GL Noble Denton

| Report Number: 14128 | 11th September 2013 |
|---|---|

# 2013 Allocation of Unidentified Gas Statement for 2014/15

| Not Restricted | GL Noble Denton |
|---|---|

**Prepared for:**

Uniform Network Code Committee

**Prepared by:**

Andy Gordon, Clive Whitehand, Tony Perchard, Mark Lingham

GL Industrial Services UK Ltd trading as GL Noble Denton

Holywell Park
Ashby Road
Loughborough Leicestershire
LE11 3GR
United Kingdom

Tel: +44 (0)1509 28 2000
Fax: +44 (0)1509 28 2141

E-mail: AUGE@gl-group.com

Website: www.gl-nobledenton.com

Version 3

**Customer Reference:** 1-6AS9AC

| Version | Changed against | Date | Status<br>(for consultation<br>For approval<br>Approved) |
|---|---|---|---|
| 1.0 | Created | 1st May 2013 | For consultation |
| 2.0 | Updated version 1.0 | 26th July 2013 | For consultation |
| 3.0 | Version 2.0 | 11th September 2013 | Approved |
| | | | |

# Executive Summary

This document contains details of the methods developed by the Allocation of Unidentified Gas Expert (AUGE) for estimating the overall level of Unidentified Gas (UG) and allocating it between market sectors, the data requested to support this analysis, and the data received following such requests. Full estimates of the total energy value of UG split by LDZ and source will be provided once the methodology described in this document has been approved by the Uniform Network Code Committee (UNCC).

In addition to the above, this document describes how the AUGE has followed the published guidelines.

This document is the 2nd Draft 2013 AUGS for 2014/15. The document describes analyses undertaken in 2013 to improve the estimate of UG and the outcome of the investigations into a number of issues arising from the consultation of the 2nd Draft 2012 AUGS for 2013/14 and the 1st Draft 2013 AUGS for 2014/15.

Section 3 of this document provides a high level overview of the methodology in general terms.

Section 4 describes the analyses carried out this year and conclusions reached. It does not describe the resulting methodology as that is covered separately. The following key topics and conclusions are summarised below:

1. Update and further examples regarding the handling of Multiple Meter Supply Points
   Having obtained information for Multiple Meter Supply Points we confirmed that meters failing the consumption calculation have been scaled correctly where they were part of a MMSP and that there were no exceptions that required different treatment. However, there are situations where a meter (not necessarily an MMSP) fails validation due to the selection of AQ where the AQ has been revised to 1. This has been investigated and addressed.

2. Review of the potential use of LSP meter reads to calculate consumption when LSP metered consumptions are invalid using a large sample of LDZs
   The comparison of the use of LSP meter reads to calculate consumption when LSP metered consumptions are invalid for multiple LDZs has not been completed for this draft AUGS. Updated meter reads, consumptions and meter asset data have been provided as described later in the statement. However, a further refresh of meter asset data is required and it has not been possible to obtain and re-run consumption calculations in time for publication in this draft (see section 4.2 for further details). This would potentially affect only 2% of the LSP consumption calculations and at this time we conclude that the most appropriate data to use for LSPs is the metered consumptions and if this fails, consumptions will be estimated using the EUC average.

3. Revision of the consumption calculations to improve handling of Read Units
   The consumption method has been updated to improve the handling of certain subsets of calculations where read unit information was not being used fully. This addresses two of the issues identified by ICoSS during consultation. This statement includes details of the improved consumption methodology.

4. Assessment of the impact of vacant sites on consumption calculations
   An assessment of vacant sites has been carried out, but there is insufficient information to determine whether a site is really vacant or not. Mod 0282 would have improved the situation but it was rejected. Until additional information becomes available such sites will be treated in the same way as we treat the general population.

5. Seasonal normal UG
   We concluded that using a seasonal normal estimate of UG protects the resulting AUG table from bias should there be periods of unseasonal cold or warm weather.

6. Temperature and Pressure (T&P) Conversion Factors
   In general, the methodology makes limited use of T&P factors although the metered consumptions provided by code parties do involve the use of T&P factors to calculate them. Where T&P factors are recorded as zero we use the default factor (1.02264). Further details are provided later in this statement.

7. Assessment of an alternative theft split method proposed by ICoSS
   An assessment of the theft split method proposed by ICoSS and reported in the 1st draft 2013 AUGS for 2014/15 identified a number of flaws in the proposed solution. Furthermore, the method relies on data sources which we have previously highlighted as not robust enough for deriving a theft split method and uses a small sample of detected thefts. In particular, we conclude that theft periods, estimated volumes and AQs used to determine sector classification are unreliable and any conclusions drawn from these data sets should be treated with caution. We have concluded that the fairest method of splitting theft is by throughput.

8. Assessment of the impact of changes to Supplier Licenses to increase theft detection rates
   We have considered the incentives introduced by OFGEM with the aim of increasing theft detection levels from 3000 to 6000 per annum. To date there is no evidence to suggest that theft detection levels have increased and without a back correction facility we have concluded that detected theft deducted from the total UG will be based on the average of previous years by year of occurrence.

9. Derivation of a Balancing Factor split formula that can be used year on year
   We have identified a number of improvements to the theft split throughput method proposed last year. This includes using our consumption estimates rather than the ODR report data, applying a seasonal adjustment to the consumptions and using a rolling average instead of a forward extrapolation. This results in a much more transparent, simple and consistent method going forward.

10. Metered volume corrections
    We investigated the impact of meter consumption corrections that occur over time with a view to estimating the amount of corrections that would be expected year on year. However, as there has been a large step change in the consumption data provided for this year (since it includes year on year changes plus additional updates not provided in 2012) we cannot assess this for this years statement. We recommend this is assessed at the next consumption data drop in 2014.

11. Update to the unregistered site methodology to handle bias in the initial AQ estimates
    Initial AQ estimates for new sites that may be unregistered or shipperless are not validated. Analysis of these identified that subsequent AQs turned out to be lower than the initial estimates. We concluded that conversion factors should be applied to the initial estimates from the snapshot reports to ensure the correct level of unregistered/shipperless site consumption was obtained.

12. CSEP Shrinkage
    We concluded that this was an issue for the Shrinkage forum.

13. Improvements to the handling of Prime/Sub meters
    We have assessed the potential impact of double counting consumption for Prime/Sub meters configurations and for meters that have LDZ re-assigned. The methodology has been updated to deal with these configurations and prevent double accounting.

14. Improvements to the handling of meters that have LDZ re-assignments
    We have assessed the potential impact of double counting consumption for meters that have LDZ re-assignments. The methodology has been updated to deal with this.

15. Improved selection of AQ for validation of consumptions
    To ensure the most accurate AQS were used to validate consumptions, the methodology used updated AQs if they had been revised part way through the year. However, where the revision set the AQ=1 perfectly valid consumptions for previous years were being rejected and replaced by the EUC 01B average this resulted in erroneous EUC band assignments. The methodology has been updated to use the AQ at the end of the formula year to prevent this.

16. Improved estimation of CSEP consumptions
    As part of our validation of the methodology we identified that the best method of estimating CSEP consumption is to use aggregate AQs rather than EUC average consumptions.

17. Use of Cosmetic and real meter exchange data to improve consumption calculations
    Cosmetic and real meter exchange data is being provided by Xoserve to improve the handling of consumption calculations when exchanges occur. Previously any meter changes during a year would automatically result in a calculation failure.

18. Handling of gaps and overlaps in consumption data
    Where there are gaps/overlaps in consumption data the consumption methodology has been amended to ensure the gaps are filled in and overlaps removed.

19. Assessment of the impact of Modifications 0424, 0398 and 0429
    We have assessed the impacts of Modifications 0424, 0398 and 0429 and these are described in this AUGS. Mod 0424 changes the content of the UG associated with Shipperless Passed to Shipper (PTS) sites and information will be provided by Xoserve in order to identify the meters affected going forward. With the information currently available Mod 0398 is not expected to impact the methodology. Mod 0429 does not impact the methodology at this time, however we await a separate modification or change to the Guidelines to adjust the AUGE process to address this.

Section 4 describes the analyses carried out this year (see list of topics above) and conclusions reached. It does not describe the resulting methodology as that is covered separately.

Section 5 describes the data used. Analyses of issues raised from previous consultations have been based mainly on existing data received in 2012. Information pertaining to Multiple Meter Supply Points (MMSPs) and Primes and Sub meters have also been provided. A complete refresh of LSP consumption and meter read data has also been received and used to improve the calculation processes. Additional SSP data up to the end of March 2013 has also been provided. A further update to meter asset data has been requested.

Section 6 describes the resulting methodology proposed for April 2014 – March 2015.

Section 7 will reference the AUG table which will be published at a later date

Section 8 references consultation feedback and responses from previous statements

As explained in the previous AUGS and at the UNCC (AUGE) meeting on 15th May 0213, following the issues encountered at the end of last year's process the AUGE will not provide initial or ballpark estimates of UG either verbally or written until the methodology has been approved and the interim AUG table is published by the Gas Transporters (GTs) on 1st November 2013.

For each area of Unidentified Gas under consideration, the AUGE has provided details of the proposed method of estimating the level of Unidentified Gas from this source, and where necessary, the method of splitting this estimate between Larger Supply Point (LSP) and Smaller Supply Point (SSP) markets.

The methodology estimates the total level of Unidentified Gas, directly calculating its individual component parts where possible, and calculating the aggregate effect of the remaining causes (i.e. those that it is not possible to estimate directly in a robust manner) by subtraction as the Balancing Factor.

We note that the recent OFGEM decision letter regarding Mod 456 suggested that there could be a better separation of methodology from data issues. In an ideal world the description of the methodology could be quite short if the data that it used did not require significant validation and manipulation. However, as documented extensively throughout the AUG process it is not that straightforward as there have been significant data issues to address and since these can result material impacts on the resulting AUG table they need to be handled as part of the methodology so that this is transparent to the industry.

The methodology described in this statement is an improvement on the methods produced in 2011, 2012 and the method described in the 1st Draft 2013 AUGS for 2014/15. All the improvements described in this statement will result in a more accurate estimate of UG and whilst we do not wish to speculate on the magnitude of the impacts of these improvements some of them have the effect of increasing total UG and some have the effect of decreasing total UG. The results are still dependent on source data accuracy and any initiatives to improve data quality within the industry will improve the estimate of UG.

# Contents

# List of Figures



# List of Tables

# 1 Introduction

## 1.1 Background

The Great Britain gas industry can be segmented into two market sectors; Larger Supply Points (LSP) and Smaller Supply Points (SSP). These sectors are defined by the Annual Quantity (AQ) of gas offtaken from the system in a year. Larger Supply Points have an AQ of 73,201 kWh and above, Smaller Supply Points have an AQ of up to 73,200 kWh. Many processes within the gas industry differ between these two sectors.

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).

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 allocated entirely to the SSP market sector (an interim amount was allocated for 2011/12). Through the approval of Modification 229 (implemented in UNC section E 10 – Mechanism for Correct Apportionment of Unidentified Gas [5]) and the appointment of an Allocation of Unidentified Gas Expert (AUGE) a methodology has been defined to ensure that UG can be estimated and charged equitably to the relevant gas sectors.

Under the current Uniform Network Code (UNC) charges are made to Shippers for the volume of gas transported, which include commodity and energy charges. For LSPs the actual value charged (via the combination of Commodity, Energy Balancing and Reconciliation invoicing) is determined by the volume of gas transported as measured by the metering equipment. For SSPs, the commodity charge is ultimately derived (via NDM Allocation and Reconciliation by Difference) by calculating the difference between the volumes of gas measured coming in to the network and the volume of gas measured by the LSPs. Each Shipper with an SSP portfolio is charged a proportion of the total SSP market in proportion to their Annual Quantity (AQ) value against the total SSP market AQ. This calculation of SSP load by subtraction leads to all lost gas being assigned to this market sector.

There had been several UNC modification proposals intended to resolve this issue (Mod 194 [2], 194a [3], 228 [4], 228a [4]), none of which have been accepted by the industry. A further modification, Mod 229 [5] provided for the appointment of the AUGE with responsibility for determining the value of UG so that relevant quantities could be allocated to the correct market sectors.

GL Noble Denton was appointed to the role of AUGE in 2011 and has developed a methodology to apportion UG fairly across both the LSP and SSP market sectors.

The initial methodology used RbD and AQ Bias to estimate UG. Items of UG that could be estimated and apportioned directly were deducted to leave a balancing factor which was split into market sectors based on a theft split. The AUG table was composed of directly estimated UG (e.g. Shipperless, Unregistered sites etc) plus a portion of the balancing factor for that market sector.

In 2012 consumption data was requested for all customers. Data was provided for customers where meter reads/consumptions have been provided to Xoserve (excluding CSEPS which use AQ information) and used to estimate total UG by subtracting the total consumption from the gas input. A methodology was produced based on this data. However, it was not possible to complete the AUGE process in time for implementation.

The methodology for 2014/15 is based very much on the 2012 method with improvements to key areas following the consultation on the previous 2nd Draft 2012 AUGS for 2013/14 and the more recent 1st Draft 2013 AUGS for 2014/15. This document describes the analysis undertaken and the resulting methodology for 2014/15.

## 1.2 High Level Objectives

The AUGE's high level objectives are:

- To determine what data is required from industry bodies to evaluate UG
- To develop and update the methodology of calculating UG
- To publish the methodology in the AUGS (this document)
- To consult with the industry bodies and respond to questions / issues raised
- To prepare an AUG table containing UG totals and rates

## 1.3 Scope

This document contains the following:

- A detailed description of the proposed methodology
- Description of areas of the methodology that are being developed further and the proposed approach to these as appropriate
- Summary of data requested, received and used, and associated assumptions
- Questions raised by the industry bodies during consultations and responses as appropriate (this is provided as a separate document)
- Details of the databases used to hold information associated with UG and used to develop the methodology
- Details of the analyses carried out in 2013 in preparation of the methodology

The final AUGS Table and financial estimates will be included in a future version of this document once the methodology has been approved.

## 1.4 Out of Scope

The AUGE is not concerned with issues regarding the deeming algorithm or the RbD mechanism.

The AUGE is not concerned with resolution of fundamental gas industry business process issues.

The AUGE process is not an opportunity to deal with/investigate issues within the gas industry that should be addressed by other workgroups (e.g. Shrinkage Forum.)

The AUGE is not concerned with transportation charges associated with UG

## 1.5 Document Status

This section provides a status summary of the Unidentified Gas methodology as contained in this version of the AUGS. Final estimates of the energy value and financial value of UG have not been made at this stage, and will only be provided when the methods detailed in this AUGS have been approved by the UNCC. Table 1 shows the status of each element of UG:

Table 1: Unidentified Gas Estimate Status

| Unidentified Gas Subject | Data Status | Methodology Status | AUGS Status |
|---|---|---|---|
| Unregistered sites | Updated data provided every two months | Updated method proposed for consultation | Methodology updated and described in full in this document |
| Shipperless sites | Updated data provided every two months | Updated method proposed for consultation in light of Mod 0424 implementation | Methodology updated and described in full in this document |
| IGT CSEPs | Updated data provided every two months | Complete | Methodology described in 2011 AUGS [10] and included in this draft for completeness |
| Shrinkage error | N/A | Complete | Status described in 2011 AUGS [10] |
| Shipper responsible theft | Theft data covering detections to 2012 received.<br>EUC Groups, meter read frequencies and meter reads and metered volumes received.<br>Updated data to end March 2013 pending. | Updated method proposed for consultation | Proposed method described in this document based on improvement to methodology described in 2012 |
| Metering errors (SSP supply point, NDM LSP Supply point, DM supply point, LDZ offtake metering) | Updated data pending | Complete | Methodology described in 2011 AUGS [10] and Section 6.2 of this document for completeness |
| Overall UG estimate: using meter reads / metered volumes | Updated historical consumption data received for LSP meters for all LDZs. New data received for consumptions up to end March 2013 for SSP and LSP meters. Updated Asset Data received and a further update of Asset Data has been requested. | Further updates provided for consultation | Methodology described in this document |

# 2 Compliance to Generic Terms of Reference

This section describes how GL Noble Denton has adhered to the Generic Terms of Reference described in Section 5 of the AUGE Guidelines [1].

**The AUGE will create the AUGS by developing appropriate, detailed methodologies and collecting necessary data.**

The AUGE has developed a detailed methodology for estimating total UG using meter read and consumption data for both LSP and SSP sectors and requested the necessary data to apply this method from Xoserve. Further enhancements to the UG calculation are also described in this document.

**The decision as to the most appropriate methodologies and data will rest solely with the AUGE taking account of any issues raised during the development and compilation of the AUGS.**

The proposed methodology and assessment of what constitutes UG has been decided solely by the AUGE based on information supplied by all parties. Comments raised by shippers relating to the AUGS documents from previous years have been considered and responses issued, as detailed in Section 8 below. All views expressed have been considered, although all final decisions are the AUGE's own.

**The AUGE will determine what data is required from Code Parties in order to ensure appropriate data supports the evaluation of Unidentified Gas.**

The AUGE has assessed what data is required to support the proposed methodology and has requested information from relevant parties. For the 2014/15 analysis, updated data sets have been requested from Xoserve for all items, and the majority of these have currently been received.

**The AUGE will determine what data is available from parties in order to ensure appropriate data supports the evaluation of Unidentified Gas.**

The AUGE has determined data available following discussions with Xoserve, as all of the data required for this analysis is held by them.

**The AUGE 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 gas.**

Questions regarding various elements of UG have been sent to Shippers during the production of the AUGS for previous years. Further communication will take place as and when necessary.

**The AUGE will use the latest data available where appropriate.**

Data for the 2013 method has been requested. This includes updates to LSP consumptions (including corrections from 2008 onwards) and new meter reads for SSPs up to end March 2013. Updates for theft data, CSEPS, Shipperless/Unregistered sites etc have also been requested and will be used to generate the interim AUG table later in the year. Xoserve have set up several processes for producing reports containing new data on a regular basis (for example the two monthly Shipperless/Unregistered site snapshots). These will continue to be supplied to the AUGE to ensure that the latest data is used for each analysis as appropriate.

**Where multiple data sources exist, the AUGE 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. Over time LSP metered volumes may be corrected, but the meter reads are not. Xoserve advised the AUGE to use metered volumes but analysis has shown that these can be erroneous, particularly for non-corrected SSP data. Therefore the decision was taken to use meter reads for SSP and metered volumes for LSP. Details of how these are determined are described in previous versions of the AUGS [10], [19]. A further review of potential use of LSP meter reads when LSP consumption calculations fail pending updated meter asset data.

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

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

**Where the AUGE considers using data collected or derived through the use of sampling techniques, then the AUGE 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. In order to achieve this, a validation process was developed that was designed to maximise the sample size whilst removing any meter points with invalid data. This is described fully in the 2012 AUGS for 2013/14 [19].

**The AUGE will present the AUGS in draft form (the "Draft AUGS"), to Code Parties seeking views and will review all the issues identified submitted in response.**

The AUGE has documented and reviewed all feedback resulting from AUGS from previous years. Section 8 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 AUGE responses contained in separate documents published on the Joint Office of Transporters website.

**The AUGE will consider any query raised by a Code Party with regard to the AUGS or the data derived, and will respond promptly with an explanation of the methodology used.**

Responses were issued to all parties who submitted comments on AUGS from previous years and previous drafts of this AUGS, and these are noted in Section 8. Separate documents provide the detail of all responses [8], [9], [11], [12], [13], [20], [27].

**The AUGE will consider any relevant query that was raised during the creation of the previous AUGS and was identified as requiring a change to the AUGS, but was not incorporated into the immediately previous AUGS.**

All queries have been carefully considered by the AUGE and where appropriate improvements to the UG calculation have been made. The evolution of the UG calculation can be seen in successive versions of the AUGS.

**The AUGE will provide the Draft and Final AUGS to the Gas Transporters for publication.**

This 2nd draft 2013 AUGS for 2014/15 is provided to the GTs for publication on 26th July 2013.

**The AUGE's final determination shall be binding on Shippers except in the event of fraud, material breach, or where The Committee unanimously considers it is so clearly erroneous for it to be inapplicable.**

This guideline has not needed to be applied at this stage.

**The AUGE 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 AUGS.**

On receipt of data, the AUGE stores the data on our secure project storage area with limited access by the consultants working on the project. The AUGE 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 Xoserve so that the anonymity of the consumer is protected.

**The AUGE 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 on our 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 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 6 and/or in previous versions of the AUGS [10], [19].

## 3.1 LDZ Load Components

Daily load (as measured or calculated at the Supply Meter Point) falls into three relevant categories as far as the reconciliation process is concerned. These are:

**1. Daily Metered (DM) Load**

This is by definition metered and known on an ongoing daily basis. However, it is subject to error and data for known errors is used to correct it.

**2. Larger Supply Point Non Daily Metered (LSP NDM) Load**

In the consumption based methodology, historical LSP consumptions and meter reads are provided including corrections that have been applied over time to calculate total LSP consumption.

**3. Smaller Supply Point (SSP) Load**

In the consumption based methodology, historical SSP consumptions and meter reads are provided.

The sum of these three load components does not equal the gas intake into the LDZ due to the presence of two further factors, as follows:

**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 Uniform Network Code (UNC) [6].

**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 CSEP setup and registration delays
- Errors in the Shrinkage estimate
- Shipper-responsible theft
- Meter errors – this includes LDZ offtakes, LSP consumer meters and SSP consumer meters

UG is currently unknown and hence must be estimated.

In addition to the above factors, there may also be a small element of Stock Change, which represents the difference between opening and closing stock on any given gas day. Given that aggregate UG is based on annual rather than daily consumptions, any adjustment due to stock change (which in this case would be the difference in stock between the start of the UG year and the end of the UG year) will be negligible. It has therefore been discounted from calculations.

## 3.2 Unidentified Gas Methodology

The original method created by the AUGE for calculating UG was described in detail in the 2011 AUGS for 2012/13 [10]. This original method was based on RbD values and will be referred to as the RbD based methodology in the remainder of this document. Since this time, major changes have occurred in two areas. These are summarised below but were described in detail in the 2nd draft 2012 AUGS for 2013/14 [19].

1. The estimation of the UG total across all market sectors:
   In the 2011 AUGS for 2012/13 this was estimated based on RbD quantities adjusted for allocation bias (resulting from underlying AQ bias), as this was the most accurate method given the data available at the time. As described in the 2012 AUGS for 2013/14 [19], meter read and consumption data is now available for all supply points (both LSP and SSP) and so an improved method based on these has been developed. The new method is theoretically more accurate than the RbD based methodology. More input data validation is required for the consumption method, but this has the advantage of being able to delve into more detail to find and resolve issues hidden by the RbD method.
2. The market sector split of undetected theft:
   In the 2011 AUGS for 2012/13 this calculation was based on detected theft levels. However, this method is highly dependent on Shipper-supplied estimates of theft (duration and value) and can be influenced by Shipper theft detection strategy. It is also highly dependent on AQs particularly with regards to sector classification. In 2012 detailed analysis of theft data and meter consumptions/reads associated with theft-affected meters indicated that this data was not fit for purpose. An alternative method based on throughput was proposed. In this document a review of a further method proposed by ICoSS has also been carried out.

### 3.2.1 Estimation of Total UG using Meter Reads/Metered Volumes

The overall concept of calculating total UG using metered consumption data is simple. Total UG is estimated by taking the difference between the calculated total NDM demand (i.e. LDZ intake minus shrinkage and DM load) and the sum of metered consumption for all NDM meter points. There are, however, a number of complexities which have been identified that must also be accounted for in the calculation. The total UG is estimated for each LDZ and formula year separately, and an overview of the process is provided in the flowchart shown in Figure 1.

