided into two categories:

Permanent UG is consumed in an unrecorded fashion and costs are never recovered.

Temporary UG 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 UG for Historic Years

The overall concept of calculating total UG using metered consumption data is simple, and is centred on the basic principle of the allocation process. Note that all available historic consumption data 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 2018/19 gas year for which the factors are being produced. The pre-Nexus 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 3.1a is the sum of all metered NDM consumptions, this allows us to rewrite it as

$$Total\ UG = NDM\ Allocation - Metered\ NDM\ Consumption \qquad (3.2)$$

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

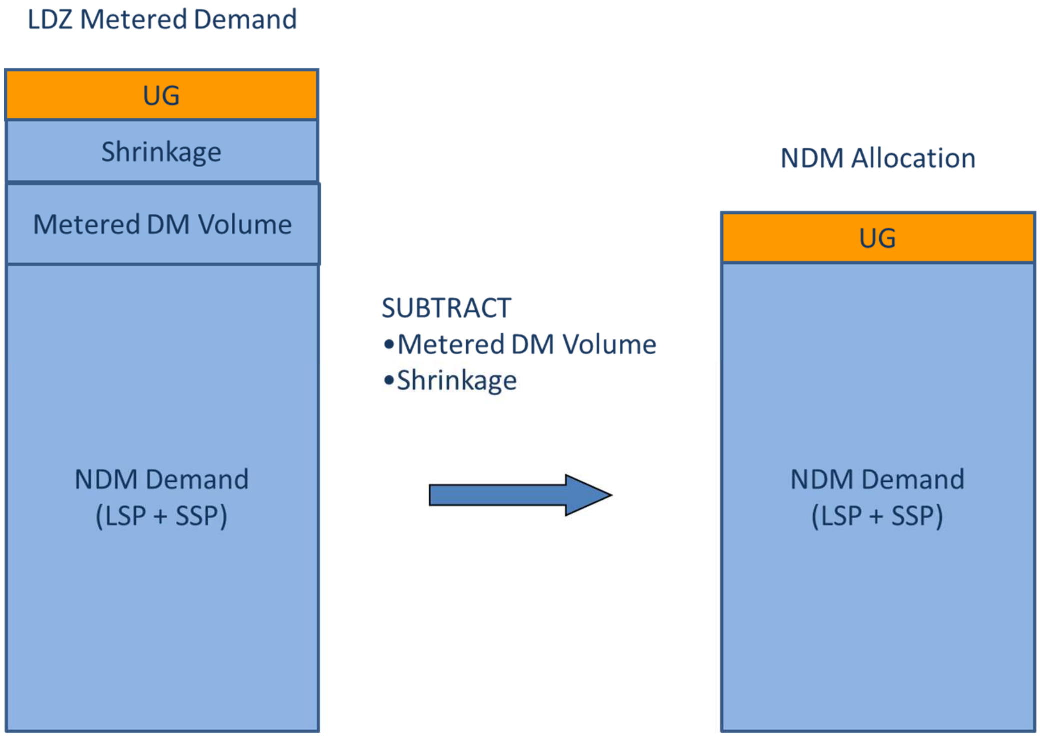

Figure 2: NDM Allocation and Unidentified Gas

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.

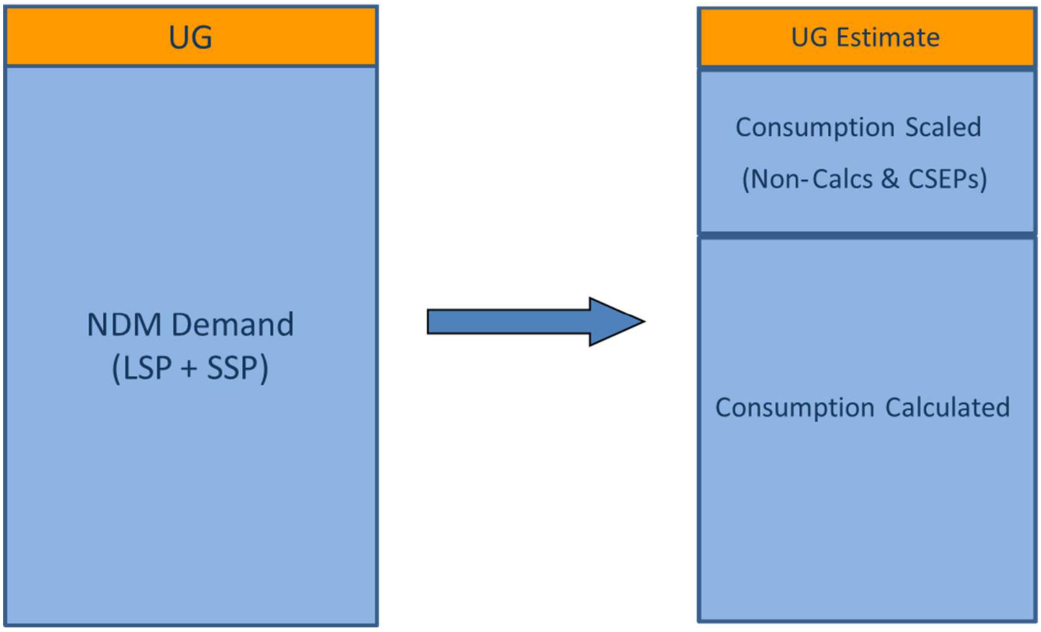

Figure 3: Calculation of Unidentified Gas from Consumptions and Allocations

This method is used to estimate total UG 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 process is used to estimate the total UG for each of the five most recent historic years for which reliable data is available. This excludes the most recent year because the number of meters where consumption can be successfully calculated is much lower due to fewer meter reads being available. The use of data from this year could therefore be subject to a large degree of uncertainty. Of the gas years which are used, the most recent years have varying degrees of reconciliations included and so represent Intermediate UIG. The older gas years for which line in the sand has passed (pre 2014) represent Final UIG.

4.3.2 Calculating Components of Total UG

Having obtained the total UG figure using the Consumption Methodology described above, the value of individual components that make up the UG total are calculated where this is possible. This also includes the calculation of the amount of this UG which is temporary for each component and how the UG 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 entirely 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 UG 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 UG 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 UG forecast must be carried out separately not only for each UG 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 five potential components of UG (Unregistered and Shipperless sites, IGT errors, CSEP shrinkage, Shipper-responsible theft and metering 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 UG component based on the quality of information available for that component.

Brief descriptions of each UG 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 UG 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 UG 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 UG calculation. UG 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 UG 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 UG. The Shrinkage estimate comes from the Shrinkage Model, and if this is biased it will affect the UG estimate.

Following a recent change to the AUGE Framework the estimation of Shrinkage Error by the AUGE has been ruled out of scope. Therefore, the Shrinkage estimate provided by the GT's 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.

For further discussion of this issue see Section 5.18.

4. Shipper-Responsible Theft

Limited reliable data on theft exists, and whilst information for detected and alleged 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 6 below), and considered over time, it forms the vast majority of that figure.

5. Meter Errors

Meter errors affect UG in different ways depending on their source. Errors in LDZ offtake metering and DM supply metering affect the estimate of total NDM demand including UG for the training period (which is all pre-Nexus), whilst NDM LSP and SSP metering errors contribute to UG 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. Balancing Factor

The Balancing Factor is calculated by taking the difference between the calculated total UG and the sum of the directly estimated components. The Balancing Factor is comprised of UG 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 UG is then given by the sum of the Balancing Factor and the Permanent components of the directly calculated components (see Figure 3).

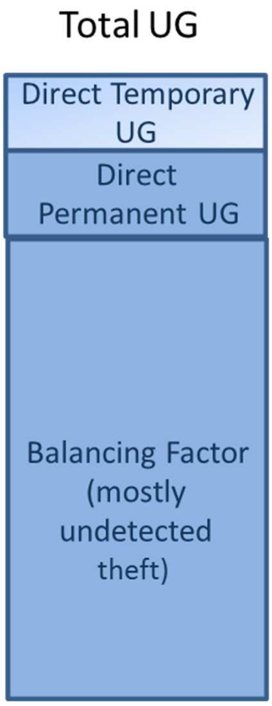

Figure 4: UG Components

4.3.3 Mapping to Post-Nexus Product Classes

The classification of UG 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 all of the training data comes from historic (i.e. pre-Nexus) years, during which time the Product Classes did not exist.

Therefore, a method of mapping existing sites to post-Nexus Product Classes is required. For the 2017/18 AUG Statement, 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 the current analysis (i.e. the statement for 2018/19) 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. Even in cases where this approach is not possible (e.g. Unregistered Sites, which by definition do not have a Product Class assigned), the actual Product data can be used to create an approximate mapping that is still far more accurate than the estimates used for 2017/18.

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

Detected Theft, Consumer Meter 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. The next step is to use as many instances of this population dataset as are available to create a trend over time. At the time of writing, data exists for Nexus go-live and November 2017. The calculated trend is then used to extrapolate the population split to the forecast year. Finally, apply this split to the Unidentified Gas calculated to arise 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. It is also recognised, however, that theft levels from Smart Meters/AMR and traditional meters are unlikely to be the same and hence a simple split by population, as described above, is not appropriate. Therefore, in this case the latest Smart Meter installation figures are used and extrapolated to the forecast year using the current installation rate. 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. This information is used together to create a map of the population by meter type and Product/EUC, which is then used as a basis for splitting the Balancing Factor. This process is described in detail in Section 7.8.

4.3.4 Projection of Permanent UG to Forecast Period

Having calculated the best estimates of Permanent and Temporary UG for each historic year for which reliable data is available (the training period), it is then necessary to calculate the projected values of Permanent UG 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 UG 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
- SSP and LSP Metering Errors
- Balancing Factor

The methods used differ based on the observed behaviour of each category of UG, and are in many cases affected by a number of UNC modifications introduced in order to address various UG 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 UG 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.

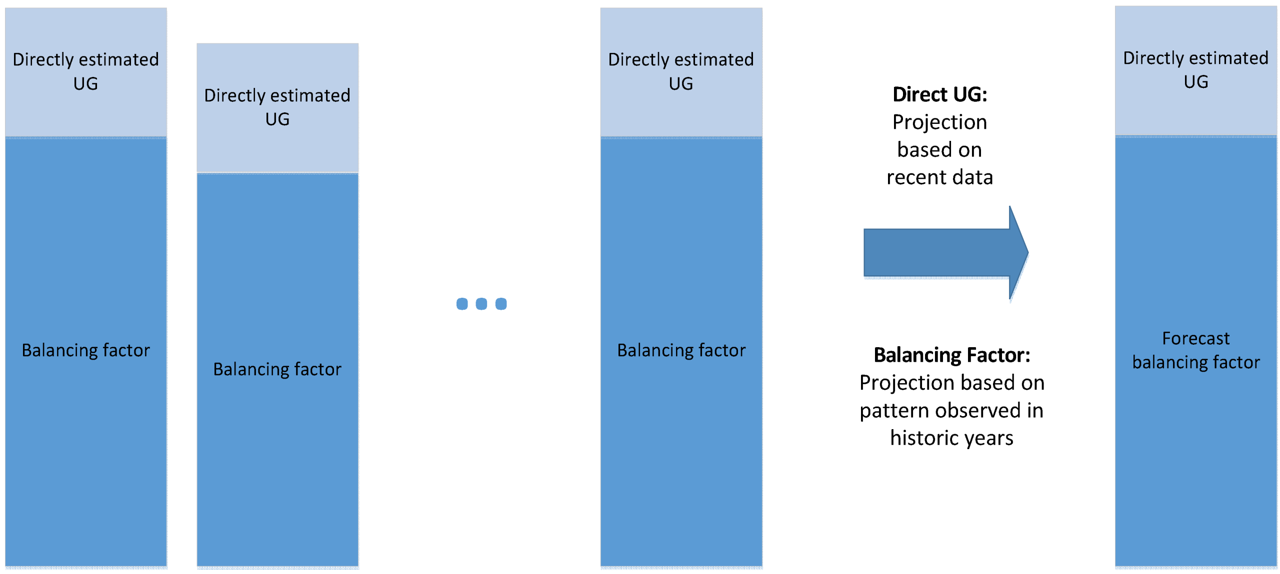

Figure 5: Projecting UG

As part of the estimation of the directly calculated UG components for the training years, an estimate of the amount of Temporary UG for each component is made as described above. The values projected forward to the forecast year are the permanent part of the UG 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 UG (100% Temporary and hence not taken forward to the forecast year).

