GL Noble Denton

Report Number: 13365	*16th December 2012*

2012 Allocation of Unidentified Gas Statement for 2013/14

Not Restricted	*GL Noble Denton*

Prepared for:

Uniform Network Code Committee

Prepared by:

Andy Gordon, Clive Whitehand, Tony Perchard

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 2

Customer Reference: 1-6AS9AC

Version	Changed against	Date	Status (for consultation For approval Approved)
1.0	First Draft 2012 AUGS for 2013/14 for Uniform Network Code Committee	30th April 2012	Superseded
2.0	Second Draft 2012 AUGS for 2013/14 for Uniform Network Code Committee	16th December 2012	For Consultation

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 splitting 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 methods described in this document have 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.

Following the approval and publication of the Allocation of Unidentified Gas statement (AUGS) covering April 2012-March 2013, this document is the second draft 2012 AUGS for the period April 2013 – March 2014.

The document describes analyses undertaken in 2012 to improve the estimate of Unidentified Gas and investigates a number of issues arising from the consultation and query processes from 2011 and 2012. Some of the topics investigated are specific to the RbD based methodology from 2011, some are specific to the proposed Consumption methodology, and there are some that apply to both. The following key topics have been covered:

- Handling of new and lost meter points
- Weather Correction Factor and Unidentified Gas
- Change in the basis for Seasonal Normal CWV
- Meter points with AQ=1 (kWh)
- Estimation of total Unidentified Gas using meter reads and metered volumes
- A revisit of the theft methodology using consumption data and investigation of sites where theft exceeds the SSP/LSP threshold in a formula year.

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 in detail. Data used is described in section 5. The resulting methodology proposed for April 2013 – March 2014 is described in Section 6.

As reported throughout the year there have been a number of delays in obtaining data for the consumption analysis. This part of the project has required a significant amount of time and effort from both the AUGE and Xoserve working closely together to prepare a data set that could be used for the analysis. There are inherent issues with the raw data which we have had to identify and account for in the methodology. Any initiatives to improve this would also improve the estimates of Unidentified Gas (regardless of the chosen methodology, since AQs also use the same set of meter reads/metered volumes and these form the basis for the RbD based methodology). A summary of data issues and how these have been addressed is given in this statement.

Over time metered volumes (i.e. volume calculated from meter reads) are corrected, although this only applies to the LSP sector in the data we have, and it was recommended these should be used in the analyses. This has been done where appropriate. We have found that the calculated metered volumes can be erroneous, however, and in such situations (examples are provided in the statement) using meter reads directly can and does provide a better result. If the resulting calculated consumption fails our validation process it is rejected. Therefore the analysis uses both meter reads and metered volumes.

Our primary focus has been the methodology based on meter reads and metered volumes since this will drive the total Unidentified Gas figure. In parallel with this analysis, we have also revisited the estimation of theft split between market sectors. This also involved obtaining consumption data associated with theft-affected sites, and similar data issues were encountered. Not surprisingly, there are more issues with meter reads and metered volumes for theft-affected sites. In the interim report we identified various methods to address shortcomings of the original method, and these have been updated in this report.

Additional analyses have also been carried out to investigate and improve the handling of new and lost meter points. A bias in initial AQ estimates for new sites has been identified, and this will be incorporated into the methodology for both unregistered sites and the consumption methodology. Improved understanding of sites with AQ=1 and non-consuming sites in general has also been incorporated into the consumption based method.

The change in basis for Seasonal Normal CWV and Unidentified Gas associated with the Weather Correction Factor are specific to the RbD based method from 2011. These analyses are included in this statement for completeness although they do not form part of the proposed consumption methodology for 2012.

During consultation concerns were expressed that the methodology is undergoing a big change and that this leads to uncertainty in Unidentified Gas figures going forward, including potential for further large changes year on year. The issue of Unidentified Gas is extremely complex, particularly with regard to the nuances of the data that is available with which to estimate it and apportion to each market sector. It is inevitable that there will be changes in the methodology as each topic evolves and additional data becomes available.

The use of meter reads and metered volumes is an improvement on the use of AQ bias and RbD to estimate Unidentified Gas and this is described in detail in Section 4.7. This method is also more intuitive and hence easier to understand, and in addition addresses one of the major concerns with the method from 2011 regarding SSP-assigned Unidentified Gas.

From the analysis carried out this year, the proposed methodology to estimate total Unidentified Gas will use meter reads and metered volumes.

The theft split method from 2011 was constrained by the data available, in that sites were assigned a single market sector for the whole period of theft (i.e. market sector could not change year on year), and as “non theft-affected” AQs were used a theft override concept could not be implemented. This year, with the introduction of meter reads and consumption data, the AUGE has been able to extend the analysis to look at metered plus non-metered consumption, and to classify market sector for each site by formula year. In addition, if the consumption calculation cannot be made for a site AQ is used to set the sector, this can now be overridden in the case where the theft in a formula year exceeds the SSP/LSP market sector threshold.

There are a variety of issues that affect the estimate of theft split when using detected theft data and these issues present a risk of volatility in the theft split going forward. The AUGE has concluded (as described in Section 4.8) that the fairest way of estimating the theft split is to base it on market sector throughput.

The Unidentified Gas totals provided in this report are not final figures and do not currently take account of a number of issues such as the majority of LDZ meter errors. The method does include provision for meter errors, but full data was not available in time for these to be included in this report. We have included some known large LDZ meter errors where this has been possible. The Unidentified Gas figures in this document are therefore provided for information purposes only and do not represent what the final Unidentified Gas estimates for the LDZs will be. Final Unidentified Gas figures will be produced and supplied to the industry once the methodology has been approved.