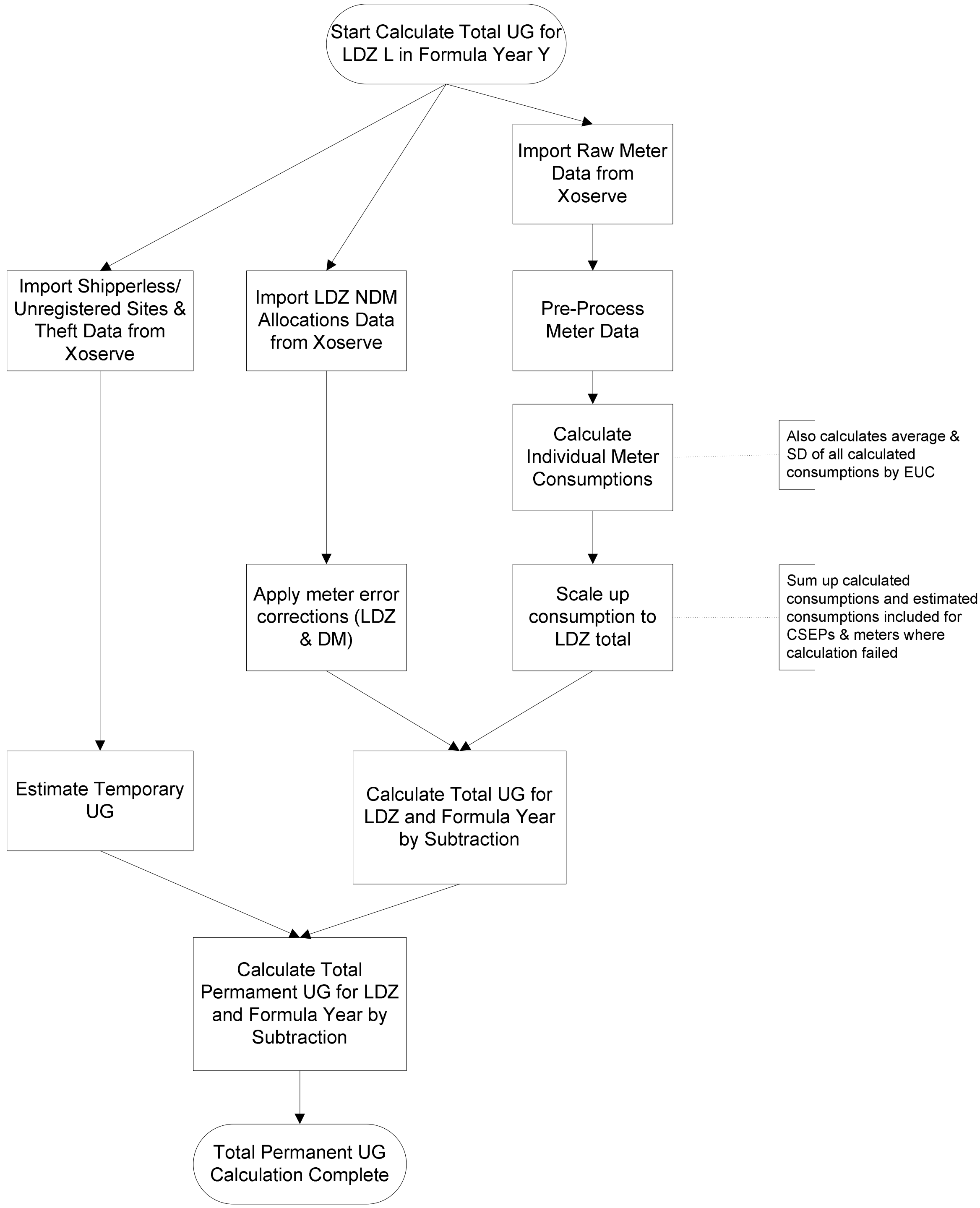


Figure 1: Overview of UG Calculation Methodology

Having obtained the total figure using the consumption methodology, the value of individual components that make up the UG total are calculated where this is possible. The difference between the calculated UG total and the sum of the directly estimated components is referred to as the Balancing Factor, and contains the remainder of UG, which cannot be calculated directly. The Balancing Factor is comprised of UG elements for which data is either unavailable or unreliable.

A key drawback of the old RbD-based method for estimating the UG total was that it was only capable of estimating the volume of UG assigned to the LSP sector. Whilst, as discussed in [10], the AUGE believed that the volume of UG assigned to the SSP sector was small, this was only true given the high bias in LSP AQ values relative to SSP AQ values at the time. The most recent data obtained in 2012 showed that this difference in AQ bias is not present and therefore invalidates the assumption that SSP-assigned UG is small. The RbD based approach to estimating UG cannot therefore be reliably used to estimate total UG without estimating the SSP-assigned UG separately. The use of consumption data in the UG estimation process allows the actual total, including both LSP-assigned and SSP-assigned UG, to be calculated. This is a key advantage of the new method.

It is known that data for each of the five potential components of UG (Unregistered and Shipperless sites, IGT errors, Shrinkage error, Shipper-responsible theft and metering errors) is available, along with meter read and consumption data for all supply points (except CSEPS and those sites that have no read history), and other general background data on RbD values, AQs, allocation algorithm coefficients, etc. The availability and quality of this data varies from component to component, and the AUGE has therefore attempted to identify the best method of calculating the total level of UG and the split between its causes based on the quality of information available for each component.

The proposed approach as derived in 2012 is therefore to first estimate the UG total for each LDZ, which can be defined as follows:

$$\text{Total UG} = (\text{Alloc SSP} + \text{Alloc LSP}) - (\text{Metered SSP} + \text{Metered LSP}) \tag{3.1}$$

This can be expressed as follows:

$$\text{Total UG} = \text{Aggregate LDZ Load} - \text{DM Load} - \text{Shrinkage} - (\text{Metered SSP} + \text{Metered LSP}) \tag{3.2}$$

Figure 2 below shows the ‘Gas into LDZ’ component. This is made up of NDM demand, DM demand, Shrinkage and UG along with their respective measurement errors. There is also an overall error in the measurement of gas entering the LDZ. Subtracting LDZ metering errors (estimated), the sum of DM metered volumes (measured but subject to measurement error) and Shrinkage (estimated) leaves the sum of NDM demand plus UG plus any error in estimating Shrinkage.

The ‘Metered Gas Out’ component is calculated using meter read information for every meter point. Where possible, the consumption for the formula year in question is calculated from meter reads or metered volumes. Where this calculation is not possible, an EUC-appropriate average value is used for this meter point. More details of this process are given in Section 6.1.

Having obtained an estimate of gas going into the network and gas being metered across all meter points, the difference between the two is our best estimate of UG plus Shrinkage estimate error. There will also be as yet undetected LDZ offtake meter errors and DM meter errors. This is shown in Figure 3. Note that the calculated total consumption across all meter points will have an error associated with it, which in turn will affect the estimate of UG. This overall consumption error consists of the error in estimating consumption at individual meter points which is based on either meter reads or average EUC consumption.

The total UG figure calculated thus far contains both permanent and temporary UG. Some elements of UG are subsequently corrected for and billed. These temporary sources of UG need to be removed from the total UG to obtain the total permanent UG. More details of temporary and permanent sources of UG are given in Section 3.3.

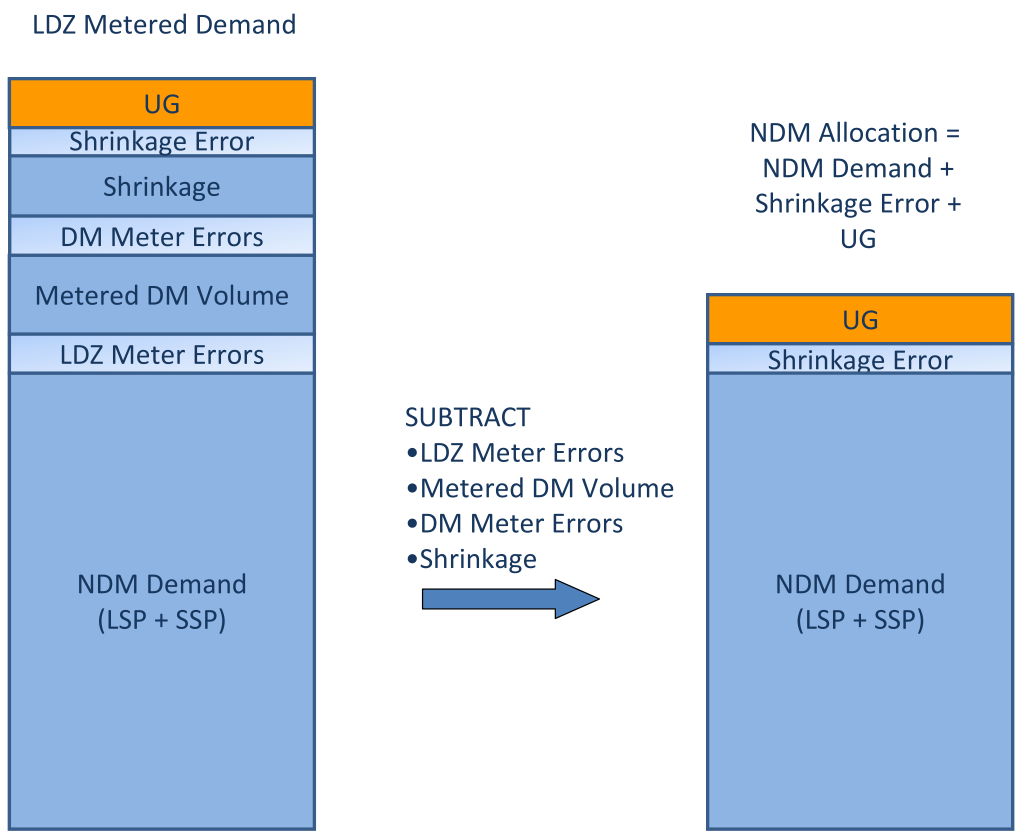


Figure 2: Derivation of Unidentified Gas

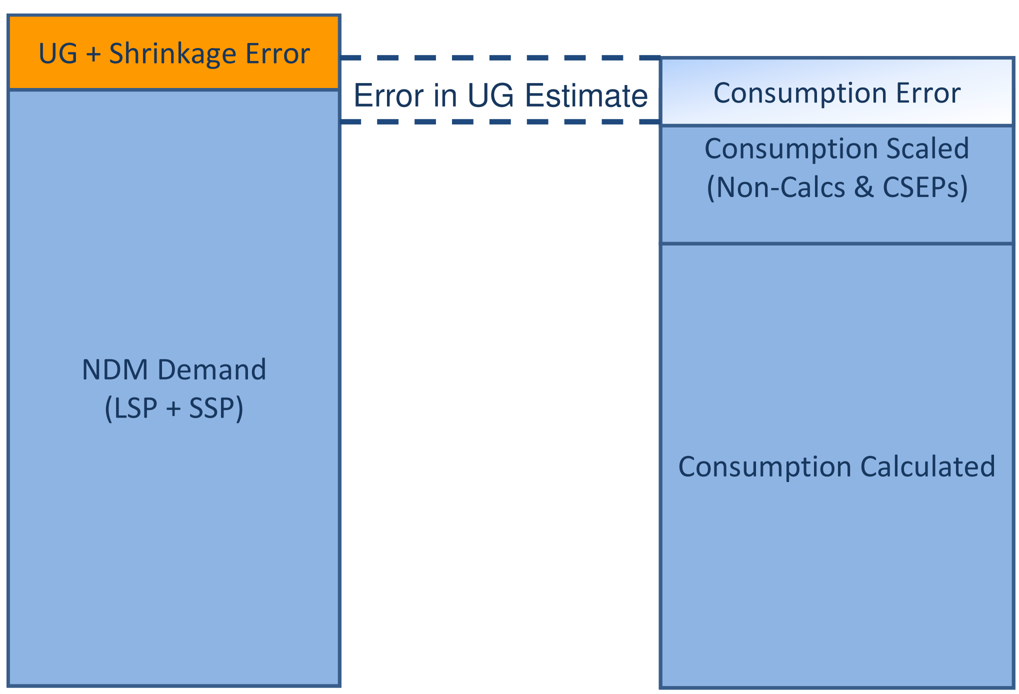


Figure 3: Calculation of Unidentified Gas from Consumptions and Allocations

### 3.2.2 Calculating Total UG Components

Elements of the UG total that have good quality data can be estimated directly, with the remaining elements for which insufficient data exists to produce a robust estimate grouped together and calculated by subtraction as the Balancing Factor.

Full details of this approach to the analysis, including full descriptions of the calculation methods for Total UG and for each individual element, are provided in previous versions of the AUGS [10], [19]. Brief descriptions of each UG element are given below.

a) Unregistered and Shipperless Sites

The AUGE believes these sites should be included in the UG calculations. The data required for this element consists of the historic number and AQ of sites either late registered or unregistered, split by cause and market sector. UG from this source is then calculated by assigning calculated consumption profiles to the validated AQ values from these sites. Unregistered and Shipperless sites that contribute to UG are split into the following sub-categories:

- Shipper Activity
- Orphaned Sites
- Unregistered <12 Months
- Shipperless PTS (Passed to Shipper)
- Shipperless SSrP (Shipper Specific Report)
- Without Shipper <12 Months

b) IGT Connected System Exit Point (CSEP) Setup and Registration Delays

IGT CSEP setup and registration delays should also be included in the UG calculation. UG from this source is due to gas networks owned by iGTs but not present in Xoserve's records, and also comes from unregistered sites on known CSEPs. The data required for this analysis consists of the number and composition of these unknown projects (number of sites and AQ split by market sector), and the number and AQ of unregistered sites associated with known projects.

c) Shrinkage Error

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

In addition to this, in the UG estimation process the figures for *Total LDZ Input minus Shrinkage minus DM Load* are calculated using allocations. Initial estimates of Shrinkage are used during the allocation process, and the final Shrinkage estimates may differ from these.

Shrinkage Model errors are very hard to quantify, given that actual Shrinkage is unknown and that the models are built on the most accurate data available. At the time they were trained they were, by definition, unbiased, and this may remain the case. If this is true, each individual instance of Shrinkage model error may affect the UG total that relies upon it, but these errors will even out over time, leaving a net effect of zero. If changing conditions over time have led to the Shrinkage model becoming biased, these effects will be picked up by the Balancing Factor, and this is therefore where this element will be captured.

d) Shipper-Responsible Theft

The AUGE believes that this element should be included in the UG calculation. Very little reliable data on theft exists, however, 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%. It is therefore very difficult to accurately estimate theft levels directly, and for this reason theft will be calculated by subtraction. It is part of the Balancing Factor, and considered over time, it forms the vast majority of that figure (based on an assumption that the Shrinkage models are unbiased, so their individual contribution can be positive or negative and will sum to a value close to zero over time).

e) Meter Errors

Meter errors can affect UG depending on their source. Errors in LDZ offtake metering and DM supply metering affect the estimate of total NDM demand including UG, whilst LSP NDM and SSP metering errors contribute to UG by affecting the NDM metered total. The AUGE has assessed this area and corrections are applied to LDZ offtakes, DM and unique site meters. SSP meter errors have been assumed to cancel out, further industry led studies would be required to evaluate this further.

The calculation processes detailed above will allow a reliable estimate of UG to be calculated based on the latest available data, which will in turn be used to populate the UG table, the format of which is given in Section 7. It also gives a sound basis for the year-on-year update of these figures, given appropriate provision of up-to-date information as requested.

## 3.3 Permanent and Temporary Unidentified Gas

Regardless of the calculation method used, certain elements of UG are permanent and others are temporary. The definitions of these terms are as follows:

**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 backbilling.

For all directly calculated elements of UG, the data supplied to the AUGE relates to all UG sources, both permanent and temporary. It is therefore necessary to split these into the correct category and only include permanent UG sources in the final calculations.

Table 2 below shows the permanent/temporary status of each element of UG.

Table 2: Permanent and Temporary UG

| Unidentified Gas Source | Type |
|---|---|
| iGT CSEPs | *Temporary* for LSP sites on CSEPs.<br>*Permanent* for SSP sites on CSEPs. |
| Shipperless/Unregistered | |
| - Shipper Activity | *Temporary* if shipper carries out site works.<br>*Temporary* if a third party carries out site works but asset meter read is the same as the shipper's opening meter read.<br>*Permanent* otherwise. |
| - Orphaned | As for "Shipper Activity". |
| - Unregistered <12 Months | As for "Shipper Activity". |
| - Shipperless PTS | *Permanent* for sites that became Shipperless prior to 25th January 2013<br>*Temporary* afterwards |
| - Shipperless SSrP | *Permanent* |
| - Without Shipper <12 Months | *Permanent* if the site would be classified as SSrP once it has been Shipperless for 12 months, or if it became Shipperless prior to 25th January 2013<br>*Temporary* otherwise. |
| Meter Errors | *Temporary* for detected errors that are corrected within the reconciliation period.<br>*Permanent* otherwise. |
| Theft | *Temporary* for detected theft.<br>*Permanent* for other theft. |

# 4 Summary of Analyses

This section contains a summary of the analysis work carried out during preparation of the 2013 AUGS for 2014/15.

## 4.1 MMSPs

During the development of the consumption method in 2012 it was noticed that some of the AQ records provided by Xoserve contained AQs and EUCs which did not correspond. Specific examples were raised with Xoserve and were identified as meter points which are part of a Multi-Metered Supply Point (MMSPs). We were informed that meter points in an MMSP are assigned an EUC based on the aggregate consumption of the supply point and as such may not match the AQs at the meter point level.

As all of the AUGEs consumption calculations are carried out at the meter point level the AUGE chose to calculate its own EUC values based on the meter point AQs. Note, as they are solely based on AQs these EUCs do not contain WAR band information.

It is particularly important that these meter point level EUCs are used when the consumption calculation fails because they are used to fill in a consumption estimate based on EUC average consumption levels. If the consumption calculation for an individual meter point in an MMSP fails then it should obviously be filled in based on that meter point's AQ/EUC rather than the aggregate AQ/EUC of the supply point which may contain dozens of meter points.

As an example, take supply point 59004 from EA LDZ which has 4 meters as shown in Table 3. The supply point has its EUC band based on the aggregate AQ of 4 meters which places it in band 03B. During the consumption calculation, it may be the case that consumptions are successfully calculated for 3 meters but for some reason fails for the remaining meter. Clearly if the consumption calculation fails for meter 772255, then the consumption for this meter should be taken as the average for EUC band 01B rather than 03B.

Table 3: Example MMSP 1

| | ID | AQ | EUC |
|---|---|---|---|
| Supply Point | 59004 | 407,937 | 03B |
| Meter Points | 761895 | 139,160 | 02B |
| | 761579 | 139,849 | 02B |
| | 771922 | 123,509 | 02B |
| | 772255 | 5,419 | 01B |

A further example is shown in Table 4. This supply point has an EUC band 05B with seven meters with EUC bans 01B to 04B. As in the previous example, if the consumption fails for meter 768834 then the average for the 02B EUC band should be used to replace it rather than the supply point level 05B band.

Table 4: Example MMSP 2

| | ID | AQ | EUC |
|---|---|---|---|
| Supply Point | 72941 | 2990341 | 05B |
| Meter Points | 761425 | 521045 | 03B |
| | 763964 | 2001867 | 04B |
| | 768834 | 151004 | 02B |
| | 771415 | 1 | 01B |
| | 771421 | 233176 | 02B |
| | 772964 | 75054 | 02B |
| | 1739797 | 8194 | 01B |

Within the AUGE's dataset there are 85,047 supply points with multiple meters. However, this number is based on a count of unique supply point IDs which change when a supply point is reconfirmed with a different shipper. The actual number of MMSPs will therefore be less than this. The AUGE believes that these meter points are being correctly calculated and that the meters are not being incorrectly "downgraded".

## 4.2 Use of LSP Meter Reads

This concerns LSP meter points with reads, but which fail the consumption calculation because of negative metered volumes or consumption estimate which is greater than five times the AQ. It has been suggested by ICoSS during consultation that in these cases, raw meter reads should be used to calculate consumption if possible.

The methodology proposed in the 2nd draft 2012 AUGS for 2013/14 used only metered volumes for LSP meters. This approach was taken because the metered volumes can be corrected if errors are discovered, but meter reads are not. It would therefore be expected that the metered volumes would always be the best data to use given they have been through further stages of validation and correction. This may not be the case if the meter units and/or T&P correction factors used to calculate the metered volumes are incorrect. However, in these cases it would be expected that the error would be identified and the metered volume subsequently corrected.

Table 5: Current SSP and LSP calculation success rates

| LSP | Total Sites | Successful consumption | % Successful consumption | Insufficient Meter Reads | % Insufficient Meter Reads | Negative Volume | % Negative Volume | AQ Check failure | % AQ fail |
|---|---|---|---|---|---|---|---|---|---|
| 2009 | 333,159 | 272,423 | 81.8 | 52,304 | 15.7 | 4,639 | 1.4 | 3,793 | 1.1 |
| 2010 | 324,415 | 264,995 | 81.7 | 52,723 | 16.3 | 5,214 | 1.6 | 1,483 | 0.5 |
| 2011 | 312,387 | 261,855 | 83.8 | 43,741 | 14.0 | 5,302 | 1.7 | 1,489 | 0.5 |
| 2012 | 309,111 | 252,929 | 81.8 | 49,750 | 16.1 | 4,391 | 1.4 | 2,041 | 0.7 |

| SSP | Total Sites | Successful consumption | % Successful consumption | Insufficient Meter Reads | % Insufficient Meter Reads | Negative Volume | % Negative Volume | AQ Check failure | % AQ fail |
|---|---|---|---|---|---|---|---|---|---|
| 2009 | 21,392,589 | 18,125,245 | 84.7 | 3,096,001 | 14.5 | 126,218 | 0.6 | 45,125 | 0.2 |
| 2010 | 21,423,817 | 19,336,952 | 90.3 | 1,940,024 | 9.1 | 104,944 | 0.5 | 41,897 | 0.2 |
| 2011 | 21,465,907 | 19,693,475 | 91.7 | 1,652,965 | 7.7 | 82,824 | 0.4 | 36,643 | 0.2 |
| 2012 | 21,490,547 | 15,107,858 | 70.3 | 6,289,794 | 29.3 | 67,179 | 0.3 | 25,716 | 0.1 |

The majority of LSP meter points which fail the consumption calculation do so because there aren't suitable meter read records to use. Table 5 shows the overall (2009-2012) LSP calculation rates. SSP calculation rates are also provided for information and show the lower number of SSP consumption successes in 2012 that we would expect given these are read less often.

For LSP meter points, only ~2% could potentially use meter reads as a fall back to metered volumes.

The success rates shown in Table 5 are similar to those achieved by Xoserve when they calculate AQs. We note that during the AQ generation process there is an amendment window which can result in a higher success rate of ~90%. However, the meter read data associated with the amendments are not included in the data sets provided to us by Xoserve, hence our overall success rate is different from the overall success rate that Xoserve report (including amendments). The amended AQs are provided in the AQ data and additional information has been requested from Xoserve to identify amended AQs.

As shown in previous sensitivity analyses (please see section 4.7.4 of [19]) a calculation success rate of 80% is, however, sufficient for the estimation of the total UG.

We are currently awaiting a refresh of the meter asset data (see section 4.17 of this statement) which we believe will further improve the quality of the data we have. However, it has not been possible to import and process the data and re-run calculations prior to the publication of this 2nd draft AUGS although the data will be updated for the preparation of interim figures in October. We do still intend to complete this analysis and the results will be reported at a later date.

For this statement, failed LSP consumptions will be replaced by EUC group averages.

## 4.3 Improved Handling of Read Units

During the consultation period for the 2nd draft 2012 AUGS for 2013/14, ICoSS identified two examples of meter points where read units had been incorrectly excluded from the consumption calculation process. The AUGE responses [20] explained the specific situation in which this occurred (which relates to meter points that change market sector). As a result the AUGE has updated the consumption algorithm to correctly handle this situation and the overall calculation method is described in Section 6.1.2.

## 4.4 Vacant Sites

During the consultation on the 2nd draft 2012 AUGS for 2013/14 a concern was raised by British Gas that although the AUGE has taken steps to account for non-consuming sites in the general population there may be an issue with vacant sites. In order for a meter to be excluded from the allocation process, the shipper must remove or disable the meter to isolate the supply. A shipper may not wish to do this where a site may remain vacant for only a short period. Modifications 0282/0282A [24] were raised by British Gas and ScottishPower to create a process for handling vacant sites including collecting data to flag such sites, but these modifications were rejected. A vacant site will therefore continue to be included in the allocation process with its latest AQ value until such time as two identical meter reads are received, at which point its consumption will be calculated as zero and its AQ will be set to one (assuming that the meter reads meet the criteria for consumption calculation). This may not happen if there is no access to the meter. It was argued that sites with no meter read data for an extended period have a higher propensity to be vacant and as such be non-consuming.

This raises the obvious questions of how long a site should be without reads before it can be assumed to be vacant and how long vacant sites are unoccupied for on average. Neither of these can be satisfactorily answered with the available data.

Table 6 below shows the number of live meters with no meter reads after the start of the formula year within the entire NDM meter population. For example, 6,357 meters did not have any meter reads from 1st April 2009 onwards. As would be expected this number is significantly higher at the end of the dataset as infrequently metered sites are awaiting a new read.

Table 6: Number of Meter Points Without Reads

| Year | Number of Meters |
|---|---|
| 2009 | 6,357 |
| 2010 | 56,036 |
| 2011 | 297,545 |

Of the 6,357 sites with no reads for 3 years, 760 have a rolled over AQ of 1. We have no information about whether the remaining 5,597 meter points are consuming or not, but we can quantify the possible effect by comparing the UG estimates under two scenarios: all such meters consuming vs. all such meters non-consuming.

We have calculated consumptions under both scenarios for all meters with no meter reads for three years using our method of scaling based on average consumptions per EUC band. The difference is approximately 90 GWh per year. Given that some of these meters will be consuming and some will only be vacant for part of a year the true difference is likely to be much less. Unfortunately Xoserve do not hold any information which can be used to determine explicitly if a site is vacant. Mod 0282/0282A proposed that a shipper confirm a site is vacant to Xoserve, who store this information. For the AUGE to be able to further quantify the effects of vacant sites, this information would be required.

Going forward our method assumes that these meters will behave in the same way as the wider population (i.e. same ratio of consuming to non-consuming meters) and will be scaled up accordingly.