4.3.5 UIG Factors

The final output of the UG analysis is a set of UIG factors rather than direct estimates of the magnitude of UG 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 UG 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 UG (for the year in which the factors will be in force) described above. Once the Final UIG 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):

$$UIG\ Factor_{PRODUCT,EUC} = 10 * Final\ UIG\ (GWh)_{PRODUCT,EUC} \ /\ Aggregate\ AQ\ (TWh)_{PRODUCT,EUC} \qquad (3.3)$$

Note that the Final UIG and Aggregate AQ have different units, and that a multiplier of 10 is also applied. These approaches are used to 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 UG 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.

5 SUMMARY OF ANALYSIS

This section contains details of analyses carried out and updates made to the UG calculation methodology since the AUGS for 2017/18. It also summarises new UNCC modifications and industry initiatives that could potentially impact on the UG calculation methodology in the future.

5.1 Meter Read Spacing for AQ Calculation

During the consultation period for the first draft 2017/18 AUG Statement, a shipper raised a question concerning the criteria used to select meter reads for estimating a gas year consumption. Specifically, the minimum of 120 days between reads compared to a minimum of 9 months specified by UNC for calculating AQs.

The AUG Expert responses [22] explained that the 120-day window was designed to balance the requirement to successfully calculate consumptions for as many sites as possible against the requirement for these calculated values to be as accurate as possible. The key factor in terms of accuracy being the bias. Any bias in the calculated consumption values will contribute to the total unidentified gas figure.

However, the rules around selection of meter reads have been in place since the 2013/14 AUG statement and numerous improvements have been made to the data used since then. Initial analysis showed that the rules could be tightened up without significantly affecting the calculation success rate.

The AUG Expert has now carried out a review of the four rules used in previous years when choosing meter reads:

- The start read is no more than 540 days from the start of the gas year in question.
- The end read is no more than 540 days from the end of the gas year in question.
- The distance between the two chosen meter readings is at least 120 days.
- The overlap between the metering period and the gas year is at least 60 days.

These have been tightened up to:

- The start read is no more than 365 days from the start of the gas year in question.
- The end read is no more than 365 days from the end of the gas year in question.
- The distance between the two chosen meter readings is at least 170 days.
- The overlap between the metering period and the gas year is at least 90 days.

This has the effect of reducing the average calculation success rate by around two percentage points and altering the estimate of total NDM consumption by 0.06%. The resulting change in the Unidentified Gas Factors is of the order of 0.01.

The historic data used for the consumption calculation shows that meter reading approximately every 6 months is quite prevalent. Increasing the minimum separation between meter reads from 6 to 9 months will result in some of these meters failing the consumption calculation. Of the meters which currently have a consumption successfully calculated, 3% are based on meter readings closer than 9 months apart.

The AUG Expert believes that using 160 days as the minimum meter read separation represents the best compromise between calculation success rate and accuracy. This position is supported by analysis carried out on data provided by the industry (see Section 5.1.1). The minimum meter read separation is now close to the pre-Nexus rules (minimum of 6 month separation between meter reads).Given that the data used to calculate consumption is all pre-Nexus data, this is consistent with pre-Nexus calculations of AQ.

In the post-Nexus world, tighter limits are potentially more appropriate given the change to rolling AQ calculation and improved data quality. Going forward, the AUG Expert will continue to monitor this to ensure that the most appropriate meter read selection rules are used. We also propose looking at alternative replacement values where consumption can't be calculated for a gas year. In most cases, the consumption calculation fails for only one year out of the five years used. In cases where the failed year is not the first or last year, it may be more accurate to use the average of the calculated consumptions for the years either side.

5.1.1 Meter Read Separation Analysis

Following the AUG consultation period in 2017, data and analysis was provided to support the view that a minimum meter read separation of 9 months is most appropriate for AQ calculation. The data provided was for 3,788 meters and included a prevailing AQ (only for 3,180 of the meters) together with a calculated AQ based on meter read separations of 6m, 9m and 12m, with the 12m figure being treated as an "actual" value i.e. it was used as a baseline to assess the accuracy of the other AQs. It is unclear how the 6m and 12m periods to carry out the calculation were chosen. To be a realistic comparison with what actually happens, the periods should have been selected at random to avoid introducing any bias e.g. the 6 month period is always over the summer.

In summary, the previous analysis carried out showed the following

- Using 9 months, the bias (average of the calculated AQ % error) was +0.55% compared to a bias of -0.56% for the 6 month AQ calculation i.e. both calculations show very similar levels of bias though in opposite directions (positive values represent underestimate of AQ compared to 12 month calculation).
- The 9 month calculation shows a much tighter error distribution with a standard deviation of 11.7% compared to 24.2% for the 6 month calculation
- Based on the sites where a prevailing AQ was available, the AQ calculation using 9 months of data was shown to be better than using the prevailing AQ in 62% of cases, but using the 6 month AQ calculation was only better than using the prevailing AQ in 41% of cases

Based on this, it was concluded that it is better to roll-over prevailing AQ values rather than use a calculated AQ based on a 6 month meter read separation (as calculated value was only better in 41% of cases). This is the case if we are interested in getting the most accurate estimate of AQ for an individual site or accurate relative AQs between sites. However, the consumption method is estimating total consumption and is not concerned with individual site consumptions. For this purpose, it is more important to get site consumption estimates which are unbiased. Using the data provided, calculation of the bias for cases where a prevailing AQ exists shows that the bias for the prevailing AQ is -14.3% compared to -0.82% for the AQ calculated using 6 month meter read separation.

Summing the prevailing AQs for all of the sites in the dataset provided overestimates the total by 21% compared to a 2.7% overestimate using the sum of the AQs calculated using a 6 month meter read separation. Our conclusion is therefore that calculating consumption based on 6 month meter read separations is preferable to using biased prevailing AQs in order to estimate total consumption.

The AUG Expert was aware of the ongoing bias in AQs and therefore does not use the prevailing AQ as a replacement value where the consumption calculation fails. The replacement value used is the EUC average. However, this analysis highlights the danger of introducing a bias regardless of what replacement value is used and the AUG Expert firmly believes that using a calculated value wherever possible is preferable.

Given the fact that the bias of the calculated consumption value over 6 months is essentially the same as the bias for the 9 month calculation, we believe that for the purposes of our calculation, the use of a 6 month minimum separation in meter reads is appropriate.

5.2 UNC Modification 0570 – Obligation on Shippers to Provide at least One Valid Meter Reading per Meter Point into Settlement per Annum

This Modification was implemented with effect from 03/11/2017 following Ofgem's decision on 31 Aug 2017.

The Modification states that it seeks to create a condition in the UNC to correspond with the requirement on Suppliers under the Standard Licence Condition (SLC) 21B of their licence to take a meter reading at least once every year for billing purposes. In the case of the UNC this would be for use within gas industry settlement and relate to non-daily metered sites.

This Modification does not have a direct effect on the Unidentified Gas calculations in that it does not affect the way in which the UIG Factors are calculated. It should, however, improve the quality of the data that feeds into those calculations.

In theory, Mod 0570 should ensure that no meters reach the line in the sand without having had a valid meter read submitted. In reality some meters will still not be read but the Modification will minimise this and should result in the line in the sand being reached without a valid meter read having been submitted for only a negligible number of meters. This will result in an increase in the number of meter reads available to the Consumption Method, which is used to calculate the Unidentified Gas total, and will lead to an increase in the accuracy of this calculated figure. No changes to the calculation methodology are required.

It should be noted, however, that the UG calculation process runs a year in arrears and hence this effect will not begin to be seen until the 2019 AUG Statement is published.

5.3 Mod 0625 – Extension from 6 Months to 10 Months for Transfer of Non-Mandatory Sites from Class 1

This Modification was implemented with effect from 28/11/2017 following Ofgem's decision on 24/11/2017.

This Modification will not affect the UG calculations because the forecast year for the current AUG Year starts in October 2018, by which time the transfer period will have expired regardless of whether it lasts for 6 months or 10 months.

5.4 Mod 0631R - Review of NDM Algorithm Post-Nexus

The workgroup report was published on 22 March 2018 and is due to be presented to panel on 19 April 2018. The recommendation is that this request is closed.

The review did not identify any specific improvements that should be progressed, but other modifications have been raised (Mod 0644: Improvements to nomination and reconciliation through the introduction of new EUC bands and improvements in the CWV and Mod 0652: Obligation to submit reads and data for winter consumption calculation).

5.5 Mod 0632S - Shipper Asset Details Reconciliation

This Modification has been implemented with effect from 10 April 2018

It will allow a more accurate assessment of meter type by Product Class, which is important because the Unidentified Gas methodology assigns different theft propensities to different meter types. In addition, it will allow mismatches to be highlighted (sites in Class 1-3 but with no recorded Smart Meter or AMR) for further investigation. As with the above Modifications, this does not directly affect the UG calculation methodology but does improve the quality of the data feeding into it.

5.6 Mod 0633V/0638V – Mandate Monthly Read Submission for Smart and AMR Sites from 01 December 2017/01 April 2018

Modifications UNC633V and 638V are described as aligning shipper Code obligations with the CMA Gas Settlement Order and Workgroup 0594R RFI findings, for mandating Smart and AMR meters to submit reads on a monthly basis from 01 February 2018 (UNC0633V) or 01 April 2018 (UNC0638V). This will enable utilisation of more consumption data throughout industry processes and will drive more accurate settlement and Unidentified Gas (UIG) positions for Shippers.

UNC0638V is to be implemented, creating a start date for mandated monthly Smart Meter and AMR readings of 1 April 2018. Their impact on the UG process is as described for Mod0570 above.

5.7 Mod 0634 – Revised Estimation Process for DM Sites with D-7 Zero Consumption

This Modification was implemented with effect from 20/11/2017 following Ofgem's decision on 16 Nov 2017.

The AUG Expert's Unidentified Gas calculations operate using annual load and hence are not affected by single days of zero consumption. Therefore, whilst this Modification will affect the daily UIG figure, with the intention of improving its accuracy, it does not affect the UIG factors or their calculation method.

5.8 Mod 0635 - Reforms to Incentivise Accurate and Timely DM Reads to Improve the Accuracy of Unidentified Gas Allocation

This Modification has been withdrawn by the proposer.

5.9 Mod 0639R – Review of AUGE Framework and Arrangements

The scope of this review is to consider the current arrangements for the AUGE and to identify any changes needed, based on both feedback given as part of the AUGE review process, as well as any relevant issues Workgroup participants consider should be subject to the review. The Workgroup will also consider the contracting and procurement arrangements set out in the AUGE Framework document.

A workgroup report was to be provided to Panel by 15 March 2018 but a 2 month extension for further assessment has been requested. As this review has not yet been concluded, no changes have been made to the process or deliverables for the current AUG year.

5.10 Mods 642/642A/643 - Changes to Settlement Regime to address Unidentified Gas Issues

These modifications, which act as alternatives to each other, are still live (marked as urgent) and so do not currently affect the Unidentified Gas calculation process. If implemented, these Mods will affect the way that the AUG Expert operates rather than its calculations.

Mod 642 proposes a solution to the current high level and volatility of the daily UIG figure, which involves fixing UIG at a figure of 1.1% of throughput and reintroducing the Scaling Factor to the allocation algorithm to ensure that the allocations sum to the correct LDZ intake total (with 1.1% UIG removed). This solution would involve an expansion of the role of the AUG Expert to produce additional factors to apply to equal-and-opposite reconciliation effects. Mod 643 proposes the same approach but with an additional retrospective correction to roll the financial position back to the point where it would have been had the Modification been implemented on Nexus go-live date.

Mod 642A proposes a solution where an initial UIG figure of 2.5% is split across all Shippers by market share, and a (positive or negative) balancing figure that represents the difference of the calculated daily UIG from the initial 2.5% is split by market share but just across NDM sites. This solution uses throughput-based splits only and hence negates the need for UIG factors (and hence also negates the need for an AUG Expert).

5.11 Mod 0644 – Improvements to nomination and reconciliation through the introduction of new EUC bands and improvements in the CWV

This modification has two parts, as follows:

1. Split the End User Categories (EUC) EUC01B and EUC02B into three and group by prepayment, market sector code of industrial and commercial, and all remaining meter points.
2. Amend the Composite Weather Variable (CWV) to include more than just wind speeds and temperature plus the creation of parameters to flex the Weather Correction Factor (WCF) and/or Daily Adjustment Factors (DAF) where they reach defined tolerances.

Although the modification does not specify an implementation date, it proposes making any changes in time for the start of the 2018 gas year i.e. 1 October 2018.

Item 1 will have an impact on the AUG process in the future. Based upon the proposed EUC groups, we envisage that the current factors can be used for the 2018/19 AUG year, as the high level EUC bands will remain unchanged. However, the consumption methodology works at EUC level and would need updating going forward to make use of the new EUC groups which will have different parameters (ALPs, DAFs etc).

Item 2 could also have an impact on the AUG factor calculation. The estimation of total UG for historic years uses Cumulative Weather Adjusted Annual Load Profile (CWAALP) data to adjust to seasonal normal conditions. Changes to SNCWV/CWV will also change the CWAALP. It is unclear from the modification whether CWAALP will be retrospectively recalculated. This won't change the methodology, but will require an updated dataset to be used in future calculations.

5.12 Mod 0651 – Replacement of the Retrospective Data Update Provisions

Modification 0434 'Project Nexus – Retrospective Adjustment' was approved by Ofgem on 21 February 2014. The Modification provided the ability for Shipper Users to replace Meter Readings and to retrospectively correct data errors associated with Meter Information, Address and Supply Points. The arrangements within Modification 0434 pertaining to the amendment of periodic Meter Readings and the subsequent automatic reconciliation were implemented at Project Nexus Implementation Date (PNID). The other elements of this modification were deferred for implementation and modification 0624R was raised to re-examine the business case for implementing the Retrospective Data Update elements of Modification 0434.

Modification 0651 proposes the following

- Remove from UNC all Retrospective Data Update elements of Modification 0434 (as amended by Modification 0610S).
- Introduce the Retrospective Data Update mechanism identified as Option 4 within Request 0624R.
- Require Shipper Users to provide relevant Meter Information as required by the Central Data Services Provider (CDSP) to enable a one-off industry 'data cleanse' exercise to be conducted.

This UNC Modification does not directly impact the AUG methodology but ensuring that an optimum process is in place for correcting asset data will improve data quality and improve the accuracy of the AUG factors.

5.13 Mod 0652 – Obligation to submit reads and data for winter consumption calculation (meters in EUC bands 3 - 8)

This modification will create an obligation for shippers to submit reads to ensure the appropriate winter consumption calculation takes place, for accurate NDM WAR band profiling.

To calculate a Winter Annual Ratio, shippers need to submit a pair of reads in the winter period (one in Nov – Dec, and a second in Mar – Apr). If either of these reads is not submitted, or fails validation, winter consumption cannot be calculated, and therefore a 'bucket' or default EUC band is assigned.

This UNC Modification does not directly impact the AUG methodology but ensuring accurate meter reads are submitted for the winter period will improve the accuracy of the consumption calculation.

5.14 Mod0654 - Mandating the provision of NDM sample data

This UNC Modification does not directly impact the AUG methodology but should improve the allocation process and therefore accuracy of UIG estimates. Any improvement in the allocation algorithm parameters will also improve the accuracy of the consumption calculation used to estimate total UG.

5.15 Product Class

During the development of the AUG Statement for 2017/18 actual Product Class data was not available because all data came from the pre-Nexus period. Therefore, in this case a method of mapping meter points to Product Classes and extrapolating to the forecast year was developed. This was based around asset information, including Smart Meter populations and installation rate, and a set of rules applied to this. This method is described in detail in the AUG Statement for 2017/18 [23].

Actual Product Class data is available for the current analysis, both for Nexus go-live and as of November 2017. This not only allows MPRNs to be mapped to their actual Product Class for a number of UG components, but also allows the rate of change to be calculated so that populations can be extrapolated to the forecast year 2018/19. Table 3, Table 4 and Table 5 below show both the count of meters and aggregate AQ split by Product Class and EUC for Jun 2017 (actual data), November 2017 (actual data) and the forecast year (extrapolated data). Note that these figures contain the full meter population including CSEPs: non-CSEP figures are from the CDSP asset data, whilst CSEP figures come from a dedicated dataset supplied by the CDSP for this purpose.

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 3: Meter Point Population and Aggregate AQ, June 2017

Number of Sites

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	20	22	18	44	64	207	194	217	277	1,063
Product 2	0	3	1	0	3	3	3	2	1	16
Product 3	55,543	1,049	528	29	1	0	0	0	0	57,150
Product 4	23,623,518	191,759	45,276	19,420	4,818	1,477	504	187	5	23,886,964
Total	23,679,080	192,833	45,823	19,493	4,886	1,688	702	406	283	23,945,193

Aggregate AQ (GWh)

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	0	3	9	65	269	2,146	4,003	9,323	40,255	56,073
Product 2	0	1	1	0	14	40	68	79	66	269
Product 3	809	172	242	25	3	0	0	0	0	1,251
Product 4	315,134	26,398	20,532	23,289	16,386	13,330	10,143	7,429	365	433,006
Total	315,943	26,573	20,783	23,379	16,671	15,515	14,214	16,831	40,686	490,599.0

Table 4: Meter Point Population and Aggregate AQ, November 2017

Number of Sites

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	0	0	0	0	0	0	0	0	299	299
Product 2	17	46	22	44	78	227	211	246	3	894
Product 3	59,111	3,329	1,686	92	1	0	0	0	3	64,222
Product 4	23,619,949	189,478	44,118	19,356	4,818	1,477	504	187	0	23,879,887
Total	23,679,077	192,853	45,826	19,492	4,897	1,704	715	433	305	23,945,302

Aggregate AQ (GWh)

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	0	0	0	0	0	0	0	0	42,338	42,338
Product 2	0	5	10	65	328	2,357	4,353	10,507	210	17,835
Product 3	890	547	771	80	3	0	0	0	210	2,501
Product 4	315,052	26,023	20,002	23,234	16,384	13,323	10,136	7,416	0	431,570
Total	315,942	26,575	20,783	23,379	16,715	15,680	14,489	17,923	42,758	494,244.0

Table 5: Projected Meter Point Population and Aggregate AQ, April 2019

The extrapolation technique works on a category-by-category basis, including CSEPs as described above. It is based on the rate of change for each individual category between Jun 2017 and Nov 2017, with that rate assumed to remain the same until the midpoint of the forecast year, April 2019. The population changes also include effects such as the compulsory transfer of non-mandatory DM sites from Product Class 1 to comply with Mod 0625, as well as SMETS1 meters reverting to dumb when there is a change of supplier. Where the latter occurs, any such meter that was not in Product Class 4 will revert to this class after the supplier change.

5.16 Theft

The analysis of detected theft has, up until now, been based purely on data supplied by the CDSP. For this year's analysis, the SPAA Schedule 33 data has also been made available from SPAA, and this new information has shown a significant disparity in the number of thefts reported and billed between the two data sources. In particular, the SPAA data shows considerably higher levels of detected theft than the CDSP's data, as shown in Table 6 below.

	Confirmed Thefts (CDSP)	Confirmed Thefts (SPAA)	Billed Thefts (SPAA)
2015	2,782	4,165	2,859
2016	2,351	4,052	3,327

Table 6: Detected Thefts: SPAA and CDSP Reporting

This table shows detected theft levels in the SPAA report are around 60% higher than those from the CDSP's data, and in both years more customers were billed for theft than were even reported as theft at all in the CDSP dataset.

This clearly shows that whilst the CDSP's data is far more detailed and contains vital information for the detected theft analysis that the summary-level SPAA report does not, it is also incomplete in terms of the theft occurrences that make it through into this dataset. Therefore, for the 2018/19 analysis, theft levels are initially calculated using the detailed information supplied by the CDSP (as was done in the past) but then scaled to the level in the SPAA report. This ensures that the higher level of detected theft is included in the calculated figures.

It should be noted that even the SPAA data is not complete, as it also depends on information returned by Suppliers. Whilst Suppliers have an obligation to return this data, non-compliance still occurs. Records for 2015/16 show that 67 out of an official total of 100 potential Suppliers returned data, although it is accepted that not all of this total of 100 will be active. There is a strong tendency for larger Suppliers (particularly those with active theft detection regimes) to return data, and as such we believe the available data includes a high proportion of the actual number of detected thefts. Additional information has been requested from SPAA to allow us to verify this.

5.17 Post Nexus AQ Calculation

Several changes have been made to the AQ calculation process as part of the Nexus implementation.

The main change has been the introduction of rolling AQ recalculation on a monthly basis rather than annually. This should lead to AQs being more representative of current demands as they pick up changes more quickly. AQs are used in multiple places throughout the UG calculations and more representative values will improve the accuracy of these calculations. The impact of this change will be small at this stage due to the short period of time that has elapsed since Nexus go-live, but will increase as more post-Nexus data becomes available and is incorporated into the calculations.

The AQ validation rules have been updated and will again lead to more accurate AQs. This will benefit the UG calculations for similar reasons, but the full impact will not be seen for some time.

5.18 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 UG 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. If, in the future, CSEP shrinkage is estimated by the GTs (but still not accounted for elsewhere in settlement) then this figure will be used for Unidentified Gas and will replace the AUGE's own CSEP Shrinkage estimate.