## 4.5 Seasonal Normal UG

The methodology in the 2nd Draft 2012 AUGS for 2013/14 involved calculating the total UG by taking the difference between the total LDZ NDM demand (LDZ demand – shrinkage – metered DM demand) and the sum of all metered demands. However, the demand in any given year is subject to non-seasonal weather effects. As the UG estimate is forward looking, it makes sense that it be based on seasonal normal weather.

The AUGE therefore proposes that the allocations and the consumptions are scaled to seasonal normal conditions prior to subtracting to obtain total UG. The resulting UG estimate will therefore be based on seasonal normal weather.

During consultation, the question of whether it is appropriate to apply seasonal normal adjustment to theft was raised. This is discussed further below. A sensitivity analysis has also been performed to assess the potential impact on the estimate of theft.

The demand of all buildings will be temperature sensitive to some extent, with the EUC bands recognising this through different deeming algorithm parameters. When a consumer is thieving it would be expected that both the temperature sensitive and non-temperature sensitive components of demand will increase. The increase in the heating load due to theft will depend on temperature i.e. the extra gas stolen for heating would be more in a cold year than in a warm year. It is therefore correct that the gas stolen is adjusted for non-seasonal weather but there is uncertainty over size of this adjustment. It should be noted that as the seasonal normal adjustment applied scales the consumption based on the temperature difference from seasonal normal, this scaling could increase or decrease the consumption.

In order to assess the scale of the adjustment applied, the ratio of seasonal normal national allocation to actual national allocation has been calculated for the 3 formula years 2009-2011. Overall, the seasonal normal adjusted allocation was 1% lower than the actual allocation (1.1% lower in 2009, 6.3% lower in 2010 and 4.9% higher in 2011).

## 4.6 Temperature and Pressure (T&P) Conversion Factors

A concern was raised by Energy UK on behalf of British Gas that large LSPs (AQ > 732,000 kWh), which should have site specific T&P factors, may have been incorrectly assigned the default T&P factor of 1.02264.

Looking at the factors provided by Xoserve, around 25% of these large sites appear to have the default conversion factor. It is difficult to assess the impact if these values are incorrect (as they may or may not have been used when the metered consumptions were calculated and may or may not feed into the AQ calculation process) but most of the non-default factors are in fact close to this value. Therefore the AUGE believes that this is unlikely to be a significant issue, especially compared to the issues with Read Units.

There are, however, clear cases of the number of dials being entered as a T&P factor, T&P factors of zero and T&P factors that look like the default factor with the decimal point in the wrong place.

As noted previously, where we have identified potential T&P factor issues we will flag these to Xoserve to follow up with the relevant Shipper. Table 7 shows the number of meters in the general population which appear to have incorrect T&P factors. Where T&P factor issues are present for larger LSP meters the impact is potentially significant.

Table 7: Suspicious T&P Correction Factors

| Correction Factor | Number of Meters |
|---|---|
| 0 | 1,334 |
| 0.02264 | 10 |
| 0.02564 | 2 |
| 0.10226 | 12 |
| 0.11264 | 2 |
| 0.12264 | 23 |
| 1 | 28,728 |
| 3 | 3 |
| 4 | 7,356 |
| 5 | 4,359 |
| 6 | 342 |
| 7 | 14 |
| 8 | 22 |
| 10.2264 | 46 |
| 26 | 4 |
| 102.264 | 6 |

These factors are not currently used by the AUGE in the consumption algorithm itself but they are used as part of the data pre-processing (see 6.1.1 for full details) to:

- Infer read units from the meter reads and metered volumes
- Calculate corrections to LSP meter volumes where there are gap/overlaps in the metered periods

Where the T&P factor is recorded as being 0 the default factor is used instead.

It should be noted however that incorrect T&P factors may have been used in the calculation of metered volumes provided to Xoserve by the Shippers. It is important, therefore, that all code parties make every effort to ensure that the T&P conversion factors are correct as these underpin calculations of consumption used in many other downstream processes including estimating UG.

## 4.7 Theft Analysis

In the 2nd Draft 2012 AUGS for 2013/14, it was proposed that throughput should be used as the basis for estimating the Balancing Factor, which we believe mainly consists of undetected theft split between the SSP and LSP market sectors. There were a number of reasons for this, as follows:

- Using detected theft to estimate the theft split for undetected thefts inherently assumes that the detected thefts are a representative sample of all thefts. This is unlikely to be the case as
    1. The detected thefts are a very small proportion of total thefts – the detected thefts are about 0.5-0.6% of the Balancing Factor (comparison of average annual theft occurring per year with the Balancing Factor from the 2011 AUG Table).
    2. Theft detection is likely to target larger thefts where the financial benefits from detection are greatest.
    3. Larger thefts are potentially easier to detect. In particular, SSP thefts other than cases where a high proportion of the AQ is stolen will be difficult to detect and prove.
    4. There are arguments that thefts are easier to detect in some market sectors rather than others and/or that the effort expended in theft detection varies between market sectors.
    5. It can be argued that undetected thefts will continue for longer, as detected thefts are generally cut short by being detected, i.e. the average length of a detected theft doesn't necessarily give an indication of the average length of an undetected theft.
- The detected theft data consists of estimates of both the time period and quantity of theft, which are both subject to error. This has a knock-on effect in terms of uncertainty in market sector classification.
- Theft split methods that rely on theft detection rates can be potentially influenced by Shippers.
- There is a disincentive to detect theft for theft split methods that rely on the amount of detected theft in a given sector.
- Other non-theft related components of the Balancing Factor should be split by throughput
- AQs for theft-affected sites are volatile because the data they are based on is generally sparse and potentially theft-affected, compared to AQs calculated for the wider population.

### 4.7.1 ICoSS Alternative Method

ICoSS suggested an alternative approach to estimating the theft split based on their analysis of the theft dataset. The detected theft data was used to estimate the average theft quantity (per instance of theft) for each market sector. ICoSS then assumed, based on qualitative arguments, that the fraction of sites stealing was the same between market sectors. Their calculations gave an average SSP theft of 23MWh and an average LSP theft of 74MWh. A theft split was then proposed using the following equation:

*Total Theft of Gas (MWh) = % theft sites x (Total # SSP sites x 23MWh + Total # LSP sites x 74MWh)*

Assuming approximately 22,070,000 SSP meters and 310,000 LSP meters, this would give a theft proportion of ~4.3% LSP.

In the responses to the consultation, the AUGE noted that there were a number of issues with the data used in the ICoSS analysis but agreed to review their calculations using the most appropriate data and to consider the merits of the proposed approach further.

### 4.7.2 Assessment of ICoSS Method

Following the assessment of the proposed method, the conclusion was drawn that there are a number of fundamental flaws in the methodology proposed and the data used by ICoSS:

1) The data set used was from a spreadsheet dating from the AUGE's first year of analysis, which included a calculation to estimate the annual theft rate for a detected theft site (i.e. the amount of gas that would have been taken over a full year if the theft had not been detected). In 2012 it was concluded that scaling up theft to an annualised rate was incorrect, primarily as there are many examples where a modest amount of theft has been estimated to have occurred over a very short time scale (e.g. a few days), resulting in unrealistically high rates of theft per day. Such cases are likely to be due to inaccuracies in the estimated theft duration, and scaling up of the theft to obtain an annual rate therefore potentially over-estimates the amount of theft that actually occurred. Table 9 shows examples where the period of theft is less than 30 days and the 'current' AQ is less than 73,200 kWh, but the resulting 'adjusted' AQ is greater than 73,200kWh. This demonstrates how relatively small amounts of theft can result in massive annual quantities when the period of theft is estimated to be so small. Many of these are unlikely to be realistic, although data on the capacity of the meters installed would be required to validate this fully. A typical domestic meter (e.g. U6) has a maximum flow rate of 6m$^3$/hr and can only pass 565,896 kWh per annum, and this would require a continuous maximum flow rate. Of course larger meters may be in place and the sites in question may in fact be LSP sites, but we do not have customer details with which to investigate further. In situations where the meter is bypassed before the governor, then the pressure of gas would be higher resulting in higher flows – however, this could potentially result in appliance problems depending on their input pressure rating. The overall effect of the annual theft rate calculation phenomenon is to increase the average AQ for SSP sites and potentially incorrectly reclassify some sites as LSP.

Table 8: Theft Records with Short Periods of Theft Resulting in Large Annual Theft Rates

| MPR | Sector | AQ/kWh | Theft (kWh) | Period of theft (days) | Theft per year (kWh) |
|---|---|---|---|---|---|
| 3001 | SSP | 21,027 | 18,580 | 1 | 6,781,700 |
| 2977 | SSP | 14,177 | 7,742 | 1 | 2,825,830 |
| 709 | SSP | 1 | 3,500 | 1 | 1,277,500 |
| 3661 | SSP | 11,752 | 20,148 | 6 | 1,225,670 |
| 2288 | SSP | 4,848 | 53,333 | 22 | 884,843 |
| 4378 | SSP | 10,963 | 34,667 | 19 | 665,971 |
| 682 | SSP | 1 | 7,149 | 5 | 521,877 |
| 6738 | SSP | 17,303 | 5,425 | 4 | 495,031 |
| 4000 | SSP | 57,274 | 6,133 | 6 | 373,091 |
| 3736 | SSP | 6,335 | 14,083 | 14 | 367,164 |
| 4190 | SSP | 25,960 | 18,401 | 21 | 319,827 |
| 4189 | SSP | 4,401 | 18,401 | 21 | 319,827 |
| 7265 | SSP | 7,654 | 6,704 | 9 | 271,884 |
| 2557 | SSP | 5,563 | 10,436 | 16 | 238,071 |
| 3692 | SSP | 3 | 667 | 1 | 243,455 |
| 6297 | SSP | 12,667 | 10,639 | 19 | 204,381 |
| 7459 | SSP | 12,247 | 8,400 | 16 | 191,625 |
| 2917 | SSP | 16,213 | 8,000 | 18 | 162,222 |
| 3287 | SSP | 28,719 | 9,545 | 25 | 139,357 |
| 2333 | SSP | 8,587 | 9,600 | 25 | 140,160 |
| 3453 | SSP | 26,646 | 3,001 | 9 | 121,707 |
| 7304 | SSP | 16,792 | 5,813 | 17 | 124,809 |
| 4336 | SSP | 1 | 5,845 | 16 | 133,339 |
| 1074 | SSP | 11,106 | 9,158 | 28 | 119,381 |
| 1360 | SSP | 43,596 | 4,000 | 17 | 85,882 |
| 2097 | SSP | 736 | 10,000 | 29 | 125,862 |
| 3467 | SSP | 10,410 | 7,573 | 25 | 110,566 |
| 6180 | SSP | 9,536 | 5,480 | 18 | 111,122 |
| 6567 | SSP | 61,633 | 2,386 | 16 | 54,431 |
| 3899 | SSP | 104 | 7,796 | 29 | 98,122 |
| 4125 | SSP | 58,000 | 2,216 | 23 | 35,167 |
| 2338 | SSP | 31,751 | 3,773 | 24 | 57,381 |
| 6865 | SSP | 7,279 | 2,000 | 9 | 81,111 |
| 1737 | SSP | 32,987 | 3,430 | 24 | 52,165 |
| 130 | SSP | 40,559 | 1,805 | 17 | 38,754 |
| 3892 | SSP | 36,523 | 733 | 7 | 38,221 |

2) ICoSS calculated an adjusted AQ based on the AQ in this original data set added to the annualised theft rate. The AQs in the data set are **current** AQs (i.e. those in place in 2011) and therefore they are not necessarily representative of the AQs that were in place at the time of the theft. Therefore, by adding the theft to the current AQ, the level of consumption is being over-estimated resulting in meters moving incorrectly to the LSP sector and the average SSP AQ being over estimated.

3) In the original spreadsheet ICoSS referenced for the derivation of this method, there is a table showing the number of sites and detected theft volumes by sector after adjusting the AQ for theft. This resulted in a theft split of 25.1% (see Table 9). This is included in this report to illustrate the variety of theft splits that can be achieved depending on what data you choose to use. As noted previously, the adjusted recent AQs should NOT be used for this calculation (since they are not

representative of the AQ at the time of the theft, and theft should only be added where the AQ is affected by unmetered consumption). Therefore this split is not appropriate to use. We highlight it here to demonstrate that very different theft splits can be obtained from the same data set depending on what assumptions are made.

Table 9: Theft Split by Volume using Adjusted AQ from ICoSS Spreadsheet

| Adjusted | Number | % of Total | TOG Volume | % of Total |
|---|---|---|---|---|
| SSP | 4,244 | 91.17% | 84,855,197 | 74.87% |
| LSP | 411 | 8.83% | 28,487,639 | 25.13% |

4) ICoSS have assumed that the number of sites likely to steal in each sector as a proportion of the population is the same. Table 10 below shows the number of theft records (from the original data that was used for their analysis), the overall population of sites for each sector (this is an average over 2009 and 2010 from the consumption data we have), and the rate of theft detections per 100,000 of the population of each sector. This indicates that the rates of detected theft by sector are not the same and there are approximately three times as many LSP thefts per 100,000 of the population compared to SSP.

Table 10: Theft Rate (sites) per 100,000 Sector Population

| Sector | No theft records | Population 100,000s | Rate per 100,000 sites in population |
|---|---|---|---|
| SSP | 4,474 | 220.7 | 20.3 |
| LSP | 181 | 3.1 | 58.4 |

5) The data for thefts that occurred prior to 2006/7 are generally of poor quality and should be treated with caution.

Notwithstanding the above, we have looked at the method ourselves and taken account of the following:

1) In our previous analyses we have generally looked at the level of theft that has occurred in a given year (rather than detected by year), and calculated an annual theft rate occurrence from that. However, when looking at theft rates per unit AQ we cannot scale thefts that occur over multiple years to a yearly rate and not scale up part year theft periods, as we would not be comparing like for like. Therefore to be consistent we have calculated the theft rate per annum although we note that the results will be skewed by those records where theft occurs over a short period of time (as shown in Table 8).
2) As we have consumption data, we can calculate metered plus unmetered consumption to give an AQ that is a fairer reflection of what the AQ was during the period of theft. As many meters fail the consumption calculation, an alternative estimation of AQ is derived from the pre-theft AQ or post-theft AQ plus the amount of theft occurring in a given year.
3) The site theft rate per market sector population has been calculated for both sectors using our AQs for sector classification to estimate what the difference in rate is across sectors.
4) Sites that are unregistered have been excluded from the dataset.
5) There are a number of issues that occur when trying to recalculate the ICoSS table using the more recent theft data and AQs:

a. The ICoSS analysis uses one AQ to cover the whole period of theft (in this case the 'current' AQ). To do this correctly we need to use an AQ that is representative of the period of theft to ensure the correct EUC band is used.

b. In our data sets we have AQs by year based on metered plus unmetered consumption, or pre-theft AQs as appropriate when the consumption cannot be calculated. The question then is which AQ to use? Should an average AQ be used over the period of theft? Should the theft rates be considered on a year by year basis (i.e. matching the theft per annum with the AQ for the year of the theft being examined)? The latter is not possible for all theft records as some go back to the 1990s and the AQ data does not go back that far. For the purposes of comparison we have used average AQ over the period of theft for a given theft record.

c. The average AQ by EUC band associated with the theft records cannot be compared directly to the ICoSS version. As noted in a) this is because multiple AQs occur over the duration of the theft. Using adjusted AQs as already noted is incorrect because it skews the AQs to LSP (as illustrated in Table 9). It is not really possible to calculate a satisfactory average AQ that can be used to compare theft rates per AQ between EUC groups as noted in b), in addition AQs may or may not be theft affected.

The results produced by ICoSS are shown in Table 11 below. Using the datasets from 2012 we have produced an equivalent set of results as shown in Table 12.

Table 11: ICoSS Results

| EUC | 01B | 02B | 03B | 04B | 05B | 06B |
|---|---|---|---|---|---|---|
| Average Theft Duration (years) | 1.5 | 1.3 | 1.1 | 1.4 | 1.1 | 0.5 |
| Average Theft Volume (MWh) | 22.7 | 73.9 | 83.9 | 51.3 | 87.1 | 11.8 |
| Average Adjusted AQ (MWh) | 23.7 | 127.8 | 463.6 | 1134.2 | 7273.3 | 18498.0 |
| Theft Volume / AQ | 0.96 | 0.58 | 0.18 | 0.05 | 0.01 | 0.00 |

Table 12: Equivalent ICoSS Results using Average Calculated AQ

| | 01B | 02B | 03B | 04B | 05B | 06B | 07B | LSP Average |
|---|---|---|---|---|---|---|---|---|
| Average Duration / yrs | 1.69 | 1.92 | 0.92 | 1.23 | 2.30 | 1.36 | 0.75 | 1.86 |
| Average Theft Volume (MWh) | 19.47 | 71.02 | 161.94 | 158.01 | 6.95 | 8.94 | 13.02 | 74.88 |
| Average AQ (MWh) | 15.58 | 121.06 | 450.89 | 1,197.20 | 5,700.47 | 12,564.50 | 22,691.74 | 416.33 |
| Theft Volume/AQ | 1.250 | 0.587 | 0.359 | 0.132 | 0.001 | 0.001 | 0.001 | 0.180 |
| Number Records | 5048 | 143 | 5 | 3 | 1 | 1 | 1 | 154 |

The ICoSS method then chose an amount of 23MWh as the average theft for an SSP site and 74MWh as the average theft for an LSP site.

Using the results of our overall analysis we would use a figure of 19.47Mwh for SSP and 74.47MWh for an LSP site.

Before we calculate the overall theft split, we calculate the difference in theft rates per 100,000 of the overall population as shown in Table 13. This is different from the figures shown in Table 10 as there are more theft records in our data set and the adjusted AQ is replaced by a better AQ approximation, which results in fewer sites being misclassified as LSP. This is then used to correct the initial theft split since the rate of thefts between populations is not the same. Table 14 shows the resulting theft split. The key difference between this and the results from the ICoSS method is the assumption that the number of thefts per unit of the population is the same across sectors, which the data shows is not the case.

Table 13: Corrected Theft Rate (sites) per 100,000 of Population

| | No. Sites | Population 100,000s | Number thefts / 100,000 of population |
|---|---|---|---|
| SSP | 5,048 | 220.7 | 22.87 |
| LSP | 154 | 3.1 | 49.68 |

Table 14: Adjusted Theft Split using Population Correction

| | LSP Split using average theft per site/% | LSP detected rate multiplier | Theft Split/% |
|---|---|---|---|
| **Theft Split** | 5.12 | 2.17 | 11.13 |

We now consider this method in terms of the potential sources of error in the previously derived methods that relied on detected theft data. In the Interim Report [13] and the 2nd draft 2012 AUGS for 2013/14 [19] we highlighted that methods using detected theft suffered from the following issues:

- The calculations are still heavily dependent on the accuracy of the estimate and duration of theft.
- Use of AQs - particularly as we can only calculate metered consumption for 50% of the data set.
- Use of AQs – using a singular AQ for the entire period of theft is not necessarily appropriate because demand varies over time and the AQ used may not represent the site's true behaviour (and classification). It is also unclear whether the AQs are theft-affected or not.
- Accuracy of the metered consumption calculation (Unit Reads, T&P factor issues).
- Potential effect of customer changes on pre/post theft AQs.
- Site classification issues – e.g. Unregistered sites.
- Assumption that the market sector split of unknown theft is the same as that of detected theft.
- The potential still exists to a certain degree for external influence on the theft split if mixed shippers focus on detecting theft in one sector over another, because this could affect the average theft amount per sector (rather than the amount of detected theft overall).

As part of this review we also took a closer look at the periods of theft. Figure 4 shows a histogram of theft periods in 30 day intervals. Note the very large spike corresponding to the bucket that contains the theft period of one year (in fact ~500 theft records have a theft period of exactly 365 days). This raises a serious question about the periods estimated for detected theft because this spike constitutes a clear outlier. Are the periods for these thefts really exactly one year? We suspect that a year has been used as a default option in many cases, and this will skew the calculated average period of theft. In addition, if when the theft is recorded the period of theft is then used to estimate the amount of theft (based on typical consumer consumption rates) then these too will be erroneous. This in turn will impact on the results of the ICoSS theft split method, and any other method that uses this information.

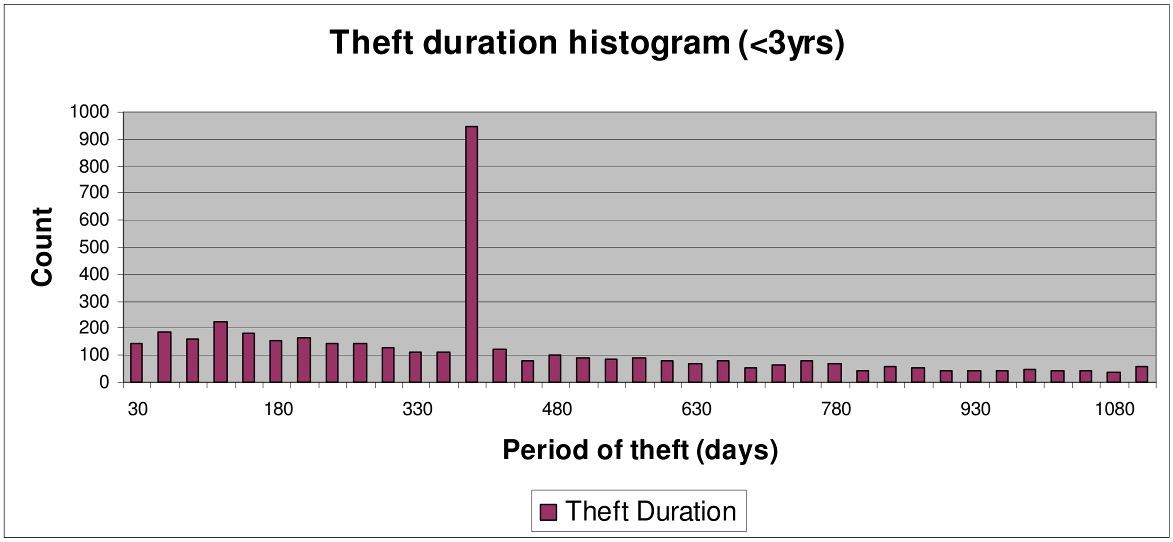


Figure 4: Theft Duration Histogram, 30-Day Intervals

### 4.7.3 Comparison of Methods

This table summarises the key data issues associated with detected theft.

Table 15: Theft-Related Data Issues

| Source Data | Issue | Comments |
|---|---|---|
| Theft period | Very short periods of theft coupled with modest theft amounts are unlikely to be correct and result in physically impossible annualised gas flows. Periods of theft are agreed with the consumer (SPAA Appendix 4 [26]) but may be different in practice, especially for estimates prior to the SPAA.<br><br>A large portion of the data set has the theft period set either exactly or very close to a one year, which is likely to be erroneous and skews the results. | We conclude that this data is unreliable and should not be used for theft analysis. |
| Theft amounts | The SPAA guidelines provide recommendations for estimating theft but they may vary from Shipper to Shipper in practice, especially for estimates prior to the SPAA. Some theft volumes, if scaled to annual level, could be physically impossible. | This data is potentially unreliable, especially if the theft period is used to estimate the theft amount. |
| AQs | Current AQs do not represent past behaviour (certainly not for thefts that occurred several years ago). Calculation of robust AQs are difficult given the scarcity of meter read data, and it is difficult to obtain a reliable average AQ for theft affected sites.<br>This issue has the potential to shift large proportions of detected thefts between market sectors. | We conclude that this data is unreliable (in the theft-affected population) and should not be relied on for theft analysis. |

The following table compares the AUGE’s theft occurrence rate method from 2011/12, the ICoSS proposed method, and the AUGE’s proposed throughput method in terms of how they use theft data and their associated assumptions.

Table 16: Comparison of AUGE Theft Methods and ICoSS Method

| | **Theft Occurrence Rate Method** | **ICoSS Method** | **Throughput Method** |
|---|---|---|---|
| Period of theft | Used to estimate the amount of theft that occurred in each year (unscaled). | Used to calculate an annual rate of theft. | Does not use period of theft. |
| Amount of detected theft | Apportioned to each AUG year covered by period of theft. | Used to calculate an annual rate of theft. | Does not use amount of detected theft. |
| AQ estimate | Uses unmetered plus metered consumption, and if this fails defaults to AQs prevailing prior to the theft period. Failing that uses post theft AQs corrected for theft. | Used current (2011) AQs. This could use AQs from theft occurrence rate method but these are different year on year. | Uses AQs to determine market sector for the whole population of meters. The population of theft-affected meters is much smaller and hence the “whole population” results are less susceptible to volatile AQs. |
| Assumption that the proportion of theft sites per unit population is the same across market sectors | Does not use this assumption. | Uses this assumption (although evidence indicates the rates are different). | Does not use this assumption. |
| Assumption that total theft rate per unit throughput is the same across market sectors | Does not use this assumption. | Does not use this assumption. | This assumption follows from using the throughput method but it is not the reason for choosing it. |
| Handling of unregistered sites for detected thefts | Unregistered sites excluded. | Unregistered sites included (but could be removed). | Method is not affected by unregistered sites. |
| Impact of “current AQ plus theft per annum” method of estimating AQ | No impact, not used. | Significant impact, AQs over-inflated. | No impact, not used. |

Having concluded that the detected theft data is not sufficiently robust to be used to split theft in 2012, the alternative method proposed by ICoSS and underpinned by the same data has the same inherent issues and cannot be considered sufficiently robust either.