In the UG analysis for 2017/18, the Imperial College estimate that between 2% and 5% of Shrinkage came from CSEPs was used as the basis for the CSEP Shrinkage estimate. For the 2018/19 analysis, a more rigorous approach is used that produces an independent estimate of the level of CSEP Shrinkage. This 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.

5.19 Performance Assurance Framework

The Performance Assurance Framework (PAF) provides the basis for the (Gas) Energy Settlement Performance Assurance Regime. The goal of this is:

“A demonstrably effective settlement regime for the gas industry where no one party adversely impacts another party as a result of its failure to operate to the defined settlement regime”.

It is expected that the PAF will help to reduce the levels of UIG, and also improve data quality by encouraging a higher level of timely meter reads, etc. These initiatives will improve the quality of industry data and therefore help the AUG Expert’s calculations, without affecting the overall methods used.

5.20 Pressure and Temperature Correction Factors

During the 2018 consultation period, a query was raised about the use of standard correction factors and their impact on UG and how it is apportioned between EUC and Product Class [27].

When converting the measured volume flow from a meter to an energy value, a conversion factor is required to account for pressure and temperature differences from the assumed standard conditions (pressure=1013.25 mbars, temperature=15°C) as described in 'The Gas (Calculation of Thermal Energy) Regulations 1996' [29].

Currently, almost all meters employ a standard factor for temperature which assumes an average gas temperature at the meter of 12.2°C. For pressure correction, a standard factor is used for all meters in EUC bands up to and including 03B. For meters in EUC bands 04 and above, the Meter Asset Manager (MAM) is responsible for calculating a site specific volume conversion factor [28].

In November 2000, Ofgem issued a consultation document [30] in which they highlighted potential changes to improve the accuracy of energy measurement, including the use of site specific temperature and pressure conversion factors.

On 21 August 2014, Ofgem published an open letter [31] following a study they commissioned into the conversion of metered volumes into energy [32]. In their letter, Ofgem concluded that "the prescribed method for converting metered volumes of gas into thermal energy remains appropriate". Ofgem also noted "The analysis demonstrates that energy calculation under the prevailing volume conversion factor is accurate to within -1.48% to +1% for 95% or more of the population. Even at the extremes, the energy calculation is accurate to within -1.569% to +2.477%, which is broadly in line with the maximum permissible error for gas meters themselves."From an unidentified gas point of view, there are two areas for investigation

- Total UG – if the correction factors are not representative of 'average' conditions, there will be a bias in the total energy calculated which will form part of UG
- UG split – the error in the correction factors are not evenly distributed between EUCs and product classes resulting in errors in the factors.

The 2014 study carried out by Dave Lander Consulting concluded that in 2011, annual GB energy was underestimated by 0.238%. This was calculated using postcode (outcode) based altitudes from Ordnance Survey (OS) and temperatures from the Met Office. It is unclear how this discrepancy breaks down between the temperature and pressure parts of the correction factor. One of the key assumptions noted in this study was "that gas use for all outcodes across each LDZ is comparable – i.e. that there is a similar mix of domestic, commercial and industrial consumers in each outcode."

For the calculation of the 2018/19 AUG factors, no allowance will be made for this discrepancy which has been demonstrated for a single year only. However, given the potential size of the discrepancy identified, the AUG Expert believes that this area warrants further work and is currently identifying potential sources of data.

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 UG Calculation (Consumption Method)	NDM allocations	Received
	NGM meter reads and volumes	Received
	LDZ, DM and Unique Sites Metering Errors	Received
	Meter Asset Information	Received
	Algorithm data (ALPs, DAFs, EWCFs)	Received
	CV 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	Outstanding
	Further investigation results for large/suspicious sites	Supplied on request
	Mod 0410A supporting data	Supplied on request
	Shipperless sites supporting data	Supplied on request
	Snapshot files (including MPR details)	Ongoing

Analysis Area	Required Dataset	Status
IGT CSEPs	Known CSEP data	Received
	Snapshot files	Ongoing
Meter Error	Meter capacity report	Received
Theft	Detected and alleged theft updated to end March 2017 - CDSP	Received
	SPAA Schedule 33 data – SPAA	Received
Shrinkage	National population statistics for mains by diameter (from DNs)	Requested
Product Classes	Meter point Product Class	Received. Further update requested

Table 7: Data Status Summary

6.2 Total UG 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 2015/16 with partial data for 2016/17.

- 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 (CSEP invoicing files). 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 UG 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 have been part of a Multi-Meter Supply Point (MMSP) during the analysis period.
- 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 UG 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 UG 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 UG 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.

- **Known CSEP Data**

  This file contains data for both registered sites on known CSEPs and Unregistered sites on known CSEPs. It 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 September 2017. 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 2017. 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 both 2015 and 2016. The key fields used in the Theft analysis are:

- Number of theft cases confirmed
- Volume of theft (kWh)

6.7 CSEP 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 Product Classes

This data set consists of:

- 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 November 2017.

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 UG for each year in the training period. This process is very similar to that used by the AUG Expert 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 UG categories are then calculated for the same period: this allows the amount of Temporary UG within the Consumption Method total for each year to be ascertained, and also allows the Balancing Factor (mostly undetected theft) to be calculated. All UG in the Balancing Factor is Permanent.

The data patterns observed in the training period for each UG component, including the Balancing Factor, are used to extrapolate to the forecast year (currently 2018/19) and provide the best estimate of each Permanent element of UG for this year. This is carried out individually for all 36 EUC/Product combinations for every UG category. Finally, these UG estimates are converted into factors by dividing the GWh UG 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 (which currently only uses pre-Nexus data) can be stated in its simplest form as:

$$\textit{Total UG} = \textit{NDM Allocation} - \textit{Metered NDM Consumption}$$

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

7.1 Correcting the NDM Allocation

The pre-Nexus NDM allocation is calculated as

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

Any subsequently detected significant errors in these three components will constitute Temporary UG 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. 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 6 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 data used in the consumption calculation is 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.

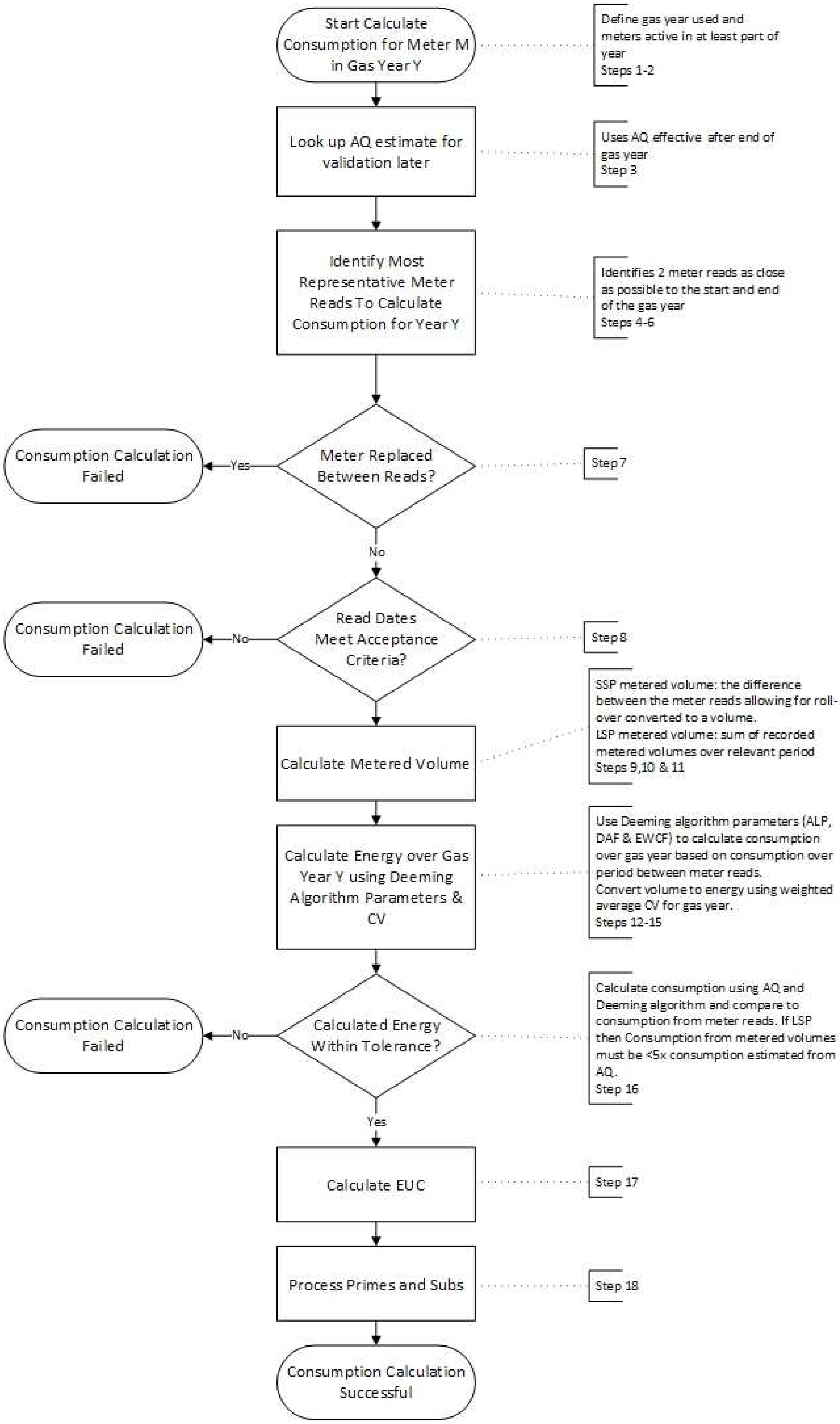

Figure 6: Consumption Algorithm Flow Chart

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 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 7 below:

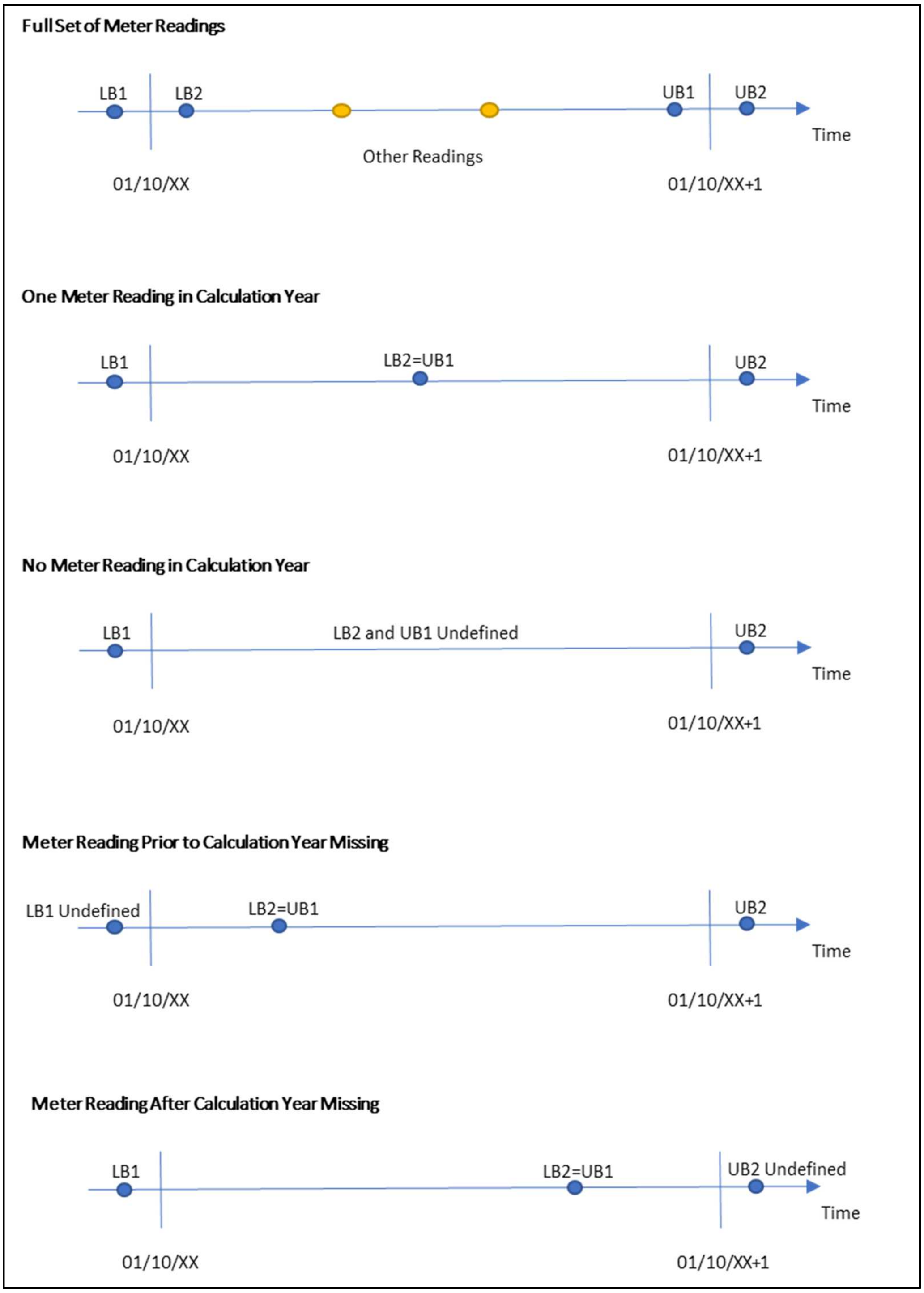

Figure 7: Meter Read Availability Scenarios

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 was replaced (and changed from imperial to metric units) 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

$$\frac{\sum_{\substack{\text{Formula Year} \\ \text{or Part Thereof}}} WAALP^v}{\sum_{\substack{\text{Metered} \\ \text{Period}}} WAALP^v}$$

   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

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

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 as follows. Note that in the case of LSPs, the volumes used have been calculated by the CDSP using site specific correction factors where relevant.

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

$$Con = \sum mv_i * CV / 3.6 \; (* 0.0283168466 \; if\ imperial)$$ 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 five 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 five 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.

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.

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.

7.2.2 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 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).
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 8 below summarises the process for obtaining a consumption value for each type of meter point.

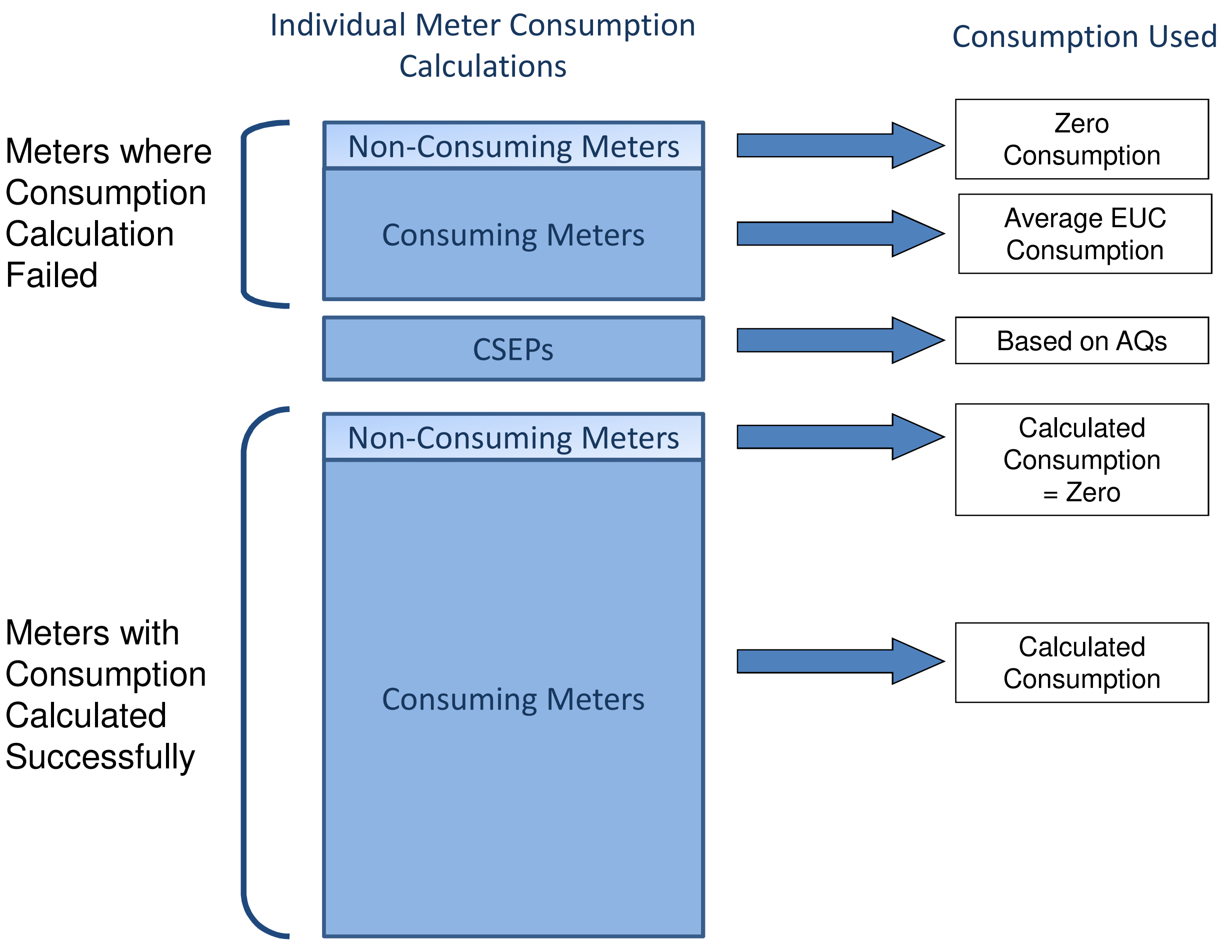

Figure 8: Consumption Method used for each type of Meter Point

UG 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 UG. Therefore, whilst giving an indication of the order of magnitude of the Permanent UG 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.

At the AUGS walkthrough meeting in 2017, it was suggested that the best source of CSEP AQ data would be the CSEP invoicing files, and these would consistently show the AQ values applicable at 1st October each year as required. For this version of the AUG Statement, the calculations have been updated to use this data which is available from gas year 2012 onwards. This removes the need to carry out an adjustment to allow for the fact that the CSEP AQ values have different effective dates for some 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

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

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 UG. 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 UG 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 UG from post-Mod sites being temporary rather than Permanent but the process required to implement this is not available in the pre-Nexus system. Therefore, all UG from post-Mod sites is only Temporary after the Nexus go-live date (i.e. for the forecast year but not the training years).

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 UG 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 UG 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 UG categories, with different fractions for pre-01/09/2013 and post-01/09/2013 sites).
   - Fraction of Unregistered UG not backbilled (for the Permanent/Temporary split for Unregistered UG 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 UG calculations.

6. "Fraction of Unregistered UG 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 UG. 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 UG calculations are now carried out. This is carried out in a series of spreadsheets, with a set of spreadsheets for each individual UG category. Each of these sheets contains a full history of all available snapshot data for the relevant UG category, which at present runs from September 2011 to September 2017. The availability of snapshot data over such a period of time means that trends can be identified within each UG 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 UG 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 UG 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 UG for each main UG 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 UG consumption. This creates a total of 1872 individual sub-categories for every main UG 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 UG 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 UG category, the snapshot period is split into a small number of separate sub-periods where the trend is consistent: these differ for each UG 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 UG 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 UG from that source at that particular point in time. Therefore, the best estimate of the UG consumed in Year Y is the average of the monthly AQ figures across the whole of the year.

Figure 9 below shows an example of the piecewise fit approach for one UG sub-category (permanent Pre-Mod Orphaned sites, Product 4, EUC 01B, EA LDZ).

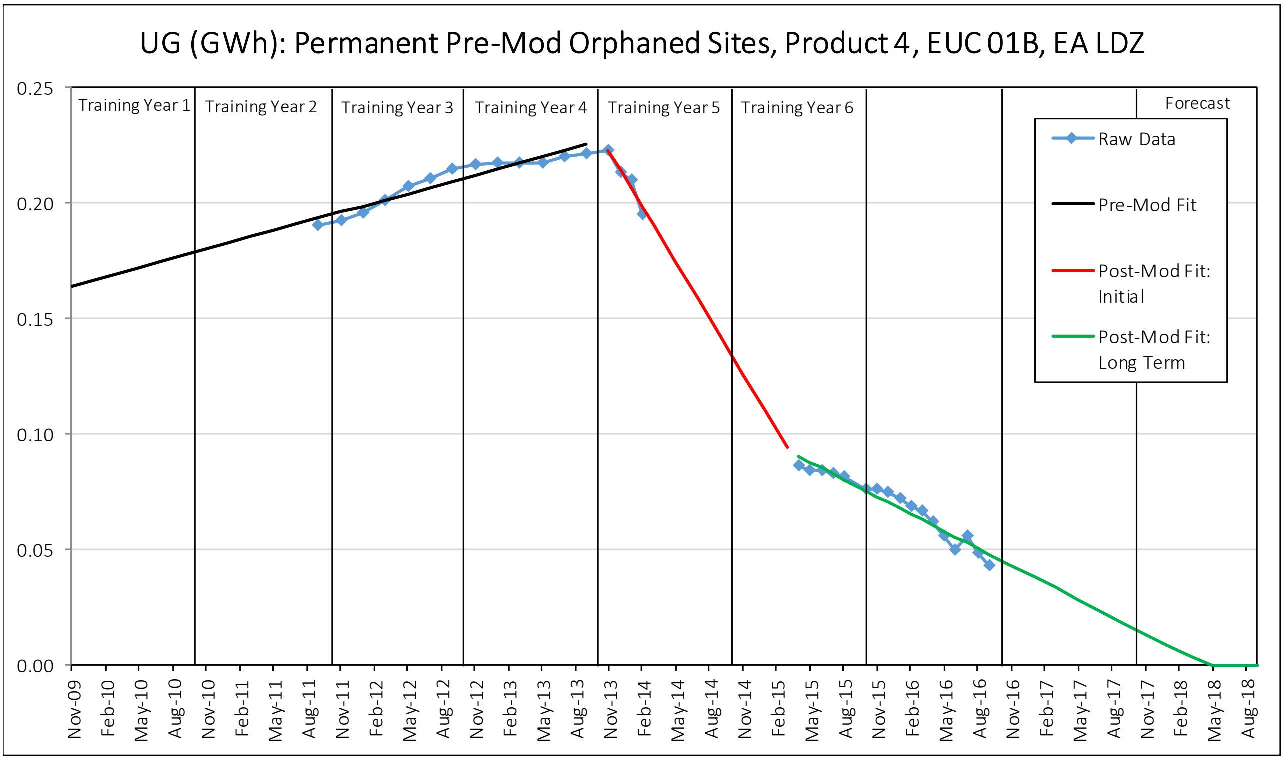

Figure 9: Piecewise Trend for Unregistered UG, Training and Forecast

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 UG 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 UG 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 UG totals (calculated as described in Step 10) to give the best estimate of the amount of Shipperless UG 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 UG 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 UG category in question.

16. The resultant values are now split by Product Class using the rules defined in Section 4.3.3 above. For all Shipperless and Unregistered UG 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 UG estimates for all Unregistered and Shipperless UG categories apart from “Awaiting GSR Visit” to be calculated. At the time of calculation these are split by:

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

For use in further UG calculations, the estimates have to 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” UG 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” UG category for this year.

Therefore, in order to estimate the UG 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 UG 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 UG 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 UG being permanent for 2012/13 and 25% permanent for 2013/14.

- For the SSrP element, all UG 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 UG 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 UG element by the "% of Year with Meter" factor. This gives the total (permanent plus Temporary) UG 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 UG total from this source. The remainder of the UG 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 UG estimate for this category, split by permanent/temporary, Year, EUC and Product Class.

The UG 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

UG 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's systems, and the CDSP is notified of such Unregistered sites on them.

For both these sources of UG from CSEPs, in the pre-Nexus market the UG from LSP sites was backbilled and therefore temporary, whilst UG 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 UG from this source is temporary under Nexus.

It is necessary to calculate both the permanent and temporary elements of IGT CSEPs UG for the training period, whilst only permanent is required for the forecast year. This allows temporary UG to be removed from the raw UG total (permanent plus temporary) calculated using the Consumption Method for the training period, leaving the final total as permanent UG only. Any permanent UG can then be extrapolated to the forecast year to facilitate the calculation of the UIG factors.

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 UG 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 has 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 standard rules defined in Section 0 and are used directly in the UG figures: these represent the estimated annual contribution to UG 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. The snapshot tables are in the format 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	3,843

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.

19. 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.

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

9. The above process gives, for each point in time, an estimate of what annual UG 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 UG is being forecast. An example of this trend across a number of snapshots is shown in Figure 10 below.

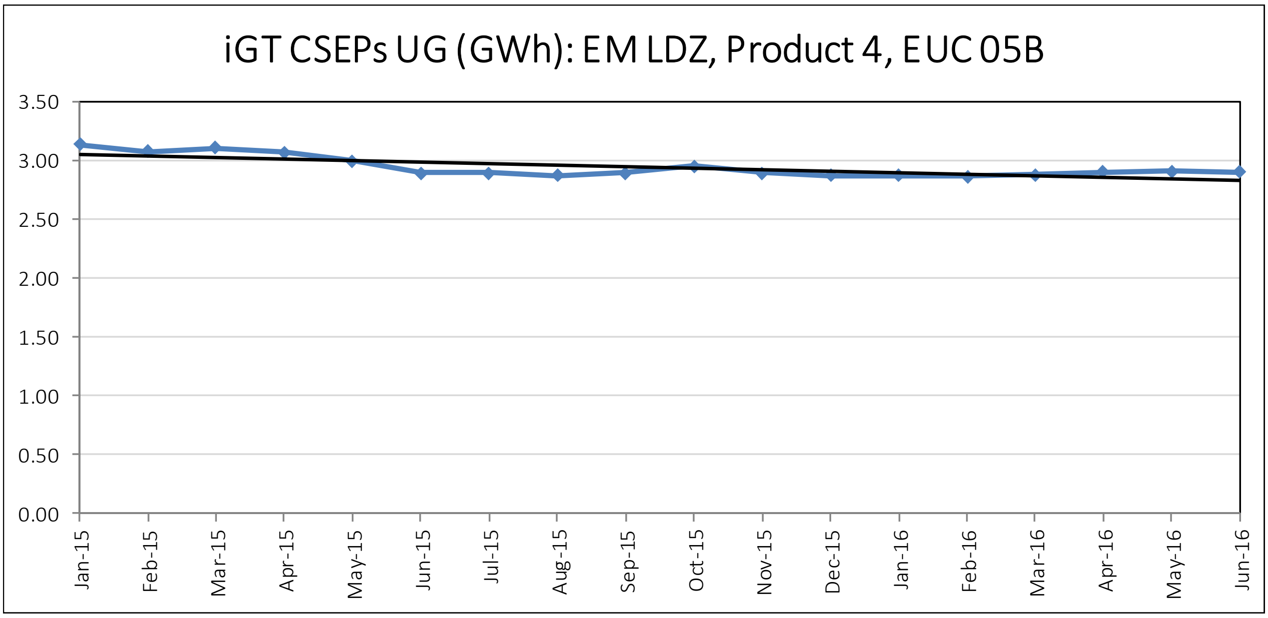

Figure 10: IGT CSEPs UG by Snapshot

10. The above process leads to the creation of a total of 468 individual trend lines for UG 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 UG 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 UG forecast year and the historic UG 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 UG and the Balancing Factor, which are based on data from the training years (currently 2011/12 to 2015/16).

11. For the pre-Nexus period (i.e. the training period), the UG calculated in this way from LSP sites is temporary. The UG calculated in this way from SSP sites is permanent. Under Nexus, the CDSP have confirmed that all UG 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 UG is temporary and does not form part of the permanent UG 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 UG 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 UG 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 UG 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 11 below.

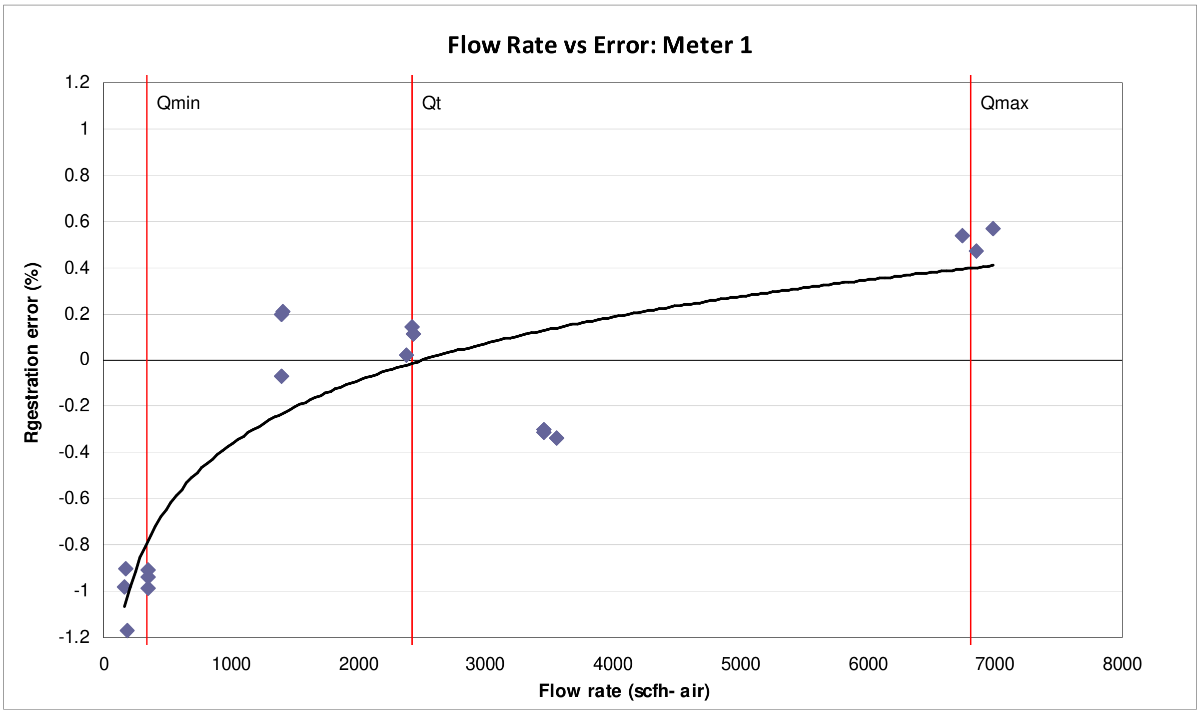

Figure 11: Typical Calibration Curve for an RPD Meter

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 UG 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 UG energy values are applied to all of the training years and the forecast year.

7.6 Detected Theft

Detected theft which occurs within the window allowing back-billing (3-4 years) is a temporary source of UG and it is therefore quantified and subtracted from the total UG for each part of the training period based on year of occurrence. For all previous AUG years, the theft calculations were based entirely on data provided by the CDSP. For the current analysis additional data, consisting of 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 UG 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 data 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 12 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's data.

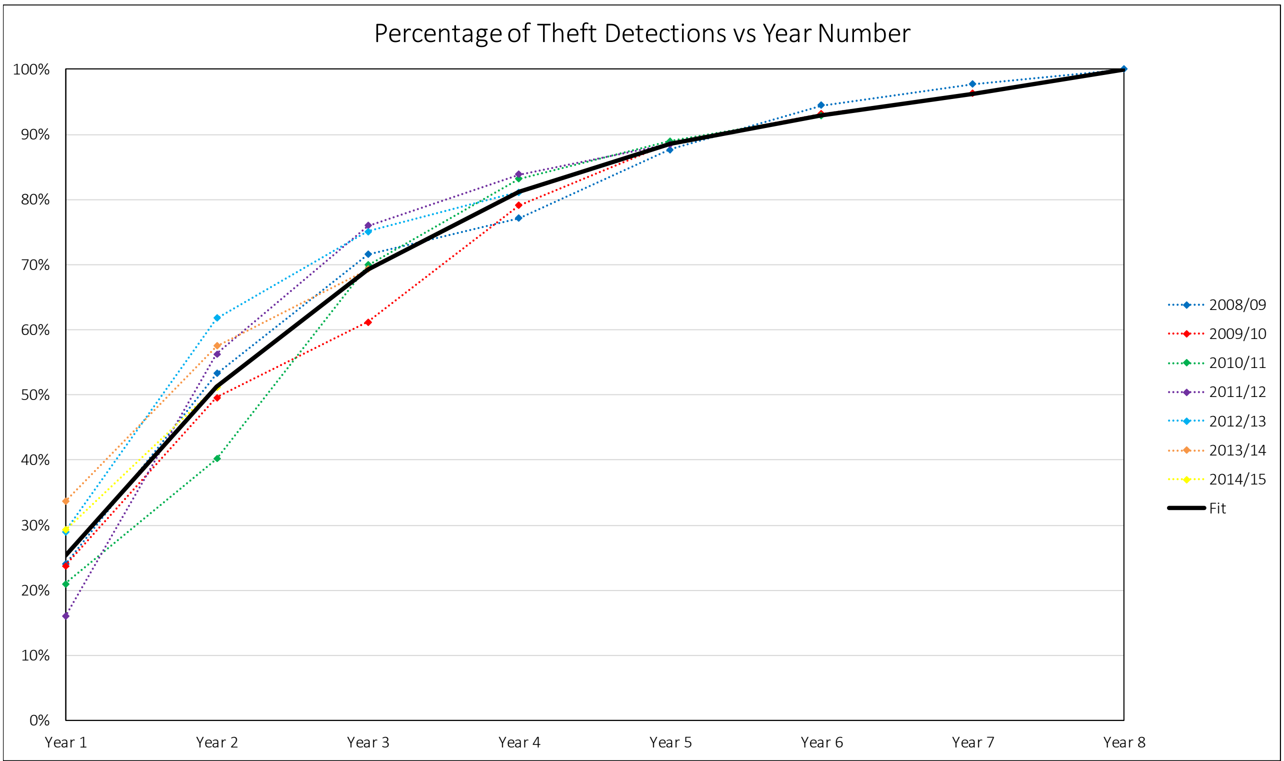

Figure 12: Percentage of Theft Detections by Year Number

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 UG, whilst the remainder are permanent.