Significant improvements would need to take place in the recording and estimating of detected theft, and a much bigger sample obtained, in order to allow a sufficiently robust and stable method of theft split to be produced going forward.

In 2012 we concluded that detected theft data was not sufficiently robust to provide a suitable theft split methodology, leading to throughput being selected as the most appropriate method. It follows from using throughput as the method of splitting theft that the theft per unit AQ is assumed to be the same across each market sector, and this assumption may or may not be true. The only data available to assess the merits of this assumption is the detected theft data, however, and this suffers from serious issues as described above.

The detected theft data as it stands does indeed show some differences between market sectors in theft per unit AQ, but until this analysis can be carried out using reliable data it cannot be used to invalidate the throughput method.

Taking all of these issues into account, and having reviewed the proposed ICoSS method in detail, we believe the only fair and most robust method for splitting theft by market sector is the throughput method.

## 4.8 Impact of Changes to Supplier Licences to Increase Theft Detection Rates

Earlier in the year OFGEM introduced changes to Supplier licences to increase theft detection rates [30]. There is an expectation that the number of detections will increase from 3000 to 6000 per annum [29].

In the proposed methodology we deduct detected theft by year of occurrence from the total estimate of UG which will be @20-25GWh based on theft data to end of June 2012.

From the consultation of the previous draft AUGS it was suggested that this figure be increased to account for the higher theft detection rates following the OFGEM changes to Supplier licences.

In the data sets provided by Xoserve the number of alleged thefts is of the order 3000-4000 per annum, however the number that are actually pursued and have theft estimates associated with them range from 500-1500 per annum and these may not all occur in the year of detection (since some thefts are not identified for several years).

There is no evidence as yet to demonstrate that higher levels of detection are actually being achieved. If the estimate was doubled, but the levels of detection not achieved (or even over achieved) there is no back correction facility in the AUGE process to subsequently correct.

We conclude that the average estimate of detected theft by year of occurrence from previous years will be deducted from the Total UG for 2014/15. If levels of detection do demonstrate an increase this can be reviewed in a future statement.

We also note that the materiality of this is relatively small compared to other elements of UG.

## 4.9 Improvements to the Theft Split Throughput Method

In the 2nd Draft 2012 AUGS for 2013/14, it was proposed that throughput should be used as the basis for estimating the theft split between SSP and LSP market sectors. At this time, throughput values were taken from the ODR1209 (OFGEM Data Request) report and extrapolated forward. However, there are some issues with using this data in its raw form. The AUGE notes that if this data were to be used, a number of corrections would need to be applied as follows

- RbD – the ODR1209 values are raw allocations and as such do not include reconciliation corrections. RbD values should therefore be used to correct these allocations. One major downside of applying the RbD corrections to the allocations is that the SSP throughout value used to calculate the split will include all UG resulting in a slightly larger SSP proportion. RbD corrections also represent corrections applied in a given year rather than corrections which should apply for a given year. It must therefore be assumed that these corrections are stationary.

- Meter Errors – the raw allocations are subject to metering errors and should be corrected for LDZ and DM meter errors in the same way as when calculating UG

An alternative approach to using the allocation values with corrections applied is to use the consumption values as calculated to estimate UG. Using the AUGE's calculated consumptions has the advantage that the effects of UG can be correctly allocated by market sector, and no assumptions are made about the statistical properties of RbD. Where UG has been calculated directly (e.g. Shipperless sites), this should be included.

It should be noted that both the allocations and the calculated consumptions are subject to non-seasonal weather related effects. It is therefore proposed that SSP and LSP consumptions are adjusted to seasonal normal conditions (using CWAALP, in the same way that AQ values are calculated). This will ensure that when projecting forward, the estimated market sector split will be on a seasonal normal basis.

The throughput method proposed in 2012 was based on a trend line extrapolated for future years based on a consistent trend of the ratio of LSP throughput vs SSP throughput. More recent data shows that this trend does not continue and the percentage of LSP throughput increased compared to SSP in 2011 and 2012. In order to produce a method that has longevity, the AUGE proposes that a rolling average based on the most recent 3 years data is used to be consistent with the period used for total UG estimation. This will be calculated by summing the seasonal normal SSP and LSP consumptions over the 3 year period and using these values to calculate the SSP/LSP split.

The recommended method is described in full in Section 6.8.

## 4.10 Metered Volume Corrections

Metered volumes are calculated from raw meter reads and asset information (meter units, T&P factors). For LSP meters, corrections can be made to metered volumes through the reconciliation process when an error has been identified. The majority of these corrections occur within the first year (see modification 0398 workgroup report [7]), but there are corrections which can go back several years. These corrections can only be applied if they fall after the reconciliation backstop. This was set at 4 years (increasing to 4 years 364 days by the end of the formula year), but the implementation of Mod 0398 has reduced this to 3 years (increasing to 3 years 364 days by the end of the formula year). The implications of Mod 0398 are discussed in more detail in Section 4.19.

The metered volume data used by the AUGE to estimate UG is a snapshot of data at a given point in time. As such, this data will be subject to corrections at some future point in time. This section discusses the issues of corrections to LSP metered volumes that have yet to be applied.

When calculating the UG for 2014/15, metered volume data for LSPs will be available for formula years 2009-2012. The AUGE does not intend to use data for 2012 as there will still be a significant number of outstanding corrections yet to be applied for this year. According to the Mod 0398 workgroup report [7], there will still be about 20% of the energy consumed still to be reconciled. This drops to about 5.5% for the preceding year (2011) and the unreconciled energy reduces further for each year prior to that.

As part of the calculation of UG for 2014/15, the AUGE proposed looking at the change in consumption based on metered volumes provided by Xoserve for the 2013/14 year and the newly calculated consumptions using the most up to date data. By assessing the change from year to year, the AUGE planned to assess whether a correction is necessary to account for the as yet unreconciled energy, and if so, how this correction should be estimated. Unfortunately, due to issues identified with the previous dataset, there will be a step change in consumption due to this and it may not be possible to gain meaningful insight from comparing the two sets of consumptions for this year's statement since the year on year corrections also include missing consumptions from last year which we cannot separate out. The AUGE recommends this is assessed in the future.

## 4.11 AQ Correction for Unregistered Sites

In the calculations of UG from Unregistered Sites (Orphaned, Shipper Activity and Unregistered <12 Months), AQ values as supplied in the snapshot files provided by Xoserve were used directly for production of the figures in the AUGSs for previous years. It has become apparent, however, that this approach will lead to errors in the resulting UG estimates.

The AQs reported in the snapshots for all categories of Unregistered sites are the Shippers' Requested AQs, which have undergone no validation whatsoever and are not necessarily representative of the real consumption of each site.

There are in fact three stages in the assignment of an AQ to a new site, and therefore this requires us to make 2 separate corrections in the UG calculation process (i.e. AQ1→AQ2 and then AQ2→AQ3).

AQ1 is the Requested AQ:
This is the initial AQ requested by the Shipper. It has undergone no validation of any sort and is whatever the Shipper asked for. It is this value that appears in the Shipperless and Unregistered snapshots.

AQ2 is the Confirmed AQ:
The Requested AQ undergoes some rudimentary validation (i.e. whether it is appropriate for the type of site). If it passes then it becomes the Confirmed AQ, if it fails then a new, more appropriate, AQ is set. In either case the value is still just an approximation.

AQ3 is the AQ Following AQ Review:
This is the first AQ for the site that is actually based on any meter read data and hence the first that can be regarded as accurate.

Analysis has been undertaken to calculate these effects by comparing AQs from each stage to the equivalents from the next stage.

- Factors have been calculated using requested and confirmed AQ data that convert from AQ1 to AQ2, with separate factors for each type of Unregistered site. This is new analysis carried out during 2013 for the 2014/15 AUGS. Based on the data supplied by Xoserve to date, the following factors have been calculated for each type of Unregistered Site.

Table 17: Requested AQ to Confirmed AQ Conversion Factors

| Type | Total Confirmed Sites | Average Requested AQ | Average Confirmed AQ | Conversion Factor |
|---|---|---|---|---|
| Orphaned Sites | 447 | 123,059 kWh | 117,614 kWh | 0.96 |
| Shipper Activity | 19 | 198,261 kWh | 198,261 kWh | 1.00 |
| Unregistered <12 Months | 10904 | 64,965 kWh | 49,573 kWh | 0.76 |

- Factors have previously been calculated using confirmed and post-review AQ data that convert from AQ2 to AQ3. In this case a single universal factor was found to be most appropriate. This work was carried out in 2012 and is described in Section 4.2 of the 2nd draft 2012 AUGS for 2013/14 [19].

Therefore, in order to convert to the final, reliable, AQ the calculation process for Unregistered Sites becomes:

*AQ1 → Factor 1 → (Estimate of) AQ2 → Factor 2 → (Estimate of) AQ3*

This process will therefore be applied to the calculations for production of the UG figures for 2014/15 and for subsequent years.

## 4.12 CSEP Shrinkage

During consultation of the 2nd draft 2012 AUGS for 2013/14; ICoSS raised the issue of CSEP Shrinkage as this is not calculated as part of the GT estimation of Shrinkage for their networks We concluded that this issue should be raised at the Shrinkage forum. The issue was raised by ICoSS at a recent Shrinkage forum meeting and National Grid Transmission have since provided an initial assessment of CSEP shrinkage [25], suggesting an overall leakage of approximately 13GWh. We will continue to monitor the progress of this and consider the conclusions that are subsequently reached by the Shrinkage forum.

## 4.13 Prime and Sub Meters

Prime and Sub meters are a special configuration that are dealt with by a dedicated Prime and Sub team within Xoserve. The following description is taken from Xoserve's Customer Discovery Day Presentation [28]

"*There are three meter types:*

- *FREESTANDING: A meter connected to the gas inlet which has no other billable meters connected downstream*
- *PRIMARY: A meter connected to the gas inlet which has one or more billable meters supplying gas to other off-take points downstream*
- *SUB DEDUCT: A billable meter connected downstream of a billable Primary meter. Gas off-taken at each of these off-take points is billed independently for each meter*

*Primes & Subs - A site can be set up with many sub meters which feed through the prime. Each meter in the configuration has its own load, including the prime. Prime & Sub meter configurations can have a* *mixture of NDM and DM meters and all meters can be supplied by different shippers. Prime Meter Reconciliation can only be performed when reconcilable reads are obtained for all meters in the configuration within a 5 day window. This is known as a set of 'Co-Terminus' reads. A Prime meter will only be reconciled once all of its Subs have been read and reconciled."*

Meter reads and metered volumes provided by Xoserve for prime meters refer to the prime meter i.e. they refer to the total flow/energy including all sub deduct meters. However, when an AQ is calculated for a prime meter, this is based on the difference between the prime meter and all of its constituent sub deduct meters. In cases where the sum of the energy for the sub deduct meters is greater than or equal to that of the prime meter, the prime meter is assigned an AQ of 1.

In our methodology when a consumption is successfully calculated for a prime meter it will include the consumption of the sub meters and therefore we need to exclude the sub meter consumptions to avoid double counting. When the calculation fails for a prime meter its consumption will be filled in using the average consumption for its EUC band. As this EUC is based on its own individual consumption this will not lead to double-counting.

We have identified 1,443 Prime and 2,799 Sub meters in our data sets. The number of instances of this configuration could reduce going forward if Mod 0428/0428A is implemented as this will prevent any new instances being created. Primes and Subs in our historic data set would still need to be handled correctly.

As noted in Xoserve's description of Primes and Subs, Primes can have Sub meters that are DMs. Of the Prime and Sub meters noted above, there are 29 Prime meters that have a DM Sub meter.

An example is shown in Figure 5. This is an MMSP with top level EUC band 08B. It consists of five meters, two of these are DM (one with an AQ of 1 for the last 7 years) and there are three NDM meters one of which is a Prime meter.

The Prime meter consists of four Sub meters, three of which are DM and one NDM. The net consumption of the Prime meter (i.e. for the demand that it feeds) is the difference between its total consumption less the sum of its sub meters resulting in an EUC band 07B.

In order to calculate the NDM consumption for the MMSP correctly, we only need the consumptions of the NDM meters (as highlighted in Figure 5). However, we do not have DM data to use in the calculation.

Our recommended solution to this is to fail the consumption calculations for those primes and sub meter configurations with sub meters that involve DMs.

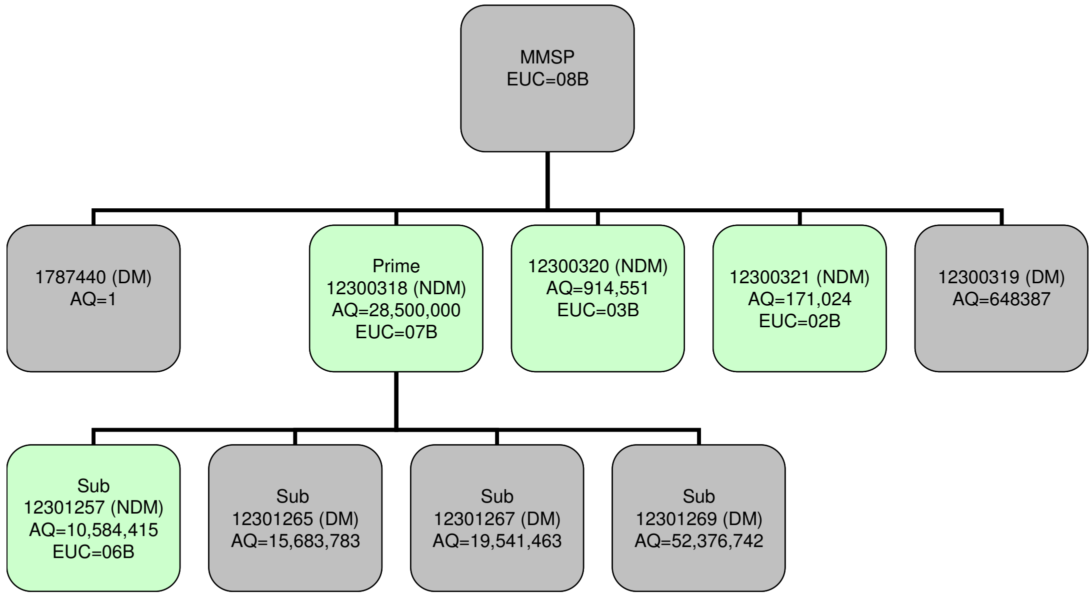


Figure 5: Primes and Sub meter example with mixed DM/NDM meters

## 4.14 Meters with LDZ reassignment

Where a meter has been moved from one LDZ to another since 01/10/2007 the meter will appear in the AQ and meter read records for both LDZs (AQ records split between LDZs based on the LDZ the meter was assigned to at the time). This causes the following issues

- As the AQ records are used to produce a list of all meters within an LDZ, a meter which has been in more than one LDZ will be included in the meter list for multiple LDZs, resulting in a "double-counting" of its consumption
- As the records for a meter are split between LDZs and processing is carried out on an LDZ by LDZ basis, the consumption calculations and subsequent validation will not be optimal as only a subset of the data is available in each LDZ

To overcome the issue, a pre-processing step is performed which identifies the current (latest) LDZ for each meter which has been in multiple LDZs. Data from previous LDZs is then copied to the appropriate data tables for the current LDZ to ensure that a full dataset is available for the consumption calculation on the current LDZ. The meter list for each LDZ is then populated to ensure that each meter appears in only one LDZ (the latest), thus preventing double counting.

The issue of meters being assigned to the incorrect LDZ means that the UG calculated on an LDZ basis will not be 100% correct. However, these discrepancies cancel out when aggregated over all LDZs.

In the current dataset, 71,500 meter points have been identified which are associated with more than one LDZ.

## 4.15 Selection of AQ for validation

The final stage in the consumption calculation is to validate our estimate against the AQ of the meter point. The algorithm selected the AQ to use for comparison using the following logic:

Select the first AQ estimate effective after the end of the formula year being calculated. If none exists after the end of the formula year use the latest value. In addition we restrict attention to the latest value from each gas year.

The problem with this approach is that for monthly metered sites we may end up using an inappropriate value, e.g. for MPR_ID 768300 we calculate a consumption of 342,868 kWh for 2010. The AQ records are shown in the table below.

Table 18: AQ selection for consumption validation

| MPR_ID | AQ_EFFECTIVE_DATE | EUC | AQ | SITE_TYPE_FLAG | LDZ |
|---|---|---|---|---|---|
| 768300 | 01/10/2007 | 05W03 | 374046 | N | EA |
| 768300 | 01/10/2008 | 05W03 | 444884 | N | EA |
| 768300 | 01/10/2009 | 05W03 | 439990 | N | EA |
| 768300 | 01/10/2010 | 05W03 | 438043 | N | EA |
| 768300 | 01/10/2011 | 03W02 | 315367 | N | EA |
| 768300 | 16/08/2012 | 03W02 | 1 | N | EA |

The meter actually stopped consuming on 30/09/2011. Based on this the AQ is set to 1 from 16/08/2012. However, we are interested in the period 01/04/2010 to 31/03/2011 when the meter was consuming the whole time. Therefore the AQ value effective 01/10/2011 is the most appropriate to use.

As an AQ of 1 was selected the AQ check fails and the site is given a ~75% chance of having zero consumption and a ~25% of 6,208 kWh (which was the average consumption of sites with AQ=1 and a successful non-zero calculated consumption) leading to an underestimate of around 340 MWh.

The restriction of only using the latest AQ from each gas year has been removed and the AQ chosen will be the next available value after the end of the formula year. This change will potentially impact the consumption estimate for at most 700 meter points; however they are nearly all LSPs including some 07B and 08Bs. The impact of this correction will increase the overall estimate of consumption and reduce total UG.

## 4.16 Estimation of CSEP Consumption

When estimating total consumption, it was previously proposed that CSEPs are treated in the same manner as meters which fail the consumption calculation i.e. they are estimated based on their EUC band. However, the AUGE has discovered that the EUC band for CSEPs is based on a nominal maximum AQ value rather than a real estimate of the AQ. Using the EUC band will therefore bias the CSEP consumption estimates.

Given the limited data available for CSEPs, the only alternative to using the EUC average is to assume that the nominated AQ values (not maximum AQs) are representative of the true seasonal normal consumption.

Using AQ values rather than EUC averages significantly reduces the estimated level of CSEP consumption which in turn leads to an increase in the estimate of UG.

## 4.17 Cosmetic and real meter Exchanges

A sample of cosmetic meter exchange data has been provided so far. This will improve the quality of the meter asset data which is used in the calculation.

## 4.18 Consumption Gaps and Overlaps

The most recent data set from Xoserve includes a refresh of the LSP data and an additional field has been added to aid in validation. Previously each meter read record had an associated read date. The new dataset includes not only the read date but the previous read date which was used in order to calculate the corresponding metered volume. Xoserve identified that there are some gaps in the data i.e. that the meter reading start and end dates may not be contiguous. The AUGE has also found a smaller number of meter points where there is an overlap in the periods over which metered volumes are calculated.

For the LSP sector where consumption is calculated by summing the metered volumes, it is important that there are no gaps or overlaps in the metered volumes. In the dataset supplied in 2012 we were unable to identify missing meter reads or overlaps because we only had meter read dates (i.e. no start dates). Going forward, the methodology has been updated to detect any such gaps/overlaps and account for them in consumption estimates. This is done by using the meter reads and asset information to calculate a metered volume. Full details are provided in section 6.1.

In the current data set 22,595 LSP meter read records have been identified with a gap or overlap.

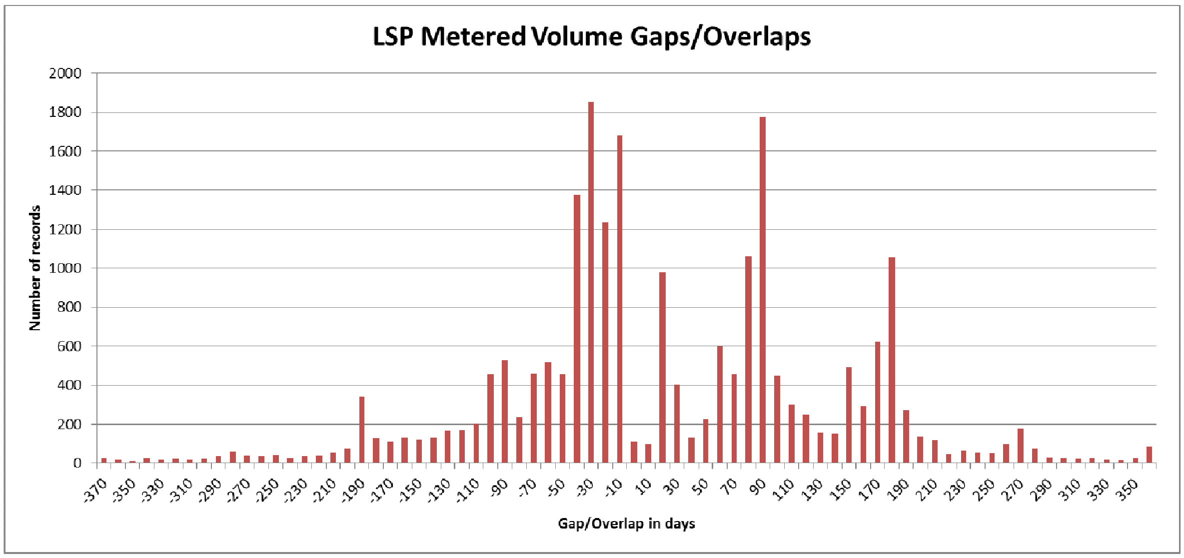


Figure 6 : Histogram of LSP gaps and overlaps

A histogram of the gaps and overlaps for LSPs is shown in Figure 6. There are clear spikes at 1, 3 and 6 months corresponding to typical meter read frequencies. In our previous estimates of total consumption the effect of gaps and overlaps will cancel out to a certain extent. By resolving this issue the impact on total consumption will be to remove the residual gap/overlap.

## 4.19 Industry Initiatives under Review

In the 2011 and 2012 AUGSs the AUGE identified a number of industry initiatives that may have an impact on UG going forward. New modifications have also been raised that, if implemented, may also have an impact on the AUGE processes. Modifications that could potentially impact on the UG calculations are summarised briefly below.

**Mod 0398 Limitation on Retrospective Invoicing and Invoice Correction**

Mod 0398 [7] proposes reducing the reconciliation window to 3-4 years. This has been approved by Ofgem and effective from 1 April 2014 [22]. Mod 0395 was a similar modification with a different reconciliation window of 2-3 years. Mod 0398 was implemented in preference to Mod 0395.

In line with this modification, a backstop date of 3 years will be set at 1 April each year. This backstop remains for the whole year giving an increasing reconciliation window as the year progresses, up to a maximum of 3 years 364 days before the backstop is then reset to 3 years for the following year.

The final modification report for Mod 0398 [7] includes data from Xoserve which shows that only a very small proportion of reconciliations occur in the 4-5 year window and concludes "Reducing the reconciliation window would therefore have a minimal impact on energy allocation". The AUGE would therefore expect the impact on UG to be correspondingly minimal. This will also align the reconciliation window and the period of data used in total UG calculation.

One exception to this may be Significant Meter Errors (SMEs). Ofgem noted in their decision letter regarding Mod 0398 that "Several respondents raised concerns about the impact UNC 0395 or UNC 0398 may have on the accurate reconciliation of Significant Meter Errors ('SMEs'), particularly in relation to an NTS-LDZ offtake meter given the length of time that an error may go undiscovered and subsequently take to resolve". Mod 0429 seeks to address this and has been approved by Ofgem; it is described later in this statement.

**Mod 0410/0410A Responsibility for Gas Off-taken at Unregistered Sites following New Network Connections**

Modification 0410/0410A [15], [21] seeks to reduce the amount of Unidentified Gas caused by Unregistered sites by introducing mechanisms to reduce this in terms of the responsible party. If approved, it intends to change the AUGE guidelines to specifically require the AUGE to take account of information from Xoserve when dealing with Unregistered sites.

Mods 0410 and 0410A are live modifications and have not yet been implemented so have no impact on the proposed AUG methodology at this time. The Modification Panel has recommended Mod 410A is implemented and the modification is awaiting OFGEM decision. The implementation timeline for Mod 410A is not currently defined.

If implemented, this modification would mean that UG arising from unregistered sites would be temporary going forward and so should be removed from the total UG estimate. However, as the proposal is not retrospectively applied, there will still be some unregistered sites which continue to contribute to the permanent UG initially, but this amount will reduce over time as the registration issues for these sites are resolved.

**Mod 0424 Re-establishment of Supply Meter Points – Prospective Measures to address Shipperless Sites**

Modification 0424 [16] seeks to reduce the impact of Shipperless sites on Unidentified Gas. This modification was implemented as of 25 Jan 2013 [23] and so will have an impact on the level of Unidentified Gas attributable to Shipperless sites for 2014/15. The AUGE will account for the effects of this modification through the information currently provided by Xoserve.