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

7.7 CSEP Shrinkage

The method for estimating CSEP Shrinkage is new for the 2018/19 analysis and as such the step-by-step method is described in Section 5 above, where new analysis areas are covered.

In future years, the step-by-step guide will appear in this section.

7.8 Balancing Factor

All of the analysis described above allows the total permanent UG to be calculated for each year of the training period. The permanent elements of the directly calculated UG categories (IGT CSEPs, Shipperless/Unregistered, consumer meter errors and Shrinkage issues) 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 7.9 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 UG. In order to do this, rules must be applied.

It is known that the Balancing Factor is composed mainly of undetected theft, and this provides the basis for how it is split. The split itself is a two-stage process, which is based on the following principles:

1. Create an initial split by aggregate AQ for each EUC/Product Class, with no reference to meter type (traditional, Smart, AMR). It is therefore assumed that all types of sites and meters experience similar rates of theft per unit AQ (i.e. they are equally easy to steal from) at this point.
2. Adjust the above EUC/Product Class split to recognise the fact that not all meters are equally easy to steal from and not all types of site are likely to steal gas.

In November 2016 the AUG Expert approached the industry for information regarding the relative difficulty of stealing from different types of meter (traditional, AMR, Smart) and whether the additional information received from AMR and Smart Meters allowed theft to be more readily detected should it occur on these meters. A number of replies were received, and these have fed into the following additional rules for splitting undetected theft between EUCs and Product Classes. Note that Electralink, were also approached but they were unable to supply us with any further information regarding the split of theft between meter types over and above the SPAA Schedule 33 data. The Balancing Factor split logic is as follows:

1. Theft is still possible and still takes place from all types of meter. If a meter is bypassed, for example, it does not matter what type of meter it is – an AMR or Smart Meter will not prevent this theft.
2. In theory, Smart Meters provide information that could be used to detect theft more readily e.g. tamper alerts. It is, however, early in the Smart Meter rollout period and so little evidence of this phenomenon in practice has yet been observed. All shippers who responded to the AUG Expert’s request for information agreed that the more detailed data that Smart Meters supply would allow thefts to be detected more quickly, but they reported that the issue of whether a greater proportion of thefts could actually be detected was less clear. Once again, this is due to it being early in the Smart Meter rollout period, which means there is little data available regarding this issue.
3. Certain common types of theft are impossible on Smart Meters (e.g. tampering with the mechanical index) and hence the level of theft from Smart Meters as compared to traditional meters will be

reduced due to this effect. Based on shipper responses to the AUG Expert, it is estimated that mechanical index tampering accounts for around 30% of thefts from traditional meters. It is acknowledged that in some cases a customer wishing to steal will choose an alternative method of doing so if they have a Smart Meter, but in others the unavailability of this very common theft method will prevent the theft taking place.

4. Out of over 15,000 detected theft records covering an 8-year period, none were from sites in EUC 09B/DM. It is concluded that the greater scrutiny on sites of this magnitude prevents theft occurring in these cases. Thefts from all other EUCs have been observed, and hence it is concluded that undetected thefts from these sites are possible.

Based on the above, when EUC 09B has been removed from the calculations, upper and lower limits for the true split of undetected theft can be defined. These are both based on a split of population into EUC/Product categories by throughput (defined as the aggregate AQ for each). The true split of undetected theft (and hence the Balancing Factor) must lie between these limits, and the rules for defining them are as follows:

- Upper Limit (i.e. highest possible Smart Meter theft)
  Smart Meters and AMR meters are just as easy to steal from as traditional meters and data from them does not allow a greater proportion of thefts to be detected. Undetected theft is therefore split across EUCs 01B-08B by throughput with no reference to meter type.

- Lower Limit (i.e. lowest possible Smart Meter theft)
  Smart Meters and AMR meters are either impossible to steal from, or if theft did take place the data from them guarantees detection. Undetected theft is therefore split across EUCs 01B-08B by throughput with Smart Meters and AMR removed (zero undetected theft from these sources).

The actual split will lie between these two limits, but due to the limited data available it is currently impossible to carry out a rigorous analysis to place this precisely. As stated above, however, the following information is available:

- 29% of thefts are by index tamper, and these cannot take place on a Smart Meter. This provides an initial placement of the best estimate at 71% of the range between the lower and upper bound.

- Some of these will be replaced by an alternative theft method, whilst others will not. This raises the estimate above 71%.

- The meter replacement process requires anyone stealing gas to successfully remove and conceal the theft before starting it again on the new meter if the theft is to continue. This will not always happen due to either detection of the theft or unwillingness of the site owner to start a new theft. This will move the best estimate back down towards, and potentially past, the 71% point.

- As noted above, it was universally acknowledged by all respondents to our information request that data from Smart Meters will help thefts to be detected more quickly. This reduces the value of each theft, and as the UG analysis deals in GWh stolen and not the number of thefts, this reduces the best estimate further.

- The government Department for Business, Energy & Industrial Strategy have estimated the relative levels of theft from Smart Meters and traditional meters as part of their Smart Meter Roll-Out Cost-Benefit Analysis [18]. In this they state that the level of theft from Smart Meters due to quicker detection is likely to be 20-33% lower than that from traditional meters (although they purposely use a conservative figure of 10% in their cost benefit calculations).

Therefore, at this stage of the analysis it is reasonable to place the split at the mid-point of the two limits, reflecting a situation where the undetected theft rate from Smart Meters and AMR (in terms of undetected kWh of theft per GWh of throughput, for example) is half of that of a traditional meter. This 50% figure is consistent with the initial 71% point from index tampers plus an additional 20-33% from quicker detection, and there is no evidence to favour a different figure over this estimate at this point in time.

The application of this logic results in the following step-by-step process that leads to the final Balancing Factor split:

1. Extrapolate the full meter population (split by EUC and Product Class) to the midpoint of the forecast year (i.e. April 2019). This process uses the actual population data for Jun 2017 and Nov 2017 provided by the CDSP, which can be found in Table 3 and Table 4 in Section 5.15 above. The extrapolated values for Apr 2019 can be found in Table 5 in the same section. The extrapolation is based on the rate of change for each EUC/Product Class category between Jun 2017 and Nov 2017, taken forward to Apr 2019.

2. Calculate the estimated Smart Meter/AMR population for Apr 2019. This is based on:

    a. Estimates of the Domestic Smart Meter population and installation rate (large suppliers only) from the Sep 2017 BEIS Smart Meter Report [34].

    b. Estimates of the Smart Meter population and installation rate for smaller non-domestics (EUCs 02B and 03B, large suppliers only) from the Sep 2017 BEIS Smart Meter Report [34].

    c. Estimates of the Smart Meter population for smaller non-domestics (EUCs 02B and 03B, small suppliers) provided by ICoSS.

    d. An assumption of 100% compliance with the licence condition that all sites in EUCs 04B and above must have AMR.

    The installation rates for the 6 quarters from Q2 2016 to Q3 2017 are used to create trends, which are then used to calculate expected installation rates for each future quarter up to Q1 2019. These estimated installation figures are then used to create the expected Smart Meter and AMR populations for Apr 2019. Actual and forecast installation rates for domestics are shown in Figure 13 below.

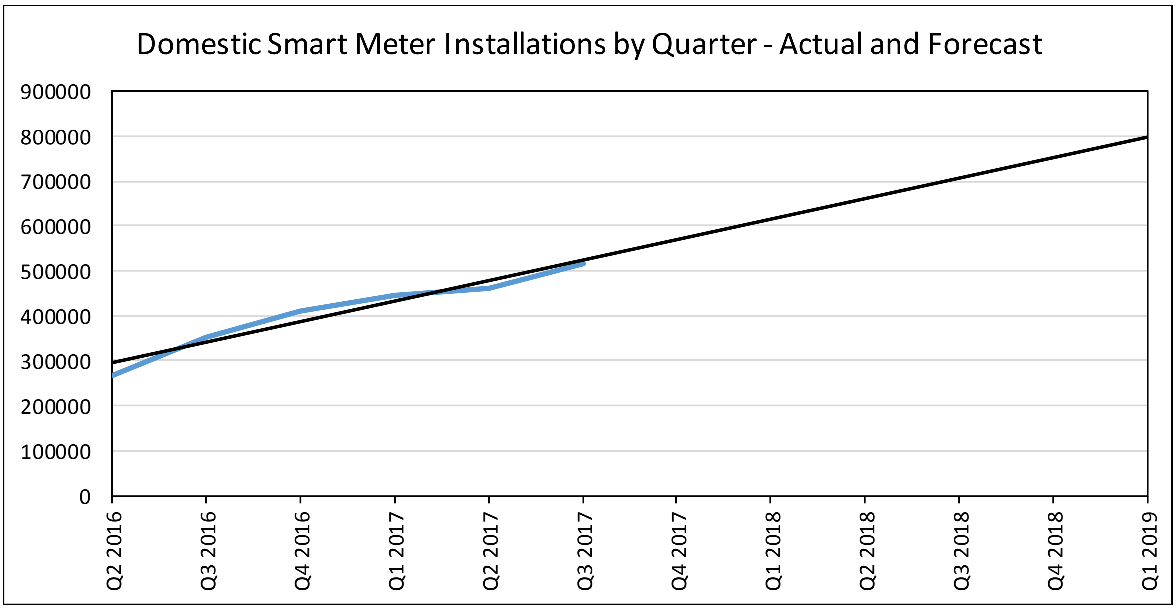

Figure 13: Actual and Forecast Smart Meter Installations, Domestic

The full calculations for both domestics and non-domestics can be found in the spreadsheet "Smart Meter Population.xlsx" that forms part of the data pack that accompanies this document. The output from this process is the expected completion percentage of the Smart Meter installation program as of April 2019, and the figures for this are as follows:

Domestic: 36.5%

Non-Domestic: 27.8%

3. Use the figures and assumptions described above to calculate the expected Smart Meter/AMR population for each EUC/Product Class category as of April 2019. This is shown in **Table 9** below. All other sites are assumed to have traditional meters.