This modification means that UG arising from Shipperless sites where the original meter is still present (Shipperless PTS report) would be temporary going forward and so should be removed from the total UG estimate. However, as the proposal is not retrospectively applied, there will still be some Shipperless sites which continue to contribute to the permanent UG initially, but this amount will reduce over time as the registration issues for these sites are resolved. Xoserve have agreed to provide the date the site became Shipperless in their snapshot files so that the AUGE can determine those sites that the modification applies to.

**Mod 0425 Re-establishment of Supply Meter Points – Shipperless Sites**

Modification 0425 [17] aims to reduce the impact of Shipperless sites on Unidentified Gas if implemented. It will place an obligation on the last registered Shipper to take responsibility for investigation and resolution of the registration of the site (either to re-register it from the date of registration or to register it with another Shipper). If implemented, the AUGE would pick up the effects of this through the information currently provided by Xoserve. If implemented it will be treated in a similar manner to Mod 0424.

Mod 0425 is a live modification and has no impact on the proposed AUG methodology at this time. A final modification report was published on 2nd July. The implementation timeline is not defined.

**Mod 0429 Customer Settlement Error Claims Process**

Modification 0429 [14] seeks to address the mismatch between the reconciliation window (currently 4-5yrs but reducing to 3-4yrs following implementation of Mod 0398) and the Limitation Act (6yrs). This results in energy invoices between Shippers and Customers that are adjusted in the Limitation Act period not being reflected in the energy allocation settlement in the current UNC process.

The Modification Panel recommended that Mod 0429 should not be made. However, OFGEM have approved Mod 0429 and will be effective from 1st April 2014.

The following text from the final modification report highlights the impact on the AUGE process:

*"Solution*

*The proposed solution to the problems identified above is the creation of a claims process for Shippers to use when major loss is incurred in the gap between the end of the reconciliation window and the Limitation Act.*

*The AUGE process may require a separate change through amendment of its guidelines statement to address the impact of a shorter reconciliation process than the Limitation Act and the creation of the proposed claims process. Any such change will be through a separate process under the change procedures for the AUGE Guidelines Statement."*

*“Impact on AUGE process*

*A key finding of the AUGE process has been that many aspects of Unidentified Gas are temporary in nature and will eventually be allocated back to an individual Shipper through the reconciliation process. The Customer Settlement Error Claims Process will not result in changes to reconciliation and so there seems to be a need to recognise, through the AUGE process, that some sources of Unidentified Gas would not be corrected owing to the current reconciliation backstop date.”*

*“AUGE Process Adjustment*

*This aspect of the change will be raised as a separate change to the AUGE guidelines statement. It is detailed here in order to give a comprehensive view of the implications of this proposal. The current AUGE process attempts to determine the scale of Unidentified Gas that is present in the settlements process. A key factor in determining the amount of Unidentified Gas that exists is determining which sources of unidentified gas are permanent (i.e. will never be allocated to an individual Supply Point) or temporary (i.e. will eventually be corrected at some point and allocated to an individual Supply Point). Shortening the current reconciliation time period will shorten the period in which settlement errors are corrected. It will therefore increase the amount of Unidentified Gas and other energy in the system that cannot be corrected through the reconciliation process and so be classified as permanent Unidentified Gas. The Settlement Error Claim process above will allow for a process to correct settlement errors beyond the reconciliation window, which may include corrections for sources of gas use which would have been originally classified as Unidentified Gas.*

*It will therefore be proposed that the AUGE should make an assessment of the amount of energy that would have been corrected (and so be classed as temporary Unidentified Gas) were it not for the close out of the reconciliation window. For the avoidance of doubt, this aspect of the change will be raised as a separate change to the AUGE guidelines statement. It is detailed here for the sake of completeness.”*

*“The AUGE would detail the materiality of this “fossilized” Unidentified Gas and adjust the resulting Unidentified Gas volumes accordingly.”*

We note that Mod 0429 is based on the assumption that the AUG methodology is the RbD method. Assuming the proposed consumption methodology which does not use RbD data is approved later this year the impact of Mod 0429 on the AUGE process may be different from initially expected.

We await a separate Modification or change to the Guidelines to adjust the AUGE process to address this.

With regards the expectation that the AUGE will identify and account for the unidentified gas that would have been corrected were it not for the close out window, there will be a fundamental reliance on data availability in order to assess this. It is likely that such corrections would need to be stored and maintained even if outside the close out process window in order to assess the likely materiality. Further discussions will be required during the formulation of the modification to change the AUGE process to ensure that such corrections can be estimated by the AUGE.

At the time of publication of the 2nd Draft 2013 AUGS for 2014 these corrections will not be addressed since this draft statement is due to be published on 31st July and subsequently approved in September and no changes to the AUGE process have been approved.

**Mod 0428/0428A Single Meter Supply Points**

This modification seeks to establish a rule that would permit only one meter per supply point for new or amended supply points. If implemented it would simplify the consumption calculation process used to estimate total consumption for the calculation of UG and reduce the confusion of supply point level AQs/EUC bands vs sub meter AQs/EUC bands. As our historic data set would still contain Multiple Meter Supply Points this would not have any impact in the short term, but over time the number of MMSPs may reduce.

The Modification Panel recommended Mod 0428 in preference to Mod 0428A and Mod 0428 was subsequently approved by OFGEM on 25th July 2013 and will be effective from 1st April 2014.

# 5 Data Used

This section describes the data requested, received and used to derive the methodology to calculate UG. The AUGE has taken care to ensure that all datasets include all components of NDM consumption, i.e. CSEPs and Scottish Independents are included throughout.

There have been a variety of issues with obtaining data in previous years. This was partly to do with the way the industry manages various processes. For example, the AUGE could not obtain a history of data relating to Shipperless/Unregistered sites over time as only current snapshots can be produced. However, Xoserve now provides regular snapshots so that trends can be identified over time.

In 2012 there were issues obtaining meter reads and metered volume data, and this is described in previous versions of the AUGS. Data requests for 2013 have been submitted to Xoserve and data preparation is underway to ensure that all required data will be available in time to generate the interim and final AUG tables later in the year. One key difference this year is that in the event of data issues (particularly with consumption data) we do have the fall back position of the data provided from 2012 which we did not have the benefit of last year.

Section 5.1 below gives a summary of the data items requested and their current status. The subsequent sections give more details about the data items for each individual element of the analysis.

As part of the AUGEs quality control process, a number of standard data checks have been defined which are run prior to performing any consumption calculations. Any anomalous data will be reported to Xoserve for further investigation. It is unlikely that all issues identified will be resolved prior to final calculation of the AUGS table, so by necessity this will be 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 UG.

## 5.1 Summary

Table 19: Data Status Summary

| Analysis Area | Dataset Requested | Status |
|---|---|---|
| Direct Total UG Calculation | Allocated NDM loads | Received |
| | Metered SSP and LSP loads | Received 7th June & 14th June |
| | LDZ, DM and Unique Sites Metering Errors | Requested |
| | Meter Asset Information | Received although an updated set is pending |
| | Algorithm data (ALPs, DAFs, EWCFs, WCFs, SFs) | Received 6th June 2013 |
| | CV data | Received 6th June 2013 |
| | CSEP AQ data | Received |
| | Non-CSEP AQ data | Received |
| Unregistered and Shipperless Sites | Asset and Shipper meter reads for new LSP sites | Received |
| | Asset meter reads for orphaned sites | Received |
| | Gas Safety Visit data | Requested |
| | Snapshot files | Received on an ongoing basis |

| Analysis Area | Dataset Requested | Status |
|---|---|---|
| iGT CSEPs | Known CSEP data | Requested |
| | Snapshot files | Received on an ongoing basis |
| Meter Error | Meter capacity report | Received |
| New Analysis | Multiple Meter Supply Point data | Received |
| | Primes and Subs information | Received |
| Additional Supporting information | Mod81 data 2012 | Received |
| Theft | Detected and alleged theft updated to end March 2013 | Requested |
| | AQs before, during and after theft | Requested |
| | Metered volumes and meter reads, Read Units and T&P factors for theft detected sites | Requested |
| | EUC groups and meter read frequencies for theft affected sites | Requested |

## 5.2 Total UG Calculation (Consumption Method)

Data has been requested from Xoserve in the following formats. In all cases, data has currently been provided for the time period 01/04/2008 to 31/03/2012. The supply of data for the 2013/14 formula year is ongoing.

- 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.
- A history of AQ and EUC data has been provided for each MPRN so that calculated consumptions can be validated against AQs and failed meter points can be replaced with an appropriate EUC average.
- 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 time period for which it was active.
- Aggregate MPRN count and AQ data by EUC for CSEPs. Meter read data is not available for these sites, but knowledge of the number and AQ of MPRNs allows them to be included in the total UG calculations when the sample consumption is scaled up to cover the full population.
- Meter asset information on an 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 a number of different parts of the consumption algorithm.

The provision of this data allows the consumption for each individual meter point, for each formula year of interest, to be calculated using the method described in the 2012 AUGS for 2013/14 [19]. The exact format of the data provided is also given in this document.

## 5.3 IGT CSEP Setup and Registration Delays

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

- Unrecognised projects summary, including
  - number of unknown projects by LDZ
  - count of supply points and aggregate AQ of unknown projects by LDZ

  This data is supplied by Xoserve in two-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.

## 5.4 Unregistered/Shipperless Sites

The following information is supplied by Xoserve concerning Unregistered/Shipperless sites. For all aggregate-level data both the number of sites and their aggregate AQ is included. All data is split by LDZ, and also between "Small AQ" and "Large AQ" categories.

Xoserve have created a regular report to ensure that new data is collated and sent to the AUGE every two months. 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, but find a meter connected to the service and flowing gas. If it is the same meter as allegedly removed 12 months ago 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 a site visit finds a new meter fitted, 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.

- Created <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.

This data is supplied by Xoserve in two-monthly snapshot files on an ongoing basis.

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

- A summary of the remaining Shipperless sites, i.e. those that have been without a Shipper for less than 12 months and hence do not yet appear in the “Shipperless PTS” or “Shipperless SSP” lists. This data comes from the records of Gas Safety Regulations visits.

- Asset meter reads for orphaned sites to determine the proportion which have been flowing gas prior to becoming registered.

- Asset and shipper details for a sample of confirmed sites. This is used to calculate the proportion of UG from Unregistered sites that cannot be backbilled.

Updated data for all of these items covering the time period up to February 2013 has been received.

## 5.5 Meter Errors

Data for meter error calculations consists of meter capacity, AQ and NDM/DM classification records for all LSP sites. This report is supplied on an annual basis, with the latest one having been received by the AUGE in July 2013. This dataset is therefore up to date and ready to be used in the UG calculations.

# 6 Methodology

This section describes in detail the methodology for each aspect of UG where the calculation method has changed since the last published analysis in 2012. Where methods have remained the same, details can be found in Section 6 of the 2012 AUGS for 2013/14 [19].or in the previous 2011 AUGS for 2013/14 [10]. Details of the Consumption Method have been included here for completeness because whilst they were given in the last AUGS they have yet to be used in a UG estimate. In addition, elements of the method have been developed since the last AUGS and so a full description of the latest version of the method is given here.

The Consumption Method can be stated in its simplest form as:

*Total UG = (Alloc SSP + Alloc LSP) – (Metered SSP + Metered LSP)*

This can be alternatively stated as:

*Total UG = Aggregate LDZ Load – DM Load – Shrinkage – (Metered SSP + Metered LSP)*

Unlike the RbD method first presented in the 2011 AUGS for 2012/13 [10], this method estimates the actual UG total, including both LSP-assigned and SSP-assigned UG. This is a key benefit compared to the RbD-based method, which estimates LSP-assigned UG only and uses this as the best estimate of total UG.

The Consumption Method in its raw form includes both permanent and temporary Unidentified Gas in its output. 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.

## 6.1 Total UG Calculation (Consumption Method)

The consumption algorithm relies on a large quantity of data, summarised in Section 5.2. 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 derive additional information and helps speed up the consumption calculation process. Appendix B describes the resulting dataset. 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.

### 6.1.1 Data Pre-Processing

In order to calculate the total UG figure it is useful to pre-process the raw data and derive some additional data fields. All data fields used by the consumption method are described in Appendix B. In this section we explain the methods used to derive those fields from the raw data.

1. ANNUAL_QUANTITY_XX(EUC_CALC)
This field is calculated using a function defined in the database called Calc_EUC_Band which returns an EUC band given an AQ using the following logic:

if AQ <= 73,200 then '01B'
else if AQ <= 293,000 then '02B'
else if AQ <= 732,000 then '03B'
else if AQ <= 2,196,000 then '04B'
else if AQ <= 5,860,000 then '05B'
else if AQ <= 14,650,000 then '06B'
else if AQ <= 29,300,000 then '07B'
else if AQ <= 58,600,000 then '08B'
else '09B'

2. FACTORS_DAILY(WC)
This is calculated as max(0.01, 1 + DAF * EWCF).

3. FACTORS_DAILY (WAALP)
This is simply ALP * WC.

4. FACTORS_YEARLY(CWAALP)
The sum of WAALP over formula year.

5. FACTORS_DAILY (VOL_WAALP)
This is WAALP / CV.

6. FACTORS_YEARLY(VOL_CWAALP)
This is CWAALP / CV.

7. METER_INFO_XX(UNITS_CALC)
This field is derived using the following method:

For each MPR, for each pair of meter reads ($mr_1$, $mr_2$) check

- The meter was not replaced in between
- $mr_2 - mr_1 > 0$
- metered volume ($mv_2$) associated with $mr_2 > 0$
- round the clock indicator associated with $mr_2 = 0$
- gap/overlap indicator associated with $mr_2 = 0$

If yes to all, calculate

$$Ratio = mv_2 / ((mr_2 - mr_1) * CF)$$

$$A = \log_{10}(Ratio)$$

$$B = abs([A] - A)$$

Where *CF* is the T&P correction factor (taken from the meter asset data) and *[A]* denotes the integer part of *A*.

If

- $0 \le [A] \le 4$
- $B < 0.002$

Then set *units_calc = 10*[A]. We then look up the latest meter_info entry prior to the meter reads and update the units_calc field. If no such entry exists a new dummy record is inserted with an installation date = $mr_1\_date$ -1 and default values for the other fields.

8. METER_INFO_XX(IMP_IND_CALC)

This field is the IMP_IND flag taken from the meter read records for the meter. It is stored here for comparison with the value from the meter asset data.

9. METER_READS_XX(BAD_READ)

The algorithm for flagging bad reads is as follows:

Given sequential meter reads $mr_1$, $mr_2$, $mr_3$ and $mr_4$ calculate:

$$con_1 = mr_2 - mr_1$$
$$con_2 = mr_3 - mr_2$$
$$con_3 = mr_4 - mr_3$$

If any of these are negative we check for meter index rollover (see Section 6.1.2.1)
If the meter was replaced we leave the consumption null
Then if the meter was not replaced during the period we check

- If ($con_3$ > 0) and ($con_2$ < 0) and ($con_1$ > 0) then we have a bad reading
    - If $con_1$ > abs($con_2$) then $mr_2$ is bad
    - Else if $con_3$ > abs($con_2$) then $mr_3$ is bad

10. METER_READS_XX(GAP_OVERLAP)

This field is the number of days between the start read date for a records and meter read date from previous record.

- A positive value indicates a gap
- A negative value indicates an overlap

Note, start read dates are only available for the records provided in the most recent data set.

11. METER_READS_XX(VOLUME_CORRECTION)

For the records with a gap/overlap this field is the difference between the volume consumed between that record and the previous read (calculated using the meter read values and relevant asset data) and the recorded metered volume.

12. NDM_DM_CHANGE(NDM_START_DATE, NDM_END_DATE)

The entries in this table are calculated manually. The first step is to list the MPRs from the AQ records which are recorded as both DM and NDM (there were 375 such meters during 2012 analysis). Then by inspecting the AQ records for each MPR determine the start and end date of its NDM status.

13. SITE_LIST (MPR_ID)

These are taken from the AQ records for each LDZ. We then identify meter points which “move LDZ” and remove their ID from the site_list for the incorrect LDZ.

14. SITE_LIST (START_DATE, END_DATE)

These are generated based on the NEW_LOST_SITE table and the NDM_DM_CHANGE table. Special care needs to be taken off “lost” sites which then become active again at a later date.

15. LDZ_MOVERS

Some meters have records associated with more than one LDZ. These meters are identified and their MPR_ID is recorded in a separate table which also includes the latest LDZ they are recorded against and any previous LDZs. Any data associated with a meter which is not for the latest LDZ is then copied to the appropriate LDZ. This is required as processing is done on an LDZ by LDZ basis with data held in separate LDZ specific tables. This copying of data ensures that when the meter consumption is calculated in the latest LDZ, all of the relevant data is present. To ensure no double counting occurs, the SITE_LIST tables are populated to ensure that each meter is only present in one LDZ (its latest LDZ).

### 6.1.2 Algorithm

In addition to the step-by-step description below, worked examples of both a standard consumption calculation and a meter index roll-over affected calculation are given in Appendix C.

1. Given a formula year Y, define the start and end dates as 01 Apr YY and 31 Mar YY+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 formula year. If none exists after the end of the formula year use the latest value. From this record store
    i. The AQ value
    ii. The EUC provided by Xoserve
    iii. The pre-calculated consumption band derived by the AUGE from the AQ value.
    iv. Market sector (SSP/LSP) based on the EUC from Xoserve

4. For each meter point find the meter reading date and value for:
    - LB1 (Lower Bound 1) – the latest metering reading prior to the start of the formula year
    - LB2 (Lower Bound 2) – the earliest meter reading within the formula year
    - UB1 (Upper Bound 1) – the latest metering reading within the formula year
    - UB2 (Upper Bound 2) – the earliest meter reading after the end of the formula year

   For SSPs we exclude those readings which have been flagged as bad by the pre-processing.
   Note that for any given meter point, only a subset of this full set of reads may be available. We need at least one lower bound and one different upper bound meter read. Possible scenarios are shown in Figure 7 below:

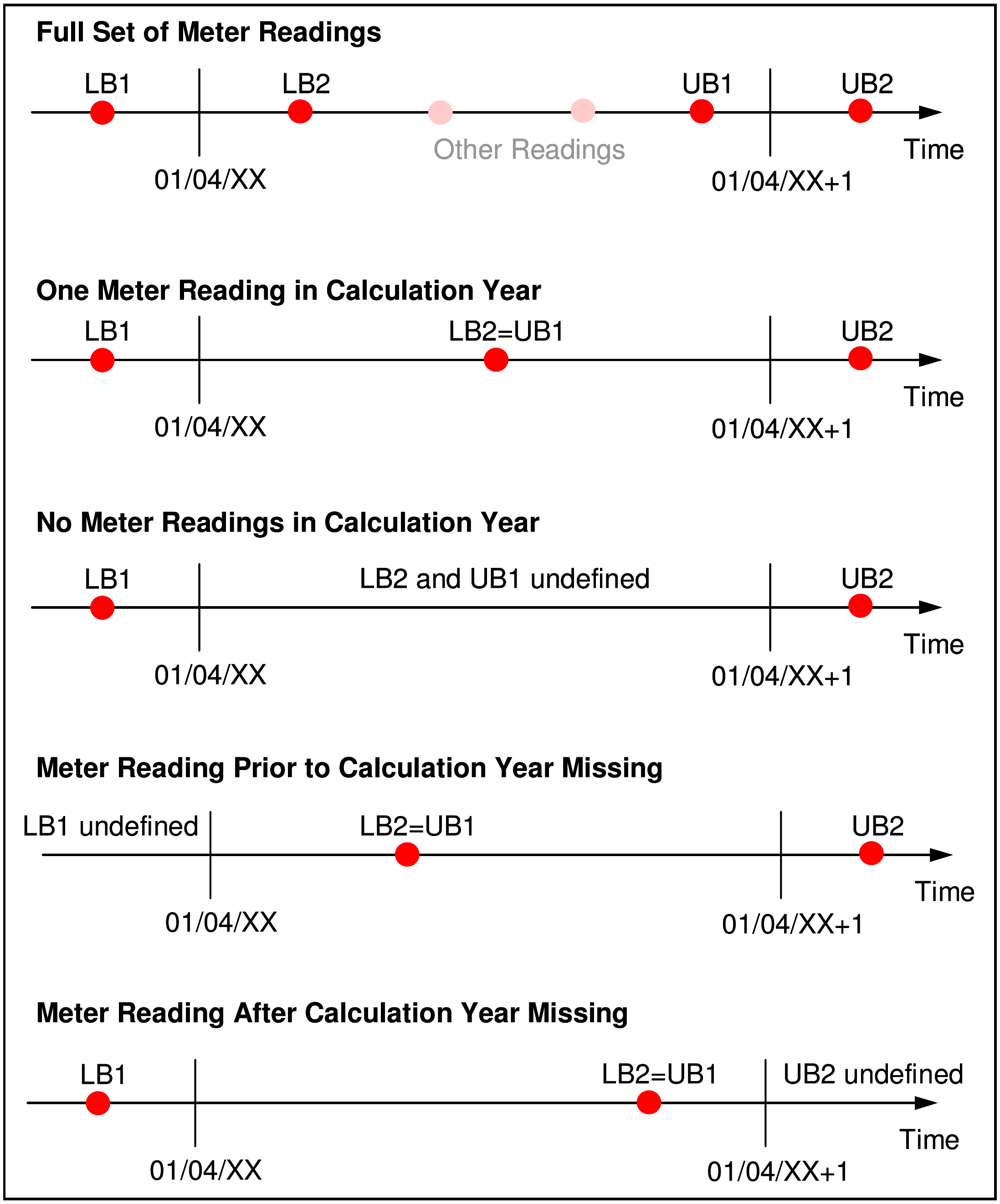


Figure 7: Meter Read Availability Scenarios

5. Set the start meter read date to LB1 unless
   - The date of LB1 is more than 540 days from the start of the formula year, or
   - 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
   - The date of UB2 is more than 540 days from the end of the formula year, or
   - 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 between LB2 and UB1 inclusive, then reject the meter point.

8. Check that:
   - The distance between the two chosen meter readings is at least 120 days
   - The overlap between the metering period and the formula year is at least 60 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 according to the market sector of the site:
   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 6.1.2.1 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.

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

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

   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 site 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 Xoserve.
       - In the case where the read units from Xoserve 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 just 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 formula year, calculated as

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

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

$$Con = (mr_2 - mr_1) * U * CF * CV\ /\ 3.6\,(*\ 0.0283168466\ if\ imperial)\ \text{for SSP} \qquad (6.1)$$

$$Con = \sum mv_i * CV\ /\ 3.6\,(*\ 0.0283168466\ if\ imperial)\ \text{for LSP} \qquad (6.2)$$

15. Calculate an AQ from this consumption

$$AQ = Con * 365 / CWAALP$$

16. If we have calculated a new AQ value from the meter readings that is more than 5 times larger than the old AQ and the new AQ puts the site in the LSP market then reject the meter point. Such sites may be manually reviewed as appropriate.

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 sub deduct consumption. See section 6.1.2.2 below for details.

#### 6.1.2.1 Meter Index Rollover Check

Given two reads $mr_1$ and $mr_2$ where $(mr_2 - mr_1) < 0$ we use the following process:

1. Estimate the number of dials from mr1

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

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 we assume meter index rolled over and use $tmp_1$. Otherwise we leave the calculated volume as negative and reject the meter point.

#### 6.1.2.2 Prime and Sub Meter Post Processing

There are four cases to consider:

- If the prime is DM we don't need to do anything as we won't have calculated a consumption for the prime so there can be no double-counting.
- If the prime is NDM and contains DM subs then we will fail the prime and replace using an EUC average consumption (see 6.1.3).
- If the consumption calculation fails for any of the subs then fail the prime and replace using an EUC average consumption.
- If the consumption calculation succeeds for the prime and all of its subs then subtract the subs' consumption from the prime.

### 6.1.3 Aggregation and Scaling-Up

When applied to each meter point in any given LDZ, the algorithm outputs a set of consumptions which 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 AUGE based on the AQ.

So for each EUC band we can calculate

1. The number of meter points with a successfully calculated consumption.
2. The number of meter points for which we do not have a calculated consumption (i.e. failed calculation)
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) we restrict attention to consuming meters only, in order to account for potential differences in the proportion of non-consuming meters within and outside the sample.
- Meters 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 meters 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 max 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. Therefore we use the aggregate current AQ as the best available estimate of consumption.
- Where the sample size for a particular EUC for a given LDZ and formula year is less than 30 the national average is used in place of the LDZ average.
- Failed meters which were only active for part of the year are assigned an average demand scaled based on the WAALPs for that part of the year.

Figure 8 below summarises the process for obtaining a consumption value for each meter point.

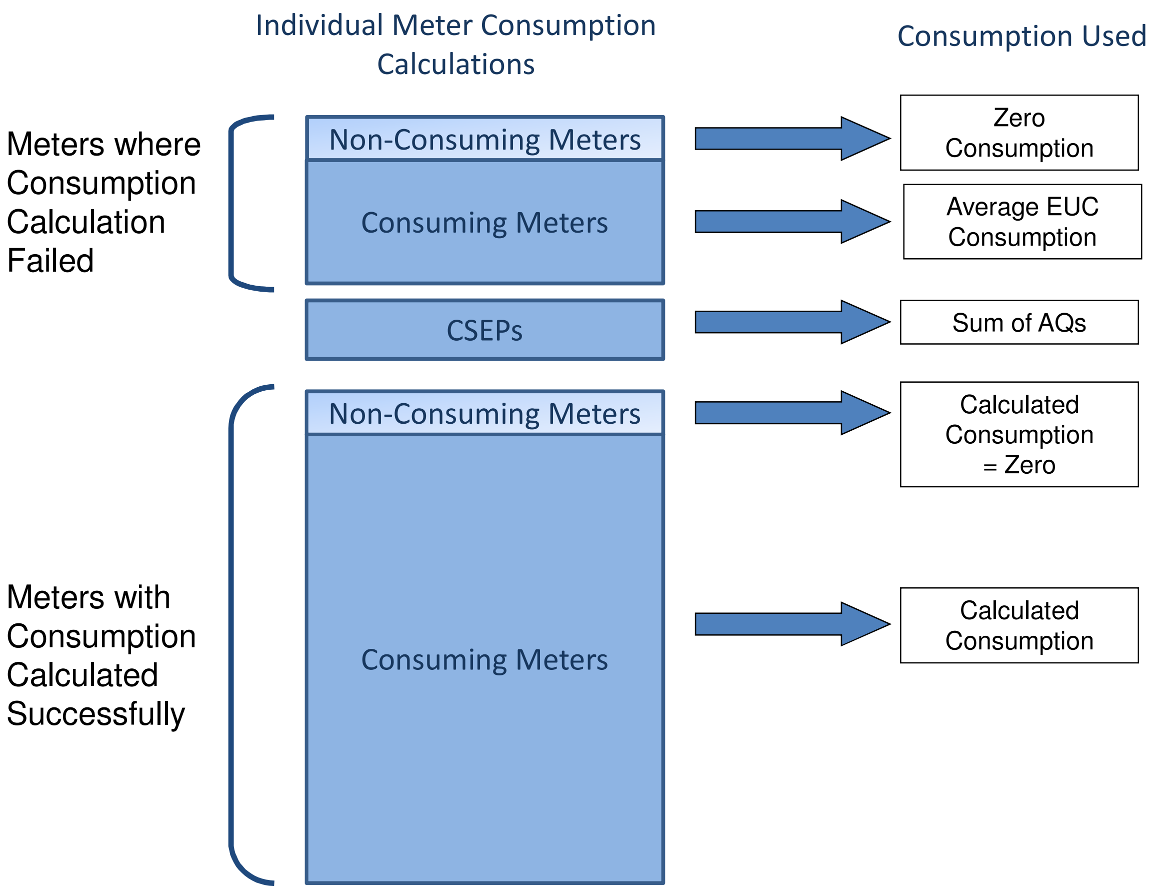


Figure 8: Consumption Method for Each Meter Point

UG for the LDZ for the formula 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 include both permanent and temporary UG, and are not corrected for either meter errors or detected theft. Therefore, whilst giving an indication of the order of magnitude of the UG total for that year, this is simply a step in the calculation process and not an estimate of the final value.

## 6.2 Known DM and LDZ Metering Errors

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 into 6-month periods so that the correct proportion of each meter error can be assigned to each formula year in which the error is active. An example of the data is given in Table 20 below.

Table 20: Sample Meter Error Data

| Billing Month | LDZ | Aggregate Energy (kWh) | Reason | 01/10/04 - 31/03/05 (kWh) | 01/04/05 - 30/09/05 (kWh) | 01/10/05 - 31/03/06 (kWh) | 01/04/06 - 30/09/06 (kWh) |
|---|---|---|---|---|---|---|---|
| May-06 | EM | 41,990,049 | 1 Large consumption adjustment. | 1,104 | 41,987,825 | 1,120 | 0 |
| May-06 | NE | -17,666,209 | 1 Large credit & 1 debit consumption adjustment. | -21,318,352 | 0 | 3,652,143 | 0 |
| May-06 | SC | -57,390,483 | 2 Large credits & 1 debit consumption adjustment. | -47,880,657 | -11,514,298 | 2,004,472 | 0 |
| May-06 | SE | 10,298,400 | 1 Large consumption adjustment. | 0 | 1,593,800 | 8,704,600 | 0 |

These adjustments are therefore applied to the Unidentified Gas calculation after the consumptions have been calculated and aggregated to EUC level. Given that for LSP load consumption corrections are already included in the calculations and for SSP load the AUGE's own validation procedure mimics this process, this leaves the following meter errors as relevant to the UG calculation:

- LDZ metering errors
- DM site errors
- Unique site errors

When considering the high-level consumption method UG equation

*Total UG = Aggregate LDZ Load – DM Load – Shrinkage – (Metered SSP + Metered LSP)*

these errors affect the Aggregate LDZ Load and the DM Load, i.e. the total figure from which metered consumption is subtracted to leave UG. Using the other form of the UG equation

*Total UG = (Alloc SSP + Alloc LSP) – (Metered SSP + Metered LSP)*

it can be seen that this total is calculated using allocations. Therefore, the three types of meter error adjustment listed above are applied to the allocation total, which is calculated at the formula year level of granularity. Corrections for

- 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

## 6.3 Permanent and Temporary Unidentified Gas

The correction for temporary UG is applied on a formula year by formula year basis after the initial total UG figure has been calculated (including the meter read corrections described above). As described in Section 3.3, Temporary UG can exist in the following categories:

- iGT CSEPs (for LSP sites only)
- Shipperless Sites: See Mod 0424
- Unregistered Sites: Shipper Activity, Orphaned and Unregistered <12 Months (if the Shipper carries out site works, or if asset and shipper meter reads match)
- Theft (detected theft only)

Unidentified Gas as calculated using the Consumption Method includes both permanent and temporary UG, and so the temporary element is removed as the final step of the "Total UG" calculation. The method for calculating the temporary element of the categories of UG listed above is defined in detail in the 2011 AUGS for 2012/13 [10] and the only change since this time is an additional adjustment to the Unregistered sites calculation to account for the consistent initial AQ overstatement. This is described in detail in Section 6.5 below.

The figures calculated for Temporary UG using this method are therefore deducted from the (meter error adjusted) total UG figure to give the final permanent UG total. This figure then feeds into the remainder of the UG calculations, where the total is split into its component parts and also split by market sector.

## 6.4 Shrinkage Error

Shrinkage Error is not strictly a component of UG, and hence no attempt is made to estimate it directly. Any residual effects of Shrinkage on the UG estimate (such as long-term bias in the Shrinkage models), should they exist, are automatically included in the UG calculation via the Balancing Factor.

Full details of the AUGE's assessment of Shrinkage can be found in Section 6.4 of the 2011 AUGS for 2012/13 [10].

## 6.5 Unregistered and Shipperless Sites

The analysis for this element of UG has been updated for the current year. For completeness, a full description of the calculation method, including the new elements, is given.

Raw data for all categories of Shipperless/Unregistered UG except "Without a Shipper <12 Months" is contained in snapshot files supplied by Xoserve every two months. In addition to the summarised data in these files, details of each individual MPRN that contributes to the summary data are also supplied.

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

- Orphaned Sites with Opening Meter Reading
  This is used to calculate the proportion of Unregistered sites with meters that flow gas before they are registered (i.e. those that have a non-zero opening meter read).
- Connection Details for Unregistered Sites
  This is used to calculate the proportion of Unregistered sites that can be backbilled for gas consumed before registration. This can only be done if the confirming Shipper is the same as the Shipper that carried out site works.
- Gas Safety Regulations Visit Details
  The gas safety visit data is used to estimate the number and AQ of sites that have been Shipperless for less than 12 months and hence do not yet appear in the snapshots as "Shipperless PTS" or "Shipperless SSrP", but are nevertheless still consuming Shipperless gas.

Further details of these data files are given in Section 5.4 above.

A flowchart of the calculation process for Shipperless and Unregistered UG is shown in Figure 9. A step-by-step procedure for the calculation of the UG estimates follows.

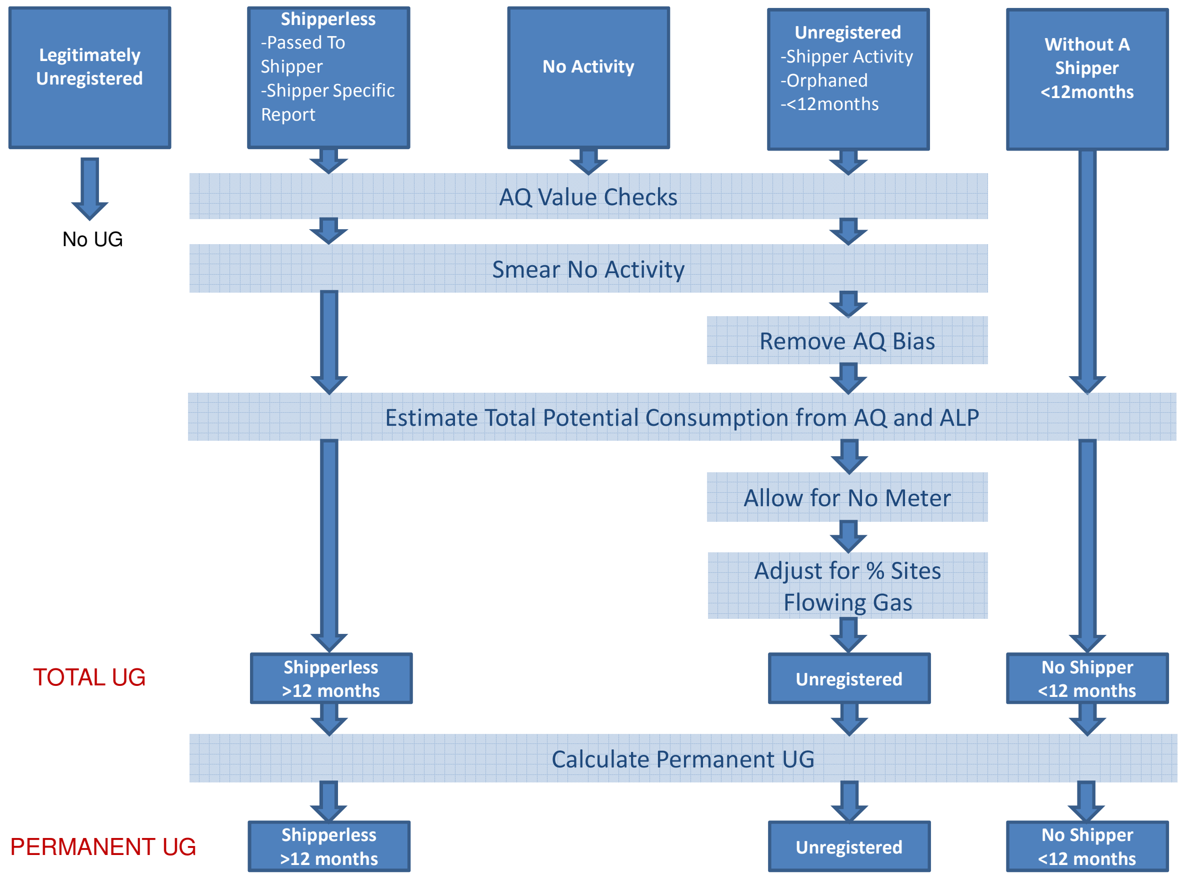


Figure 9: Shipperless and Unregistered UG Calculation Process

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

1. In the backup files containing data for each individual MPRN, each is assessed and flagged for further investigation by Xoserve 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 LSP AQ is flagged.
   - Any DM site (i.e. with an AQ greater than 58.6 GWh) is flagged.

   The resultant list of flagged sites is sent to Xoserve.

2. Xoserve will 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 aggregated to LDZ level for each category of Shipperless or Unregistered site for each snapshot. The data in the relevant snapshot file is then amended to account for these differences. Seven consecutive two-monthly snapshot files are required to calculate the Shipperless and Unregistered UG for a year. Sites where Xoserve have no further information are left as is.

3. Before the analysis is run, the following coefficients are also updated if new data is available.
   - Fraction of opening meter reads with gas flow
   - Fraction of UG not backbilled

4. “Fraction of opening meter reads with gas flow” is calculated using the “Orphaned Sites with Opening Meter Read” spreadsheet. This file contains a list of Orphaned meters and includes their opening meter reading. A meter is defined as having no gas flow whilst Unregistered if its opening reading is either zero or close to the meter index zero (e.g. 99999). If the meter read does not satisfy either of these conditions it is defined as having gas flow whilst Unregistered. The number of meters with gas flow is expressed as a proportion of the total number of meters in the sample. This proportion is applied to the total number of sites with meters to give an estimate of the sites that are actually flowing gas in the Unregistered UG calculations.

5. “Fraction of UG not backbilled” is calculated using the “Connection Details for Unregistered” spreadsheet. This contains sets of meter readings for new confirmed LSP sites. For this analysis only the opening meter reading is required for each site, and so the remainder are discarded. For each meter, two variables are now defined:

   - Gas Flow: if gas flow has occurred at the site before confirmation (i.e. the asset meter read is different from the confirmation meter read) this is set to 1, otherwise 0.
   - Gas Flow with Different Shipper: if gas flow has occurred and the confirming Shipper is not the same as the asset Shipper, this is set to 1, otherwise 0.

   The number of sites that have gas flow with a different Shipper is expressed as a proportion of the total number of sites with gas flow, and this proportion is used to split each type of Unregistered UG calculations into Permanent and Temporary elements.

6. Once the data has been validated and updated where necessary, the first step in the calculation process is to smear the “No Activity” data. The AQ for this category is divided between all other categories in proportion to their relative AQs (except Legitimately Unregistered sites, which do not contribute to UG). The No Activity category plays no further part in calculations because the UG from these sites will be calculated as part of the remaining categories.

7. The raw Shipperless/Unregistered UG calculations are now carried out. The calculations are carried out using VBA code contained in spreadsheets previously supplied to Shippers for their perusal. A formula-based version of the calculation is also available where the individual calculation steps can be followed and verified for a single LDZ at a time. Note that for Orphaned Sites, Shipper Activity and Unregistered <12 Months, both the total UG (including that which will subsequently be backbilled) and the Permanent UG are calculated. The difference between these figures is the Temporary UG from this source and this is used to adjust the UG total as described in Section 6.3 above. Before this adjustment is carried out, both total and Permanent UG figures are modified as described in Step 9 below.

   The total consumption for Shipperless and Unregistered sites is first estimated using AQ data from the snapshot files (amended as described above). To do this, the most recent seven bi-monthly snapshots are used. Seven snapshots are required to cover a full year because each two month period of consumption is calculated from the average aggregate AQ across two snapshots for any given Shipperless/Unregistered UG category. Each of these averages is then multiplied by a factor based on the sum of the ALP over that two month period, with this factor normalised such that the sum of the factors over the six periods is equal to one. The estimate of total annual consumption is therefore given by

$$Consumption = \sum (AQ_m - AQ_{m-2}) \times \frac{P_m}{2}$$

where

$AQ_m$ = Aggregate AQ from snapshot for month $m$
$AQ_{m-2}$ = Aggregate AQ from snapshot for month $m$-2, i.e. the previous snapshot
$P_m$ = Normalised Profile Factor for month $m$ calculated as

$$P_m = \frac{\sum ALP_d(2\,month\,period)}{\sum ALP_d(full\,year)}$$

where

$ALP_d$ = Value of Annual Load Profile for day $d$

8. 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.

9. 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 4 above.

10. The UG estimate for each type of Unregistered site is split into Permanent and Temporary elements based on whether the site will be backbilled or not, as described in Step 5 above.

11. For past iterations of this calculation, the resultant UG values were used without further modification in the AUGS. Factors are now 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*

12. The UG estimates produced in Step 6 above are therefore multiplied by the appropriate combination of these factors. This is done as follows:
    - Shipperless sites (PTS, SSrP): no adjustment
    - Unregistered (Orphaned, Shipper Activity and Unregistered <12 Months): adjust using composite *Factor1(n) x Factor2*.

13. Following the implementation of Mod 0424, Shipperless PTS Unidentified Gas is no longer all Permanent. UG from sites that became Shipperless before the implementation of the modification on 25th January 2013 will remain Permanent, but sites becoming Shipperless after that date will be backbilled and hence UG arising from them is Temporary. The AUGE has requested that Xoserve start to provide a flag in the MPRN details that shows whether each site became Shipperless before or after

the threshold date, and this will be used to divide the UG from this component into Permanent and Temporary elements. The Temporary element will then be used to adjust the UG total as described in Section 6.3 above.

At the time of writing, all sites that appear in the snapshot files will by definition have become Shipperless before the threshold date because they only appear in the file when they have been Shipperless for 12 months or more. The AUGE is required to forecast UG for 2014/15, however, when sites that became Shipperless after 25/01/2013 *will* qualify for inclusion. In addition, the "Without a Shipper <12 Months" category will also be affected because it is composed of sites that will end up in both the PTS and SSrP categories. UG from the SSrP sites will remain Permanent in 2014/15, but that from the PTS sites will be Temporary because in this year, any site that has been without a Shipper for less than 12 months will have become Shipperless after 25/01/2013.

Therefore, it will be necessary to estimate the impact of the 25/01/2013 changeover on future UG. This can be done by establishing the level of "churn" in the Shipperless PTS dataset over time – i.e. the proportion of sites from one snapshot that disappear from the dataset (either by having their meter removed or by being reconfirmed) in subsequent snapshots. The trend in churn over a period of N snapshots can be estimated using the backup MPR data supplied by Xoserve for the full snapshot history currently available.

The statistic plotted in this analysis is the proportion of sites (in terms of number and AQ) that appeared in the Shipperless PTS category in Snapshot X also appear in Snapshot X+1. This value is calculated over a number of successive snapshots (i.e. from Snapshot X to Snapshot X+N, with the proportion of Snapshot X sites still Shipperless in Snapshot X+N always reported). These figures are used to derive a relationship between N and the proportion of sites from Snapshot X still remaining Shipperless. This relationship is then used to derive the Permanent/Temporary split of this UG category by applying it to the January 2014 snapshot, which will be the last that contains fully Permanent Shipperless PTS UG. No more Permanent UG from this source can be added after this time because the next snapshot, Mar 2014, will lie more than a year after the Mod 0424 threshold date, and so all new sites appearing as Shipperless PTS will, by definition, be Temporary.

The decay in the Permanent UG from the January 2014 snapshot is modelled using the churn equation, and the difference between the total UG in each snapshot and the remaining Permanent UG from the churn equation applied to the Jan 2014 snapshot is the Temporary UG from this category.

Analysis of the churn statistics shows that there is a steady decrease in the proportion of remaining sites for the first seven snapshots, after which only long-term Shipperless sites remain and the proportion does not decrease any further. This is illustrated for the SSP market sector in Figure 10 below, which shows the actual and fitted values for the proportion of AQ remaining after 9 snapshots.

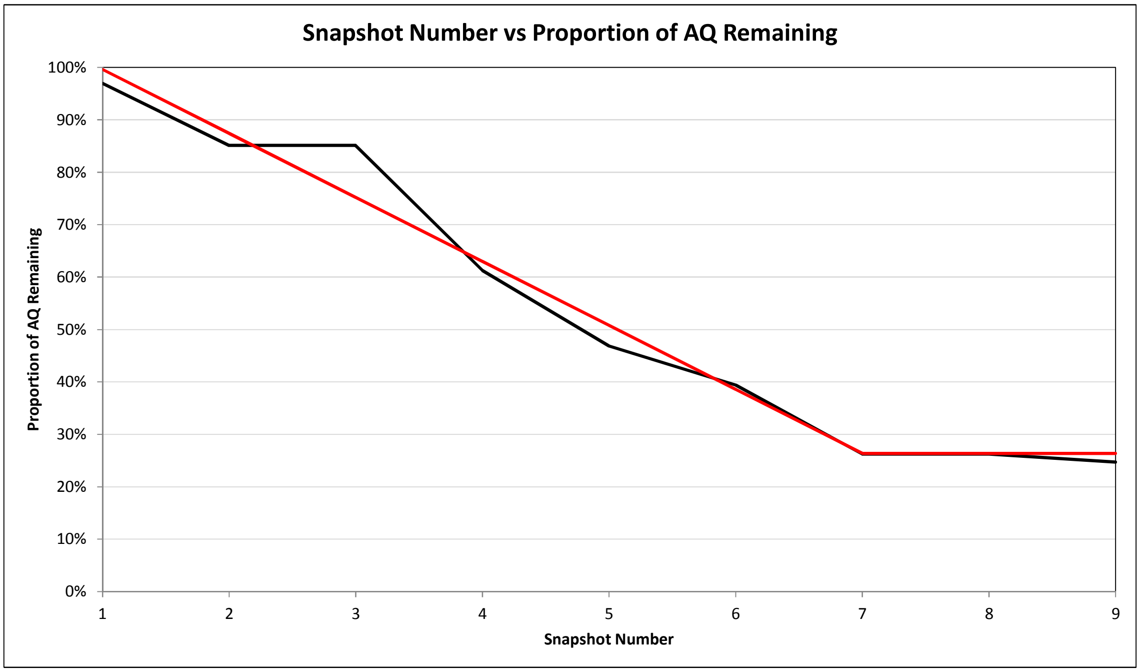


Figure 10: Shipperless PTS Churn

In order for this relationship to be accurately implemented taking into account the annual profile in consumption, calculations based on this equation are implemented in the AUGE's Shipperless Contribution Calculator model. For each bi-monthly period of the year being forecasted, the percentage remaining AQ of sites contributing to Permanent Shipperless PTS UG is calculated, and this is multiplied by the normalised annual profile value for the equivalent time period, taken from the ALPs for the relevant market sector. The resultant values are multiplied by the PTS UG total for the last fully-Permanent snapshot (i.e. January 2014: when carrying out calculations before this date, values for this snapshot must therefore be estimated). The final result of this calculation is the remaining Permanent Shipperless PTS UG.

- Total Shipperless PTS UG is calculated from the snapshots in the usual way. This, as always, involves calculating the UG for the most recent time period for which we have snapshots and assuming that the process remains steady, and this is therefore the best estimate for the UG year being forecasted.
- Permanent UG is subtracted from Total UG (both calculated as defined above), with the remainder being Temporary UG.
- Permanent and Temporary UG thus calculated are then treated as appropriate in the overall UG calculations.

14. The final figures produced as described above are now used to populate the AUGS for all categories except "Without a Shipper <12 Months".

15. "Without a Shipper <12 Months" UG is calculated using Gas Safety Regulations visit data contained in the "GSR Passed To Shipper" spreadsheet. 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 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 Shipperless for less than 12 months and hence make up the "Without a Shipper <12 Months" UG category for this year.

Therefore, in order to estimate the UG from this category, the AQs from the gas safety visit data are aggregated by LDZ and SSP/LSP split. Given that the sites in question will have been becoming Shipperless at a steady rate throughout the year, they will on average have been Shipperless for 6 months each. Therefore each aggregate AQ total is divided by 2 to give the final total UG estimate for this category. This figure then needs to be split into PTS and SSrP components.

All sites that fall into the "Without a Shipper <12 Months" category will end up either as Shipperless PTS or Shipperless SSrP if they remain Shipperless for more than a year. Under the terms of Mod 0424, any UG from Shipperless PTS sites that became Shipperless after 25/01/13 is Temporary and hence should be removed from the "Without a Shipper <12 Months" Permanent UG total. The Gas Safety Regulations visit data does not group sites into those with an existing meter and those with a new meter, and so the split of "Without a Shipper <12 Months" UG into PTS and SSrP categories is done using ratios obtained from the Xoserve snapshot data. PTS UG estimated in this way is classified as Temporary, whilst SSrP UG is classified as Permanent.

The current AUGS provides UG estimates for the 2014/15 formula year. During this year, any site that has been Shipperless for less than 12 months must, by definition, have become Shipperless after 25/01/13, and therefore UG from all sites that would go on to become PTS is Temporary. Therefore the Temporary/Permanent split for "Without a Shipper <12 Months" UG is the same as the split between PTS and SSrP, and no further modification is necessary.

## 6.6 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 loads within them or entire iGT networks are not recognised by the Xoserve system and are thus taking gas in an unrecorded manner.

### 6.6.1 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 Xoserve's systems. Regular meetings are held between the iGTs and Xoserve in order to resolve these issues and reduce the number of Unknown Projects.

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

For both these sources of UG from CSEPs, that from LSP sites is backbilled and is hence **temporary**. UG from SSP sites on CSEPs is not backbilled and is hence **permanent**.

It is necessary to calculate both the permanent and temporary elements of iGT CSEPs UG so that the raw UG total calculated using the Consumption Method can be reduced to account for the temporary element, leaving the final total as permanent UG only.