	01B	02B	03B	04B	05B	06B	07B	08B	09B	Total
Product 1	0	0	0	0	0	0	0	0	299	299
Product 2	17	46	22	44	78	227	211	246	3	894
Product 3	59,111	3,329	1,686	92	1	0	0	0	3	64,222
Product 4	8,577,040	50,291	11,044	19,356	4,818	1,477	504	187	0	8,664,717
Total	8,636,168	53,666	12,752	19,492	4,897	1,704	715	433	305	8,730,132

Table 9: Estimated Smart Meter and AMR Population, April 2019

4. Calculate the Balancing Factor low limit. This is based on an assumption that all meters are equally susceptible to theft and hence represents a population split by AQ (with no consideration of meter type), with EUC 09B removed. The Balancing Factor low limit is shown in **Table 10** 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%	0.01%	0.07%	0.52%	0.96%	2.33%	0.00%
Product 3	0.20%	0.12%	0.17%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%
Product 4	69.78%	5.76%	4.43%	5.15%	3.63%	2.95%	2.25%	1.64%	0.00%

Table 10: Balancing Factor Low Limit

5. Calculate the Balancing Factor high limit. This is based on an assumption that there is no theft from Smart Meters or AMR and hence represents a population split by AQ with all Smart Meters and AMR removed (i.e. traditional meters only). The Balancing Factor high limit is 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	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Product 4	85.47%	8.14%	6.39%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table 11: Balancing Factor High Limit

6. The final split by EUC and Product Class for the Balancing Factor is the midpoint of these two extremes, and is shown in Table 12 below. These percentages are applied to the Balancing Factor total, and the resultant figures form a part of the calculation of the final factors (which are given in Section 8).

	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%	0.01%	0.04%	0.26%	0.48%	1.16%	0.00%
Product 3	0.10%	0.06%	0.09%	0.01%	0.00%	0.00%	0.00%	0.00%	0.00%
Product 4	77.63%	6.95%	5.41%	2.57%	1.81%	1.48%	1.12%	0.82%	0.00%

Table 12: Balancing Factor Split

7.9 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 snapshots available until mid or late 2017.

The analysis 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 UIG factors that form the final output are calculated on a national basis, however, and so figures are aggregated across all LDZs once they have been calculated in order to achieve this.

The basis for the factors is always Final UIG (i.e. UIG at line in the sand, with all reconciliations up to this point applied) divided by estimated throughput, for each EUC/Product Class. Final UIG is unknown, however, and is hence estimated using the total permanent Unidentified Gas, which is calculated as described in this document.

The first step of this process is to calculate the total Unidentified Gas (permanent plus temporary) for the training period using the Consumption Method, which is described in detail in Section 7.2. The permanent UG total for each year is then obtained by subtracting temporary UG from it. For the forecast year, total permanent UG is obtained by extrapolating each category forward individually to the relevant year and summing the resultant values (both the Balancing Factor and the directly calculated elements). This gives the permanent Unidentified Gas total for the forecast year for each EUC/Product Class, which is the best estimate of Final UIG. These energy values are then converted into factors using the method described in Section 7.10 below.

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

7.9.1 Unregistered and Shipperless

The principles of extrapolating this element of UG 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 UG. The implementation of these Mods during the training period results in the UG pattern changing at specific points (e.g. the implementation dates). As many individual time segments as necessary are used to accurately model the UG 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" UG, all UG from this source is temporary from 2014/15 onwards.
- For the SSrP element of "Awaiting GSR Visit" UG, following Nexus go-live, any site without a record of the installation date of the new meter will contribute permanent UG 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 UG element by the "% of Year with Meter" factor. This gives the total (permanent plus temporary) UG 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 UG total from this source. The remainder of the UG for the Shipperless SSrP element of "Awaiting GSR Visit" is temporary.

7.9.2 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 UG category post-Nexus and the trends are only used for the training period. This is due to the fact that only permanent UG is calculated for the forecast period and under Nexus all IGT CSEPS UG is temporary.

For the training period, only UG 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.

7.9.3 Consumer Metering Errors

This category of UG is typically small. An assumption is made that the process is steady, and therefore single estimates of the figures for each LDZ/EUC/Product Class combination are applied to each historic year and taken forward to the forecast year.

7.9.4 Detected Theft

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

7.9.5 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, calculated as described in Section 5.18 above, is applied to these.

For the training period, only an aggregate figure for UG 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.

7.9.6 Balancing Factor

The Balancing Factor is calculated individually for each gas year of the training period by subtracting temporary UG plus all directly estimated permanent UG from the total UG 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 3,775 GWh.

7.9.7 Total Permanent Unidentified Gas

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

The forecast of total permanent UG for 2018/19 is 3,837 GWh.

7.10 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 UG from it.

They are therefore calculated using the detailed estimates of the Final UIG (for the year in which the factors will be in force) described in Section 7.9 above. Once the Final UIG 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, extrapolated to this point using the method described in Section 5.15 above). This AQ data has been supplied to the AUG Expert by the CDSP. Each factor is calculated as follows:

$$\text{UIG Factor}_{PRODUCT,EUC} = 10 * \text{Final UIG (GWh)}_{PRODUCT,EUC} / \text{Aggregate AQ (TWh)}_{PRODUCT,EUC}$$

The Final UIG and Aggregate AQ have different units, and a multiplier of 10 is also applied. These approaches are used to ensure that the resulting factors give sufficient precision when expressed to 2 decimal places as required.

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 UIG FACTORS

The current draft Unidentified Gas factors, calculated as described in this document, for the 2018/19 gas year are as follows:

Supply Meter Point Classification	Class 1	Class 2	Class 3	Class 4
EUC Band 1	0.22	42.25	43.69	94.01
EUC Band 2	0.22	42.25	44.61	104.08
EUC Band 3	0.22	42.25	42.58	105.90
EUC Band 4	0.22	42.25	42.05	43.38
EUC Band 5	0.22	42.25	43.23	42.11
EUC Band 6	0.22	42.71	43.23	42.34
EUC Band 7	0.22	31.98	43.23	42.02
EUC Band 8	0.22	4.56	31.13	41.94
EUC Band 9	0.22	0.22	0.22	0.22

Table 13: Draft AUG Table

Whilst the general pattern of these factors remains similar to that from the 2017/18 analysis, there are a number of differences as described below. The largest issue driving these changes is the availability of actual Product Class data for this year's analysis. For the previous year, pre-Nexus asset data was used, with a number of rules applied to map this information to post-Nexus Product Classes. This process necessarily relied on a number of assumptions, not all of which proved to be accurate – in particular the take-up of Product Classes 2 and 3 has been much lower than anticipated.

The application of these rules for 2017/18 resulted in factors that were generally similar for Product Classes 2, 3 and 4, with the only major difference being higher factors for EUC bands 01B-03B in Product Class 4 due to the presence of a large number of traditional meters in these categories. This aside, the factors for each of these classes remained at a relatively consistent level up to EUC 06B, followed by a gradual drop-off to the minimum value taken for EUC 09B.

Whilst the Smart Meter installation programme continues, a large number of traditional meters remain in Product Class 4 EUCs 01B-03B and so the higher factors for these categories continue. The use of actual Product Class data has now allowed the genuine differences between Product Classes 2, 3 and 4 to be highlighted. The drop-off in factor values with increasing EUC is still present for both Product Classes 2 and 3, although the two show different rates of change, with Product Class 2 dropping more quickly. The pattern of Product Class 3 is the closest to what was observed in the previous year.

Product Class 4 shows no drop-off at all, however, with the factors remaining consistent right up to EUC 08B, and only dropping down for 09B. This shows that when meter readings are periodic, the site magnitude does not affect the relative amount of Unidentified Gas assigned (i.e. assigned UIG remains proportional to the AQ regardless of EUC).

The final point to note is that over and above these observations about the distribution of the factors, their overall magnitude has also dropped. This reflects an ongoing fall in the expected total permanent Unidentified Gas for the forecast year, which has fallen from 4113GWh for 2017/18 to 3837GWh for 2018/19. Whilst the changes in the pattern of the drop-off at high EUCs results in a few cases where the 2018/19 factor is higher for a particular category, in general the new factors are lower than their equivalents from the previous year.

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 and 2018/19 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 Eon Response - DNVGL Comments" [22]
- "AUGS 2018-19 British Gas - DNVGL Comments" [24]
- "AUGS 2018-19 ICOSS Response - DNVGL Comments" [25]
- "AUGS 2018-19 Corona Response - DNVGL Comments" [26]
- "AUGS 2018-19 Eon Response - DNVGL Comments" [27]

Note that all responses contained in these documents relate to the UG 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 UG 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

Table 14: Responses to the First Draft of the 2017/18 AUG Statement

Organisation Name	Date of Communication
British Gas	14/03/2018
ICoSS	14/03/2018
Corona	14/03/2018
E.ON	09/03/2018

Table 15: Responses to the First Draft of the 2018/19 AUG Statement

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 Meter Roll-Out Cost-Benefit Analysis, August 2016

[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 Eon Response - DNVGL Comments, April 2017

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

[24] AUGS 2018-19 British Gas - DNVGL Comments, April 2018

[25] AUGS 2018-19 ICOSS Response - DNVGL Comments, April 2018

[26] AUGS 2018-19 Corona Response - DNVGL Comments, April 2018

[27] AUGS 2018-19 Eon Response - DNVGL Comments, April 2018

[28] Code of Practice for Gas Meters Asset Managers, V1.0, September 2005

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

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

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

[32] Gas Energy Measurement in Consumer Billing, DLC/0059, D.F.Lander, 23 July 2014

[33] Unidentified Gas Summary 2018_19.xlsx

[34] Department for Business, Energy & Industrial Strategy: Smart Meters Quarterly Report to end September 2017, 30 November 2017

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.
CDSP	Central Data Services Provider. This role is carried out by Xoserve
CF	Correction Factor. Used in calculation of energy from volumetric meter reads. Corrects measured volume to equivalent volume under standard temperature and pressure conditions
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.
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
EUC	End User Category

EWCF	Estimated Weather Correction Factor (deeming algorithm output - alternative to WCF based on CWV rather than demand)
Final UIG	The value of UIG which would be calculated using the data available at 'line in the sand' i.e. all reconciliations have been applied. This will contain only permanent elements of unidentified gas
Found Meter	A meter being supplied by a Shipper but for which the CDSP have no record
GBNA	Graphical Based Network Analysis, software used by Cadent to perform computer simulation of networks
GFD	Gas Flow Day
GSR	Gas Safety Regulations
IGT	Independent Gas Transporter
Initial UIG	This is the same as UIG as defined in UNC
Intermediate UIG	The value of UIG which would be calculated using the data available at some intermediate point in time between initial allocation and the 'line in the sand'. i.e. some, but not all reconciliations have been applied
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
MAMCoP	Meter Asset Manager Code of Practice
Model Error	The statistical error associated with any modelling or estimation process. It is an inherent part of any statistical model and does not imply that the model itself is inadequate or incorrect.
MPRN	Meter Point Reference Number
NDM	Non-Daily Metered
ODR	OFGEM Data Request
OUG	Own Use Gas
PNID	Project Nexus Implementation Date
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
SPAA	Supply Point Administration Agreement
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
UG	Unidentified Gas
UIG	UNC definition (Section H 2.6.1) of a daily estimate of Unidentified Gas calculated 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
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

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
METER_CAPACITY	Number
READ_FACTOR	Number
METER_MECHANISM	Char[30]

MMSPS

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

Name	**Description**
MPR_ID	Number - Unique ID for meter used across ALL data
SUPPLY_POINT_ID	Number - Unique dummy ID for supply point
LDZ	Char[2] - LDZ identifier e.g. EA
CONF_EFFECTIVE_DATE	Date
CONF_END_DATE	Date

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.

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.