### 6.6.2 Data

*Unknown Projects* data is supplied by Xoserve in bi-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 Xoserve. 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 SSP/LSP CSEP throughput percentages 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 Xoserve have been notified that the supply point is Unregistered rather than the number of sites that are actually Unregistered. Xoserve 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.

### 6.6.3 Process

Processing is done in spreadsheets, which are supplied to the industry to allow auditing of the AUGE'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 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 market sector (SSP/LSP) and by LDZ is calculated for registered sites. These figures are then used to calculate the percentage split of CSEP site count by market sector and the percentage split of CSEP AQ by market sector. This split is used in the calculations for Unknown Projects, described below.

4. The total site count and AQ by market sector (SSP/LSP) and by LDZ is calculated for Unregistered sites. The AQ figures produced 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 20 below. In these tables, the “Year” field refers to the year in which the CSEP came to the attention of Xoserve. 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.

Table 21: Unknown Projects Snapshot

| YEAR | LDZ | Count of Unknown Projects | Sum Of AQ | Count of Supply Points |
|---|---|---|---|---|
| 2012 | EA | 39 | 4,606,764 | 361 |
| 2012 | EM | 40 | 9,717,609 | 609 |
| 2012 | NE | 4 | 680,926 | 27 |
| 2012 | NO | 6 | 1,981,423 | 27 |
| 2012 | NT | 50 | 8,082,751 | 673 |
| 2012 | NW | 33 | 3,610,353 | 251 |
| 2012 | SC | 16 | 24,549,597 | 131 |
| 2012 | SE | 19 | 3,578,899 | 173 |
| 2012 | SO | 18 | 1,448,603 | 114 |
| 2012 | SW | 21 | 3,912,544 | 306 |
| 2012 | WM | 37 | 6,090,727 | 467 |
| 2012 | WN | 5 | 2,755,000 | 31 |
| 2012 | WS | 13 | 2,482,641 | 199 |
| | Total | 301 | 73,497,837 | 3,369 |

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 SSP and LSP market sectors 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 market sector (in terms of supply point count and AQ), with this composition again calculated from registered sites on known CSEPs.

8. The total SSP AQ and LSP AQ across all LDZs plus Unknown LDZ is calculated. These figures represent the best estimate of annual consumption in Unknown Projects at the time the snapshot was produced.

9. 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.

10. The above process gives, for each snapshot, 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 11 below.

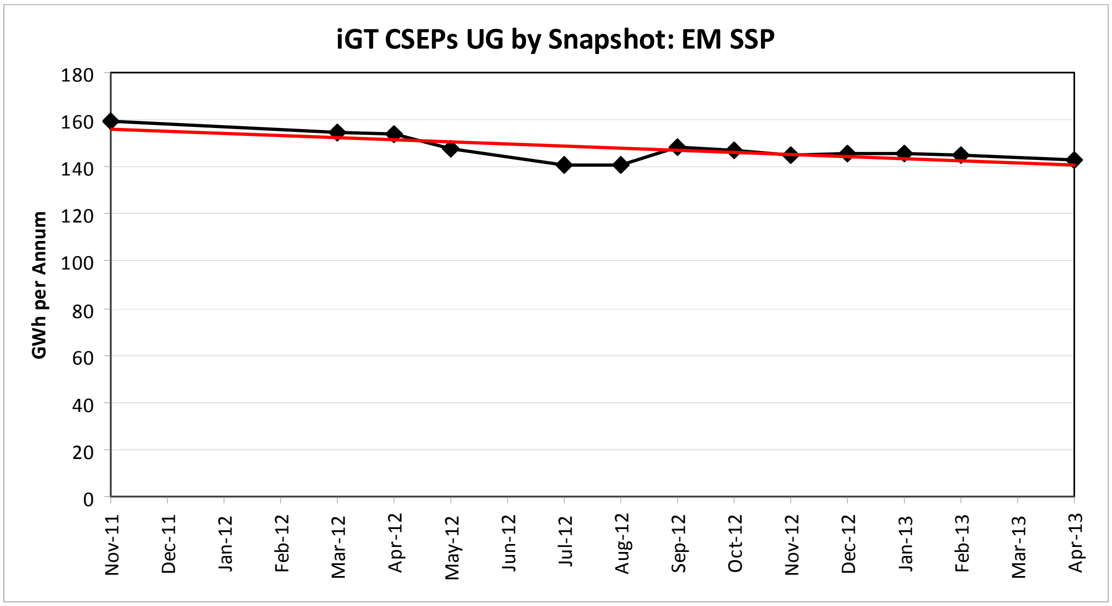


Figure 11: iGT CSEPs UG by Snapshot

11. The identified snapshot-to-snapshot trend for each LDZ/market sector combination is used to extrapolate either forwards or backwards to each time period of interest. The UG for each time period used in the analysis is calculated using the (estimated) values from all the snapshots that fall 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 for the final UG figures, 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 2009/10 to 2011/12).

12. The UG calculated in this way from LSP sites is **temporary**. The UG calculated in this way from SSP sites is **permanent**. The final figures are therefore carried forward into the overall UG calculations and used as appropriate.

## 6.7 LSP and SSP Metering Errors

The analysis for this element of the UG calculation remains the same as described in the 2011 AUGS for 2012/13 [10] and is included here with reference to the Consumption Method for completeness. The figures will be updated based on the latest data when the final UG estimates for 2014/15 are produced.

The effects of LDZ metering errors and known DM/Unique Site supply point errors are discussed in Section 6.2 above. In addition, errors in SSP and NDM LSP 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.

The GL Noble Denton Metering Team were asked to provide input for the analyses identified above, and the conclusions drawn in this section come from them. The following conclusions were drawn from the investigation:

- 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 (i.e. SSP loads and smaller LSP loads) 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 meter drift to be carried out, a large amount of data would have to be collected via a national meter survey similar to that conducted many years ago, as noted in the AUGE's responses to comments on the second draft of the 2012 AUGS for 2013/14 [20]. 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 6.2. Data regarding such adjustments is supplied to the AUGE by Xoserve 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 an NDM LSP RPD meter is shown in Figure 12 below.

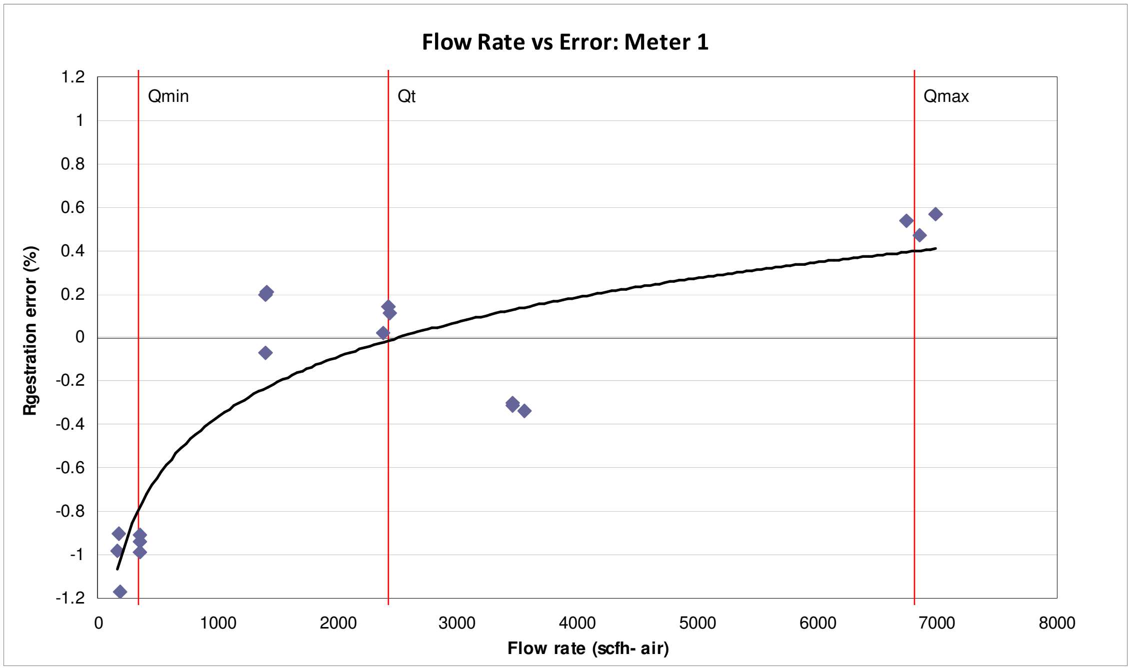


Figure 12: Typical Calibration Curve for an RPD Meter

Data for this graph was provided by the GL Noble Denton 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 was limited to the meter capacity and AQ of each LSP site, and this required the AQ to be used to estimate average hourly load, which could 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.

- Sites with an average hourly flow of less than 1% of meter capacity were considered to be likely to be operating at or around $Q_{min}$ when gas was flowing. These were 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 were considered to be likely to be operating at or around $Q_{max}$ when gas was flowing. These were 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 then represents the net over- or under-read in the population.

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

## 6.8 Shipper Responsible Theft

As described in Section 4.7, undetected theft (which forms the vast majority of the Balancing Factor) will be calculated using the Throughput Method. This is a very simple method that splits this element of UG in the same proportion as SSP/NDM LSP throughput. This has a number of advantages over other methods of splitting theft between market sectors, as follows:

- This method acts as an incentive to reduce theft as it removes the situation where detecting a theft would increase the theft split percentage for that market sector. Instead, prevention and detection of theft will reduce the total UG figure, which in turn will result in a lower residual figure for the Balancing Factor. This will result in a lower figure of UG in each sector.
- It is simple and transparent to calculate.
- It cannot be manipulated or affected by different detection rates.
- It does not rely on estimates of theft and estimates of periods of theft.
- It does not rely on estimation of AQs and correct sector assignment of detected thefts which is fundamental to any theft split method.
- Other elements of the Balancing Factor (i.e. those elements bundled in with theft) are also apportioned by throughput.
- Issues concerning treatment of unregistered theft-affected sites and use of pre/post theft AQ are removed.
- The risk of large changes in theft split percentages year on year is reduced.

Going forward, the market sector split for theft will not be calculated using allocations from the ODR report. Instead, the split will be calculated using the seasonal normal adjusted consumption values calculated to estimate UG. This has the advantage that the effects of UG can be correctly allocated by market sector and makes no assumption about the statistical properties of RbD.

$T^{LSP}$ = the ratio of LSP seasonal normal throughput to the total seasonal normal throughput defined below.

$$T^{LSP} = \frac{Tput_{SN}^{LSP}}{\left(Tput_{SN}^{LSP} + Tput_{SN}^{SSP}\right)}$$

where $Tput_{SN}^{LSP}$ and $Tput_{SN}^{SSP}$ are the total LSP and SSP throughputs calculated over the same 3 formula years as the total UG, corrected to seasonal normal conditions. These values also include any directly calculated UG associated with the relevant market sector. $Tput_{SN}^{LSP}$ is calculated as follows (and $Tput_{SN}^{SSP}$ is calculated in a similar way):

$$Tput_{SN}^{LSP} = \sum_{LDZ=1}^{13}\sum_{yr=1}^{3}\left(SNCons_{LDZ,yr}^{LSP} + UGPerm_{LDZ,yr}^{LSP}\right)$$

where

$SNCons_{LDZ,yr}^{LSP}$ = Calculated seasonal normal annual consumption for all LSP meters

$UGPerm_{LDZ,yr}^{LSP}$ = Total permanent directly calculated component of UG for all LSP meters (includes shipperless sites, unregistered sites and iGT CSEPs) occurring over the relevant year.

The resulting factor ($T^{LSP}$) is used when calculating the LSP permanent UG as described in Section 6.10.

## 6.9 DM LSP Market Sector

In the 2011 AUGS for 2012/13, the UG attributed to DM LSP sites was concluded to be negligible. This is based on the following assumptions:

- There is no theft from DM sites.
- Any Unregistered DM sites are backbilled.
- DM sites do not become Shipperless.
- There are no unknown DM sites.

In addition, it is known that DM sites on unknown CSEPs will be backbilled because this applies to all LSP sites on CSEPs.

This leaves only unknown meter error for DM sites, and as described in the Worked Example in Section 6.11 below, current data indicates that there is little or no over-read on DM sites due to meters working at the very low end of their range.

Updated data received during the course of the current formula year will be reviewed to confirm whether these assumptions still hold or not. At the time of this draft DM LSP UG is concluded to be negligible.

## 6.10 Aggregation of Final National UG Figure

Although the analysis of UG has been carried out on an LDZ by LDZ basis, the final national figure will be based on the combined total.

For 2014/15 this will be based on the total LDZ consumptions for formula years 2009-2011 subtracted from their corresponding total NDM allocations with corrections for meter error and temporary UG which includes detected theft.

The resulting figure will be averaged over the years used, and split into the LSP and SSP sectors.

The total permanent UG for the LSP sector ($FUG_{LSP}$) is calculated by adding the directly calculated permanent UG for the LSP sector to the LSP portion of the Balancing Factor.

$$FUG_{LSP} = \frac{1}{3}\sum_{LDZ=1}^{13}\sum_{yr=1}^{3}\left(BF_{LDZ,yr}^{LSP} + DCPUG_{LDZ,yr}^{LSP}\right)$$

where,

$DCPUG_{LDZ,yr}^{LSP}$ = Directly Calculated Permanent UG for LSP by LDZ, year

$BF_{LDZ,yr}^{LSP}$ = Balancing Factor for the LSP sector by LDZ, year as defined below

$$BF_{LDZ,yr}^{LSP} = BF_{LDZ,yr}^{All} \times T^{LSP}$$

where,

$T^{LSP}$ = the ratio of LSP seasonal normal throughput to the total seasonal normal throughput as defined in Section 6.8.

$BF_{LDZ,yr}^{All}$ = Total Balancing Factor quantity for a given LDZ, year as defined below,

$$BF_{LDZ,yr}^{All} = Alloc_{LDZ,yr}^{SSP} + Alloc_{LDZ,yr}^{LSP} + LDZcorr_{LDZ,yr} + DMcorr_{LDZ,yr} - Cons_{LDZ,yr}^{LSP} - Cons_{LDZ,yr}^{SSP} - UG_{LDZ,yr}^{Temp}$$

where,

$Alloc_{LDZ,yr}^{SSP}$ = SSP Allocations for each LDZ for each year used

$Alloc_{LDZ,yr}^{SSP}$ = LSP Allocations for each LDZ for each year used

$LDZcorr_{LDZ,yr}$ = Total LDZ level meter corrections per LDZ per year

$DMcorr_{LDZ,yr}$ = Total DM and Unique site meter corrections per LDZ per year

$Cons_{LDZ,yr}^{SSP}$ = Calculated total SSP consumptions per LDZ per year

$Cons_{LDZ,yr}^{LSP}$ = Calculated total LSP consumptions per LDZ per year

$UG_{LDZ,yr}^{Temp}$ = Total temporary UG by LDZ by year from the direct calculated UG process including total detected theft

## 6.11 Worked Example

In order to illustrate how the above techniques are applied in practice, the following worked example is provided for an unspecified LDZ (referred to as XX LDZ) and year. This shows how each element of UG is calculated and how it contributes to the final total. The values used throughout this example are for illustrative purposes only and do not relate to real figures from any LDZ.

The UG calculation takes places in stages, as follows:

1. Calculation of total UG using the consumption method. At this stage this includes both permanent and temporary UG and in the 2012 AUGS was calculated over the formula years 2009 to 2010 and seasonally adjusted. The calculation is carried out as described in Section 6.1 above, and for XX LDZ the total calculated UG is 550.0 GWh.

2. The temporary UG total is now calculated for the categories of UG listed in the table in Section 3.3 and calculated as described in the 2011 AUGS for 2012/13 [10]. This is deducted from the total UG figure

calculated in Step 1 above to give the total permanent UG. In this example, temporary UG totals 50.0 GWh and hence:

Permanent UG = Total UG – Temporary UG = 550 GWh – 50 GWh = 500 GWh

3. The next stage of the process is to calculate the directly estimated components of UG. This is done separately for SSP and LSP, thereby giving a breakdown by market sector as well as the total for each component.

4. The iGT CSEPs calculation is based on data provided by Xoserve in the Unknown Projects Summary, along with information about live and Unregistered sites on known CSEPs. Figures are as follows for XX LDZ:

   Unknown Projects = 100
   Supply Point Count = 1305
   AQ Total = 18.0 GWh

   From known CSEPs in XX LDZ:

   SSP Supply Point proportion = 99.5%
   LSP Supply Point proportion = 0.5%
   SSP AQ proportion = 84.0%
   LSP AQ proportion = 16.0%

   These figures are used to split the unknown project supply point count and aggregate AQ by market sector:

   For unknown projects:

   SSP Supply Points = 1299
   LSP Supply Points = 6
   SSP AQ = 15.0 GWh
   LSP AQ = 3.0 GWh

   Data regarding Unregistered sites on known CSEPs is supplied by Xoserve and is as follows:

   SSP Supply Points = 3000
   LSP Supply Points = 10
   SSP AQ = 45.0 GWh
   LSP AQ = 0.1 GWh

   Total UG from this source is the combination of these two, plus a proportion of 10 unknown projects with unknown LDZ smeared across all LDZs:

   SSP Supply Points = 4400
   LSP Supply Points = 16
   SSP UG = 62.0 GWh

LSP UG = 3.5 GWh

Note that the LSP UG calculated here is temporary in nature and forms part of the 50 GWh subtracted from the initial total UG in Step 2. It is therefore not taken further into the final UG categorisation. The SSP UG is permanent and is taken forwards.

5. Shipperless and Unregistered sites are split into six categories. Calculations for each category are very similar, so a single typical example - LSP Shipper Activity Sites - is given here.

   Site count and AQ data is supplied in the two-monthly snapshot files. Figures for XX LDZ are:

   Snapshot 1 AQ: 2.6 GWh
   Snapshot 2 AQ: 3.2 GWh
   Snapshot 3 AQ: 3.0 GWh
   Snapshot 4 AQ: 3.2 GWh
   Snapshot 5 AQ: 2.8 GWh
   Snapshot 6 AQ: 3.0 GWh
   Snapshot 7 AQ: 2.9 GWh

   The gas consumed between snapshot x and snapshot y is calculated as the average AQ across these two snapshots, multiplied by the appropriate factor from Table 7 in the 2011 AUGS for 2012/13 [10] to reflect the time of year:

   Snapshots 1-2: Average AQ = 2.9 GWh
   Time of year factor = 0.065
   Percentage of orphaned/shipper activity sites with non-zero opening reads = 36.8%
   Percentage of occurrences that are not backbilled = 31.25%
   Permanent UG = 2.9 GWh * 0.065 * 36.8% * 31.25% = 21,678 kWh

   Similar calculations for the remaining snapshots give the following consumptions:

   Snapshot 1-2: 21,678 kWh
   Snapshot 2-3: 24,955 kWh
   Snapshot 3-4: 65,205 kWh
   Snapshot 4-5: 96,600 kWh
   Snapshot 5-6: 86,250 kWh
   Snapshot 6-7: 50,370 kWh
   Total: 0.35 GWh

   Calculations for each other category of Shipperless or Unregistered site are similar. The final totals of permanent UG across all categories of Shipperless/Unregistered sites for LDZ XX are:

   SSP UG = 10.6 GWh
   LSP UG = 75.5 GWh

6. For meter errors, sites with an average hourly consumption (calculated from the AQ) of 1% or less of their $Q_{max}$ value are considered to be consistently operating in the "under-read" area. Sites with an average hourly consumption of 95% or more of their $Q_{max}$ value are considered to be consistently

operating in the “over-read” area. The average levels of under-read and over-read are taken from calibration curves, as described in detail in the 2011 AUGS for 2012/13 [10].

Average under-read: 1.5%
Average over-read: 0.5%

Total sites in under-read zone for XX LDZ: 5000
Aggregate under-read: 2.0 GWh

Total sites in over-read zone for XX LDZ: 5
Aggregate over-read: 0.1 GWh

Net contribution to UG: 2.0 GWh – 0.1 GWh = 1.9 GWh

This is the error arising from the NDM LSP market and hence this is where the full 1.9 GWh is applied.

7. The sum of the directly measured UG components calculated in Steps 4-6 above gives the figure for total directly measured permanent UG. The SSP and LSP elements are summed and deducted from the total UG figure (calculated in Step 2 above) to give the total for the Balancing Factor. At this stage the Balancing Factor is a single figure, the sum of SSP and LSP elements.

   Balancing Factor = 500.0 GWh – Total Directly Measured = 350.0 GWh

8. All elements of the Balancing Factor other than Theft are either small or will sum to zero over time. Therefore it is reasonable to split the Balancing Factor volume between the SSP and LSP market sectors using the percentage split for Theft, as defined in Section 6.8 above.

   SSP proportion = 76.7%
   LSP proportion = 23.3%

   For XX LDZ for a single year:

   Total UG = 550.0 GWh
   Temporary UG = 50.0 GWh
   Total Permanent UG = 500.0 GWh
   Directly Measured UG = 150.0 GWh
   Aggregate Balancing Factor = 350.0 GWh

   SSP Balancing Factor = 350.0 * 0.767 = 268.5 GWh
   LSP Balancing Factor = 350.0 * 0.233 = 81.5 GWh

9. Finally, total UG from each sector is calculated by summing the components, values for all of which have now been populated:

   SSP UG = 62.0 GWh + 10.6 GWh + 268.5GWh = 341.1 GWh
   LSP UG = 75.5 GWh + 1.9 GWh + 81.5 GWh = 158.9 GWh

Total UG = 341.1 GWh + 158.9 GWh = 500 GWh

These calculations are then repeated for each LDZ and year to give an estimate over 3 years.

# 7 Unidentified Gas Estimates

The AUG table will be provided separately once the methodology has been approved and will reference the approved statement as appropriate. It will include the estimation of SAP price, estimates of UG by LDZ and source of UG and the final AUG table.

# 8 Consultation Questions and Answers

This section captures a history of the questions raised by industry bodies during the consultation periods and the AUGE's responses. These relate to all drafts of the 2011 AUGS for 2012/13 and the 1st and 2nd draft 2012 AUGS for 2013/14 and the recent 1st draft 2013 AUGS for 2014/15. 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 external documents:

- "AUGS Query Responses 30_09_2011" [18]
- "AUGS Draft2 Query Responses 14_11_2011" [9]
- "AUGS Query Responses 19_03_2012" [11]
- "AUGE Responses to 1st Draft 2012 AUGS" [12]
- "AUGE Responses to Interim Report Consultation" [13]
- "AUGE Responses to 2nd Draft AUGS Consultation 12032013" [20]
- "AUGE Responses to 1st Draft AUGS Consultation 25th June 2013" [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.

Table 22 below contains a list of organisations that responded to the first draft of the 2011 AUGS for 2012/13.

Table 22: Responses to the First Draft of the 2011 AUGS

| Organisation Name | Date of Communication |
|---|---|
| National Grid Transmission | 06/05/2011 |
| Corona Energy | 23/05/2011 |
| E.On | 23/05/2011 |
| British Gas | 15/06/2011 |
| EDF Energy | 16/06/2011 |
| GDF Suez | 16/06/2011 |
| Gazprom | 17/06/2011 |
| ScottishPower | 17/06/2011 |

Table 23 below contains a list of organisations that responded to the second draft of the 2011 AUGS for 2012/13.

Table 23: Responses to the Second Draft of the 2011 AUGS

| Organisation Name | Date of Communication |
|---|---|
| Npower | 31/10/2011 |

| | |
|---|---|
| ICoSS | 31/10/2011 |
| Total Gas and Power | 31/10/2011 |
| ScottishPower | 31/10/2011 |
| British Gas | 31/10/2011 |

Table 24 below contains a list of organisations that responded to the final version of the 2011 AUGS for 2012/13.

Table 24: Responses to the Final Draft of the 2011 AUGS

| Organisation Name | Date of Communication |
|---|---|
| British Gas | 20/02/2012 |
| Inexus | 08/03/2012 |
| Shell Gas Direct | 08/03/2012 |

Table 25 below contains a list of organisations that responded to the first draft of the 2012 AUGS for 2013/14.

Table 25: Responses to the First Draft of the 2012 AUGS

| Organisation Name | Date of Communication |
|---|---|
| Energy UK | 15/06/2012 |
| ScottishPower | 15/06/2012 |
| ICoSS | 29/06/2012 |

Table 26 below contains a list of organisations that responded to the September 2012 Interim Report

Table 26: Responses to the September 2012 Interim Report

| Organisation Name | Date of Communication |
|---|---|
| Energy UK | 28/09/2012 |
| Gazprom | 28/09/2012 |
| Corona Energy | 28/09/2012 |
| Npower | 28/09/2012 |
| Total Gas and Power | 28/09/2012 |

Table 27 below contains a list of organisations that responded to the 2nd Draft of the 2012 AUGS for 2013/14.

Table 27: Responses to the 2nd Draft of the 2012 AUGS

| Organisation Name | Date of Communication |
|---|---|
| ICoSS | 01/03/2013 |
| DONG Energy | 01/03/2013 |
| Energy UK | 01/03/2013 |
| RWEnpower | 01/03/2013 |
| ScottishPower | 01/03/2013 |
| SSE Energy Supply | 01/03/2013 |
| British Gas | 01/03/2013 |

Table 27 below contains a list of organisations that responded to the 1st Draft of the 2013 AUGS for 2014/15.

Table 28: Responses to the 1st Draft of the 2014 AUGS

| Organisation Name | Date of Communication |
|---|---|
| ICoSS | 12/06/2013 |
| DONG Energy | 12/06/2013 |
| RWEnpower | 12/06/2013 |
| Scottishpower | 12/06/2013 |
| British Gas | 12/06/2013 |

# 9 Contact Details

Questions can be raised with the AUGE at AUGE@gl-group.com

# 10 References


[1] Guidelines for the Appointment of an Allocation of Unidentified Gas Expert and the provision of the Allocation of Unidentified Gas Statement V3.0, 24th February 2011

[2] Mod 0194 Framework for correct apportionment of NDM error

[3] Mod 0194a Framework for correct apportionment of LSP unidentified gas

[4] Mod 0228/0228A Correct apportionment of NDM Error – Energy

[5] Mod 0229 Mechanism for Correct Apportionment of Unidentified Gas implemented in UNC Section E 10 v3.54 26th April 2011,

[6] Uniform Network Code (UNC) Transportation Principal Document

[7] Mod 0398: Limitation on Retrospective Invoicing and Invoice Correction Version 6.0, 28 August 2012

[8] AUGS Draft 1 Query Responses, September 2011

[9] AUGS Draft 2 Query Responses, November 2011

[10] AUGS Final (Version 4), December 2011, GL Noble Denton

[11] 2011 AUGS for 2012/13 Query Responses 19_03_2012, March 2012

[12] AUGE Responses to 1st Draft 2012 AUGS for 2013_14 Consultation, July 2012

[13] AUGE Responses to Interim Report Consultation 17102012, October 2012

[14] Modification 0429 Version 4.0, 27 January 2013

[15] Modification 0410 Version 7.0, 09 January 2013

[16] Final Modification Report 0424 Version 4.0, 20 December 2012

[17] Draft Modification Report 0425 Version 1.0, 21 March 2013

[18] AUGS Query Responses 30_09_2011, GL Noble Denton

[19] 2012 AUGS for 2013/14 Final Version, December 2012, GL Noble Denton

[20] AUGE Responses to 2nd Draft AUGS Consultation 12032013, March 2013

[21] Modification 0410A Version 7.0, 12 November 2012

[22] Mod 0398 OFGEM decision letter 1 March 2013

[23] Mod 0424 OFGEM decision letter 24 January 2013

[24] Mod 0282 Introduction of a process to manage Vacant sites 2010/11

[25] Estimation of CSEP Leakage, Roy Malin, National Grid Distribution, 22nd February 2013

[26] SPAA Schedule 33 “Theft of Gas Code of Practice” from version 9.4 of SPAA ,1st April 2013

[27] AUGE Responses to 1st Draft 2013 AUGS Consultation, 25th June 2013

[28] Xoserve Customer Discovery Day – Reconciliation presentation, 22nd November 2012

[29] Tackling Gas Theft: Final Impact Assessment, OFGEM, 26th March 2012

[30] Direction under paragraph 8 of condition 12A of the Standard Conditions of the Gas Supply Licence to introduce the Theft Risk Assessment Service, OFGEM, 7th January 2013

## Glossary

| | |
|---|---|
| AGI | Above Ground Installation |
| ALP | Annual Load Profile (deeming algorithm parameter) |
| AQ | Annual Quantity. An estimate of annual consumption under seasonal normal conditions |
| 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. |
| 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 |
| 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) |
| IGT | Independent Gas Transporter |
| LSP | Larger Supply Point |
| MAM | Meter Asset Manager |
| Model Error | The statistical error associated with any modelling or estimation process. It an inherent part of any statistical model and does not imply that the model itself is inadequate or incorrect. |
| MPRN | Meter Point Reference Number |
| NDM | Non Daily Metered |
| ODR | OFGEM Data Request |
| OUG | Own Use Gas |
| PSND | Pseudo Seasonal Normal Demand calculated using AQ values rather than being based on historic metered demands |
| RbD | Reconciliation by Difference |
| RbD-based Method | The methodology developed and approved in 2011 AUGS for 2012/13 |

| | |
|---|---|
| SF | Scaling Factor (deeming algorithm output) |
| SNCWV | Seasonal Normal Composite Weather Variable |
| SND | Seasonal Normal Demand |
| SSP | Smaller Supply Point |
| TPD | Transportation Principle Document |
| UIP | Utility Infrastructure Provider |
| UNC | Uniform Network Code |
| UG | Unidentified Gas |
| WAALP | Weather Adjusted Annual Load Profile |
| WCF | Weather Correction Factor (deeming algorithm output) |
| WSENS | Weather Sensitivity (deeming algorithm parameter used in EWCF definition. Sensitivity of an EUC to difference in CWV from seasonal normal) |

# Appendix A Raw Data Description

This appendix describes the raw data provided by Xoserve for the consumption method.

## A.1 ALLOCATIONS

This data contains all allocations including CSEPs from 01/04/2008 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 |

## A.2 ANNUAL_QUANTITY

This data includes all meter points active at any point from 01/04/2008 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 |

## A.3 CSEPS

This data contains information for formula year 2008 onwards.

| Name | Description |
|---|---|
| FORMULA_YEAR | Date - Formula 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 formula year |
| COUNT_OF_SUPPLY_POINTS | Number - Count of supply points at start of formula year |

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

## A.4 FACTORS

This data is provided from 1st April 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.) |

## A.5 METER_READS

This data includes all meter reads from 01/04/2008 onwards. Multiple records for a meter point with the same date are filtered by Xoserve using the following methodology.

Where there is an A (Actual) Read Type and an E (Estimate) Read Type Xoserve remove the E and retain the A Read. Where there are Read Types of R (Replacement) Xoserve 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 |

The read type codes are as follows:

| Code | Description |
|---|---|
| A | Agreed between Shippers |
| AR01 | Actual Read (Replacement) |
| B | Xoserve estimated unbundled or opening read |
| C | End user read (bundled) |
| D | Xoserve estimated unbundled final read |
| E | Estimated / Automatic |
| F | Final read for metering transaction |
| G | Gas card Read (Opening) |
| I | Information read |
| J | Further read agreed between Shippers, used for final unbundled meter reads |
| K | End user read provided by the Shipper |
| L | Further read not agreed between Shippers, used for final unbundled read |
| M | Estimated (manual) |
| N | A Normal / Firm read |
| O | Opening read for metering transaction |
| P | Opening read for corrector transaction |
| Q | Shipper Provided Estimated Read |
| R | Replacement read |
| S | Shipper provided read |
| T | Transfer of ownership |
| U | Meter reading organisation read, provided by the Shipper |
| V | Cyclic read from MRA and is used for Shipper transfer |
| W | Cyclic read from Shipper used for transfer |
| X | Remote Reading Equipment Read (Normal) |
| XROx | Remote Reading Equipment Read (Replacement) |
| Y | Remote reading Equipment Read (Opening) |

## A.6 METER_INFO

| 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 - Correction factor (P & T) |

## A.7 NEW_LOST_SITES

This data contains all meter points with a first confirmation date or an end date from 01/04/2008 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 |

## A.8 METER_ERRORS

| Name | Description |
|---|---|
| METER_TYPE | Data for all of the following meter point categories is required: DM, Unique, CSEP, LDZ Offtake |
| LDZ | Char[2] - LDZ identifier e.g. EA |
| START_DATE | Date - Start date of error |
| END_DATE | Date - End date of error |
| ADJUSTMENT | Number - Value of adjustment in kWh |

# Appendix B Consumption Algorithm Database Description

This appendix describes the data structure used by the AUGE to store the data required for the consumption analysis.

The majority of data is stored in separate tables for each LDZ. The two letter abbreviation for each LDZ is appended to the name of the relevant tables. This is denoted below by _XX. Where a database field is described as raw data it contains unprocessed data from Xoserve. All other fields are derived from this information.

There is a database package which encodes the consumption algorithm. It is run by calling consumption.calculate_all(p_year=>XXXX);

There are also two packages POPULATE_SITE_LIST and PROCESS_METER_READS which help with the necessary pre-processing of the data.

## B.1 ANNUAL_QUANTITY_XX

| Name | Description |
|---|---|
| MPR_ID | Raw data |
| AQ_EFFECTIVE_DATE | Raw data |
| EUC | Char[5] – Strip LDZ and year from full EUC Code to give e.g. 08W02 |
| AQ | Raw data |
| SITE_TYPE_FLAG | Raw data |
| LDZ | Char[2] - taken from first 2 digits of EUC |
| EUC_CALC | Char[3] - EUC consumption band calculated from AQ e.g. 01B – Needed to handle supply points containing multiple meter points. |

## B.2 FACTORS

There is a daily version containing the following information for 01-Apr-2008 onwards

| Name | Description |
|---|---|
| LDZ | Char[2] - LDZ identifier e.g. EA |
| EUC_BAND | Char[5] – EUC band e.g. 08W02 |
| GAS_DAY | Date |
| ALP | Raw data |
| DAF | Raw data |
| EWCF | Raw data |
| WC | Number - Max(0.01, 1 + DAF * EWCF) |
| CV | Raw data |
| WAALP | Number - ALP * WC |

| Name | Description |
|---|---|
| VOL_WAALP | Number - WAALP / CV |

Then to help speed up the consumption algorithm there is a yearly version which aggregates the WAALP and VOL_WAALP by formula year.

| Name | Description |
|---|---|
| LDZ | Char[2] - LDZ identifier e.g. EA |
| EUC_BAND | Char[5] – EUC band e.g. 08W02 |
| F_YEAR | Number |
| CWAALP | Number - Sum WAALP |
| VOL_CWAALP | Number - Sum VOL_WAALP |

## B.3 METER_INFO_XX

| Name | Description |
|---|---|
| MPR_ID | Raw data |
| LDZ | Raw data |
| NUM_DIALS | Raw data |
| IMP_IND | Raw data |
| METER_FITTED_DATE | Raw data |
| UNITS | Raw data |
| CORRECTION_FACTOR | Raw data |
| UNITS_CALC | Number - Units estimated from meter reads |
| IMP_IND_CALC | Char[1] – Indicator flag taken from meter read records. |

The UNITS_CALC field is calculated by looking at the ratio of the difference between meter reads and the recorded volume (since read unit data is unreliable).

## B.4 METER_READS_XX

| Name | Description |
|---|---|
| MPR_ID | Raw data |
| START_READ_DATE | Raw data |
| METER_READ_DATE | Raw data |
| IMP_IND | Raw data |
| METER_READ_VAL | Raw data |
| METERED_VOL | Raw data |
| ROUND_THE_CLOCK_IND | Raw data |
| AQ | Raw data |
| METER_READ_FREQ | Raw data |

| Name | Description |
|---|---|
| SSP_LSP | Raw data |
| EUC | Char[5] – Strip LDZ and year from full EUC Code e.g. 08W02 |
| READ_TYPE_CODE | Raw data |
| LDZ | Char[2] - taken from first 2 digits of EUC |
| BAD_READ | Char[1] - Indicator ="Y" don't use meter read |
| GAP_OVERLAP | Number – Positive indicates a gap, negative indicates an overlap |
| VOLUME_CORRECTION | Number- Difference between metered volume and calculated volume since the last read |

## B.5 NDM_DM_CHANGE

This table is pre-calculated from the AQ records

| Name | Description |
|---|---|
| MPR_ID | Number - Unique ID for meter point used across ALL data |
| NDM_START_DATE | Date – date when site becomes NDM |
| NDM_END_DATE | Date – date when site becomes DM |
| LDZ | Char[2] - LDZ identifier e.g. EA |

## B.6 RESULTS_XX

| Name | Description |
|---|---|
| MPR_ID | Number |
| LDZ | Char[2] |
| EUC | Char[5] – taken from AQ record |
| F_YEAR | Number |
| CONSUMPTION | Number – Consumption (in kWh) for formula year calculated using meter reads / metered volumes |
| OLD_AQ | Number - AQ (in kWh) chosen during consumption algorithm. |
| NEW_AQ | Number – Updated AQ estimate (in kWh) based on consumption value. |
| METER_READS | Char[1] - Indicator ="Y" found two meter reads which satisfy the criteria listed in the algorithm |
| POSITIVE_VOLUME | Char[1] - Indicator ="Y" positive volume calculated after possibly correcting for meter index rollover |
| AQ_CHECK | Char[1] - Indicator ="N" if FY_MR_CON puts the site into the LSP market and is >5 times the consumption calculated using the AQ |
| YEAR_FRACTION | Number – (0<= <=1) fraction of the year for which the site was active (calculated using the ALPs) |
| EUC_CALC | Char[3] – Consumption band calculated based on consumption from meter read data if calculated successful, else on the AQ |

| Name | Description |
|---|---|
| OLD_AQ_DATE | Date – Effective date for the old AQ |
| METER_ASSET_DATE | Date – Installation date for the meter in place during the metered period |
| START_READ_DATE | Date – Date of the start meter read used in the calculation |
| END_READ_DATE | Date – Date of the end meter read used in the calculation |
| CALC_TIMESTAMP | Date – Date and time calculation was carried out |

## B.7 SITE_LIST_XX

This table is populated with a unique list of MPR ids from the AQ table. Start and end dates taken from NEW_LOST_SITES and NDM_DM_CHANGE tables.

| Name | Description |
|---|---|
| MPR_ID | Number |
| START_DATE | Date – Date from which the site is active and NDM |
| END_DATE | Date – Date from which the site ceases to be active or NDM |

# Appendix C Worked Example of Consumption Algorithm

This appendix shows the consumption algorithm, described in Section 6.1.2, applied to example data.

## C.1 Full Example

To calculate the consumption for MPR_ID 913600 (which is in EA LDZ) for formula year 2009 the following steps are taken:

1. Check the site is active and NDM in 2009: Yes

2. Select a representative AQ. In this case the AQ from 1/10/2010 is used as the meter reads it is based on are most representative of the demand for 2009/10.

| MPR_ID | AQ EFFECTIVE DATE | EUC | AQ | SITE TYPE FLAG | LDZ | LATEST | EUC_CALC |
|---|---|---|---|---|---|---|---|
| 913600 | 01/10/2007 | 01B | 7544 | N | EA | Y | 01B |
| 913600 | 01/10/2008 | 01B | 5523 | N | EA | Y | 01B |
| 913600 | 01/10/2009 | 01B | 9457 | N | EA | Y | 01B |
| **913600** | **01/10/2010** | **01B** | **10477** | **N** | **EA** | **Y** | **01B** |
| 913600 | 01/10/2011 | 01B | 11505 | N | EA | Y | 01B |

3. Find candidate meter read dates (see meter read table overleaf)

| LB1 | LB2 | UB1 | UB2 |
|---|---|---|---|
| 19/03/2009 | 28/04/2009 | 18/01/2010 | 12/04/2010 |

4. Choose the best two

LB1-’01-apr-2009’ < 540 and no meter replacement since 20/11/2008 so use LB1

UB2-’31-mar-2010’< 540 and no meter replacement since 20/11/2008 so use UB2

5. Validate the choice of meter reads

   UB2 – LB1 > 120

   (LB1,UB2) intersection ( '01-apr-2009', '31-mar-2010') > 60

   So we have found two valid reads

| MPR ID | METER READ DATE | IMP IND | METER READ VAL | METERED VOL | ROUND THE CLOCK IND | AQ | METER READ FREQ | SSP_LSP | EUC | LDZ | READ TYPE CODE | BAD READ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 913600 | 20/11/2008 | N | 5707 | 211 | 0 | 5523 | 6 | SSP | 01B | EA | U | N | |
| 913600 | 02/03/2009 | N | 6229 | 534 | 0 | 5523 | 6 | SSP | 01B | EA | U | N | |
| **913600** | **19/03/2009** | **N** | **6275** | **47** | **0** | **5523** | **A** | **SSP** | **01B** | **EA** | **L** | **N** | **LB1** |
| 913600 | 28/04/2009 | N | 6400 | 128 | 0 | 5523 | A | SSP | 01B | EA | U | N | LB2 |
| 913600 | 28/08/2009 | N | 6455 | 56 | 0 | 5523 | A | SSP | 01B | EA | U | N | |
| 913600 | 18/01/2010 | N | 6964 | 521 | 0 | 9457 | A | SSP | 01B | EA | U | N | UB1 |
| **913600** | **12/04/2010** | **N** | **7438** | **485** | **0** | **9457** | **A** | **SSP** | **01B** | **EA** | **U** | **N** | **UB2** |
| 913600 | 01/06/2010 | N | 7518 | 82 | 0 | 9457 | A | SSP | 01B | EA | U | N | |
| 913600 | 14/12/2010 | N | 7928 | 419 | 0 | 10477 | A | SSP | 01B | EA | U | N | |
| 913600 | 22/08/2011 | N | 8665 | 58 | 0 | 10477 | A | SSP | 01B | EA | U | N | |
| 913600 | 15/11/2011 | N | 8844 | 183 | 0 | 11505 | A | SSP | 01B | EA | U | N | |
| 913600 | 04/02/2012 | N | 9340 | 507 | 0 | 11505 | A | SSP | 01B | EA | U | N | |
| 913600 | 27/07/2012 | N | 9968 | 642 | 0 | 11505 | A | SSP | 01B | EA | U | N | |

6. Calculate the volume consumed between the two meter reads:

   Site is an 01B so calculate volume as difference of meter reads

   Difference = 7,438 – 6,275 = 1,163

   This is positive.  The meter is metric and has been predetermined to have read units=1

   Therefore the final volume is 1,163 * 1.022640 = 1,189.33 $m^3$ (Compared to 1,190 $m^3$ if we had used the metered volumes in this case)

7. Calculate consumption for formula year 2009 based on meter reads

   The meter was active for the whole year, so

   Volume taken over the year is = 1,189.33 * sum volume profile over 2009 / sum volume profile over metered period

   =1,189.33 * 9.40 / 10.19

   =1,097.12 $m^3$

   Weighted average CV for 2009 is = 370.46 / 9.40 = 39.40

   Therefore consumption = 1097.12 * 39.40 / 3.6 = 12,007.51 kWh

8. Calculate a new AQ based on this consumption and compare it to the AQ chosen earlier

   New AQ = consumption * 365 / CWAALP = 12,007.51 * 365 / 370.46 = 11,830 kWh

   This makes the site still 01B and is consistent with the old AQ estimate

## C.2 Example of Meter Index Roll Over Detection Algorithm

Given the following meter reads:

| MPR ID | METER READ DATE | IMP IND | METER READ VAL | METERED VOL | ROUND THE CLOCK IND | AQ | METER READ FREQ | SSP_LSP | EUC | LDZ | READ TYPE CODE | BAD READ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16608022 | 17/11/2008 | Y | 8601 | 21782 | 0 | 22310 | 6 | SSP | 01B | WN | U | N |
| 16608022 | 28/05/2009 | Y | 9086 | 49598 | 0 | 22310 | 6 | SSP | 01B | WN | U | N |
| **16608022** | **19/11/2009** | **Y** | **9257** | **17487** | **0** | **22826** | **6** | **SSP** | **01B** | **WN** | **U** | **N** |
| **16608022** | **15/02/2011** | **Y** | **299** | **-916081** | **0** | **19974** | **6** | **SSP** | **01B** | **WN** | **U** | **N** |
| 16608022 | 16/08/2011 | Y | 572 | 1050558 | 1 | 19974 | 6 | SSP | 01B | WN | K | N |
| 16608022 | 29/02/2012 | Y | 967 | 1063034 | 1 | 19974 | 6 | SSP | 01B | WN | U | N |

We initially calculate the difference between the reads to be = 299 – 9,257 = -8,958

As this is negative we test for meter index roll-over

num_dials = round_up(log(10, start_mr)) = round_up(log(10, 9,257)) = round_up(3.97) = 4

max_read = $10^{num\_dials} = 10^4$ = 10,000

num_years = (15/02/2011 – 19/11/2009) / 365 =1.24

new_diff = (max_read - start_mr + end_mr) = 10,000 – 9,257 + 299 = 1,042

The check is: new_diff / max_read / num_years < 0.25

We have 1,042 / 10,000 / 1.24 = 0.08 > 0.25 so we set the difference to be 1,042 and continue.

# Appendix D Theft Analysis Database Description

The database structure used by the AUGE for the theft an analysis is very similar to the structure used for the consumption analysis.

The theft analysis results were generated using the THEFT package.

NOTE: The MPR_IDs used for the theft data are not consistent with those provided with the data for the consumption analysis (Xoserve currently hold the conversion mapping).

## D.1 ANNUAL_QUANTITY

| Name | Description |
|---|---|
| MPR_ID | Raw data– dummy MPR ID |
| START_DATE | Raw data |
| END_DATE | Raw data |
| EUC | Char[5] – Strip LDZ and year from full EUC Code to give e.g. 08W02 |
| AQ | Raw data |
| SITE_TYPE_FLAG | Raw data |
| LATEST | Char[1] - Indicator ="Y" latest AQ record within gas year |
| EUC_CALC | Char[3] - EUC consumption band calculated from AQ e.g. 01B – Needed to handle supply points containing multiple meter points. |

EUC_CALC is as for the consumption algorithm to determine EUC group from an AQ.

## D.2 FACTORS

There is a daily version containing the following information for 01-Apr-2006 onwards

| Name | Description |
|---|---|
| LDZ | Char[2] - LDZ identifier e.g. EA |
| EUC_BAND | Char[5] – EUC band e.g. 08W02 |
| GAS_DAY | Date |
| ALP | Raw data |
| DAF | Raw data |
| EWCF | Raw data |
| CV | Raw data |
| ENERGY_PROFILE | Number - ALP * (1 + DAF * EWCF) |
| VOL_PROFILE | Number - ENERGY_PROFILE / CV |

Then to help speed up the consumption algorithm there is a yearly version which aggregates the ENERGY_PROFILE and VOL_PROFILE by formula year.

| Name | Description |
|---|---|
| LDZ | Char[2] - LDZ identifier e.g. EA |
| EUC_BAND | Char[5] – EUC band e.g. 08W02 |
| F_YEAR | Number |
| ENERGY_PROFILE | Number - ALP * (1 + DAF * EWCF) |
| VOL_PROFILE | Number - ENERGY_PROFILE / CV |

## D.3 METER_DIALS

| Name | Description |
|---|---|
| MPR_ID | Raw data– dummy MPR ID |
| LDZ | Raw data |
| DATE_FITTED | Raw data |
| NUM_DIALS | Raw data |

## D.4 METER_READS

| Name | Description |
|---|---|
| MPR_ID | Raw data – Dummy MPR ID |
| METER_READ_DATE | Raw data |
| IMP_IND | Raw data – Imperial/Metric indicator |
| METER_READ_VAL | Raw data – Meter Read |
| METERED_VOL | Raw data – Calculated consumption as provided by Xoserve |
| SSP_LSP | Raw data – Market sector |
| EUC | Char[5] – Strip LDZ and year from full EUC Code e.g. 08W02 |
| LDZ | Char[2] - taken from first 2 digits of EUC |
| CORRECTION_FACTOR | Raw data – T&P correction factor |
| UNITS | Raw data – Read Units |
| BAD_READ | Char[1] - Indicator ="Y" don't use meter read |

The algorithm for flagging bad reads is as for the consumption algorithm.

## D.5 T_RESULTS(_V2)

There are two versions of this table. The consumptions in T_RESULTS are calculated using default meter index units and the consumptions in T_RESULTS_V2 are calculated using the meter index units from the meter asset data provided by Xoserve.

| Name | Description |
|---|---|
| MPR_ID | Number – dummy MPR ID |
| SSP_LSP | Char[3] – Sector classification calculated from consumption+theft process |
| F_YEAR | Number – Formula year |
| THEFT | Number – theft amount that occurred within the formula year |
| CONSUMPTION | Number – consumption estimate using meter reads |
| NEW_AQ | Number – Updated AQ estimate based on theft algorithm |
| OLD_AQ | Number – AQ to be used if consumption estimate fails |
| AQ_DATE | Date – effective date of OLD_AQ |
| CALC_METHOD | Number – (1,2 or 3) indicates whether OLD_AQ is pre, post or during theft and only used if consumption calculation fails |
| LDZ | Char[2] – LDZ that the theft occurred in |

## D.6 TOG

This table contains the raw theft record data

| Name | Description |
|---|---|
| MPR_ID | Raw data – dummy MPR ID |
| FROM_DATE | Raw data – Estimated start date of theft |
| TO_DATE | Raw data – Estimated end date of theft |
| LDZ | Raw data – LDZ the theft occurred in |
| SSP_LSP | Raw data – Current SSP/LSP market sector classification at the time of data extract |
| AQ | Raw data – Current AQ at the time of data extract |
| THEFT | Raw data – The estimated amount of theft that occurred during the period of theft in kWh |
| EUC | Char[3] – Consumption band portion of full EUC, e.g. 03B |
| METER_READ_FREQ | Raw data – Frequency of meter reads (A, 6 or M) |

**END OF DOCUMENT**