The AUGE has also monitored network code modifications and assessed whether any of these impact the methodology or need incorporating into calculations.

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.

It should be noted that the latest calculation method is still based on a technique of estimating 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. The elements of Unidentified Gas included in the Balancing Factor are:

- Theft
- Errors in the Shrinkage Estimate
- Open Bypass Valves
- Meters “Passing Unregistered Gas”
- Unknown Sites
- Additional Common Cause Variation

Meter read and metered volume data for the remaining 3 LDZs was received on 14th December and will now be processed. LDZ offtake meter error data has also been received and will be incorporated in the final figures in due course. There are some Shipperless and Unregistered sites reports to be validated before they can be applied to generate the final set of Unidentified Gas figures and we will be working with Xoserve to address these.

The AUGE believes the proposed methodology provides an improvement over the 2011 methodology for the previous year in terms of the accuracy of the estimation of Unidentified Gas and allocation to market sectors, and provides improved stability of the estimates going forward.

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 73,201kWh and above, Smaller Supply Points have an AQ of up to 73,200kWh. 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 [7]) 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 is determined by the volume of gas transported as measured by the metering equipment. For SSPs, the commodity charge is derived 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 [3], 194a [4], 228 [5], 228a [5]), none of which have been accepted by the industry. A further modification, Mod 229 [7] provided for the appointment of the AUGE with responsibility for determining of 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.

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 database used to hold information associated with UG used to develop the methodology
- Details of the analyses carried out in 2012 in preparation of the methodology

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

1.4 Out of Scope

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

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 be provided when the methods detailed in this AUGS have been approved by the UNCC. Table 1 below 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	Complete	Methodology described in 2011 AUGS [20]
Shipperless sites	Updated data provided every two months	Complete	Methodology described in 2011 AUGS [20]
IGT CSEPs	Updated data provided every two months	Complete	Methodology described in 2011 AUGS [20]
Shrinkage error	N/A	Complete	Status described in 2011 AUGS [20]
Shipper responsible theft	Theft data covering detections to 2012 received EUC Groups, meter read frequencies and meter reads and metered volumes received	Updated method proposed	Proposed method described in this document
Metering errors (SSP supply point, NDM LSP Supply point, DM supply point, LDZ offtake metering)	Complete	Complete	Methodology described in 2011 AUGS [20] and Section 4.7 of this document.
Overall UG estimate using approved method from 2011 AUGS for 2012/13	Complete	Complete for 2012/13	Methodology described in 2011 AUGS [20]
Overall UG estimate: new method using meter reads/metered volumes	Received for 13 LDZs	Complete subject to no further issues arising in remaining LDZs	Proposed method 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, particularly with regard to Theft, 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 2011 AUGS for 2012/13 and previous drafts of the 2012 AUGS for 2013/14 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 this draft, updated data sets have been requested from Xoserve which includes consumption data, and at the time of writing data for the remaining 3 LDZs to complete the set of 13 have now 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 error.

Questions regarding the estimation of theft volumes were sent to Shippers on 12/07/2012. A data request for consumers who changed (occupant) was also issues and partial data received.

The AUGE will use the latest data available where appropriate.

The most recent data available has been requested and, in most cases, received. Xoserve have set up several processes for producing reports containing new data on a regular basis. 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.

Meter reads and metered volumes have been provided. Over time LSP metered volumes may be corrected, but the meter reads are not. Xoserve advised the AUGE to use metered volumes, however, analysis has shown that these can be erroneous particularly for non-corrected SSP data. In this case the meter reads are used for SSP and metered volumes for LSP. Details of how these are determined are described in this statement.

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 Section 4.7. In addition, a sensitivity analysis has been carried out regarding the sampling of calculated consumptions and this is described in section 4.7.4.

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 and previous drafts of this AUGS. 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 on 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 [17, 19, 24, 26, 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 AUGS is provided to the GTs for publication on 16th December 2012.

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 has stored 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 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 of each and subsequent methodology (as appropriate) is described in Section 6 and/or in the 2011 AUGS for 2012/13 [20].

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:

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.

Larger Supply Point Non Daily Metered (LSP NDM) Load

The deemed load is first calculated using the allocation algorithm on a daily basis. It is then corrected when genuine meter reads become available, with reciprocal corrections being made to the Smaller Supply Point load via Reconciliation by Difference (RbD). At present, the effect of RbD is usually to increase SSP NDM load. This is evidenced by the fact that across the three gas years from 2008/09 to 2010/11, 79.8% of monthly RbD values were positive, and the average monthly reconciliation quantity (including both positive and negative values) was 35.0 GWh.

Smaller Supply Point (SSP) Load

This is calculated using the same allocation process used for LSP NDM load on a daily basis. When actual LSP NDM readings become available, this is subject to RbD, the effect of which is usually to increase the SSP load as described above.

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:

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) [14].

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 is described in detail in the 2011 AUGS for 2012/13 [20]. This original method is based on RbD values and will be referred to as the RbD based methodology in the remainder of this document. This document (the 2012 AUGS for the 2013/14 formula year) contains an overview of all methods, plus extended details where the methods have changed from the previous version. Major changes have occurred in two areas:

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 September 2012 Interim Report [25], meter read and consumption data is now available for all supply points (both LSP and SSP) and so an improved method based on these is now proposed.
2. The market sector split of undetected theft
   In the 2011 AUGS for 2012/13 this calculation was based on detected theft levels. This method can be influenced by Shipper theft detection strategy, however, and is highly dependent on Shipper-supplied estimates of theft duration and value. Therefore an alternative approach based on market sector throughput is now proposed.

3.2.1 Estimation of Total UG using Meter Reads/Metered Volumes

The overall concept of calculating total UG using metered data is simple. Total UG is calculated 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.

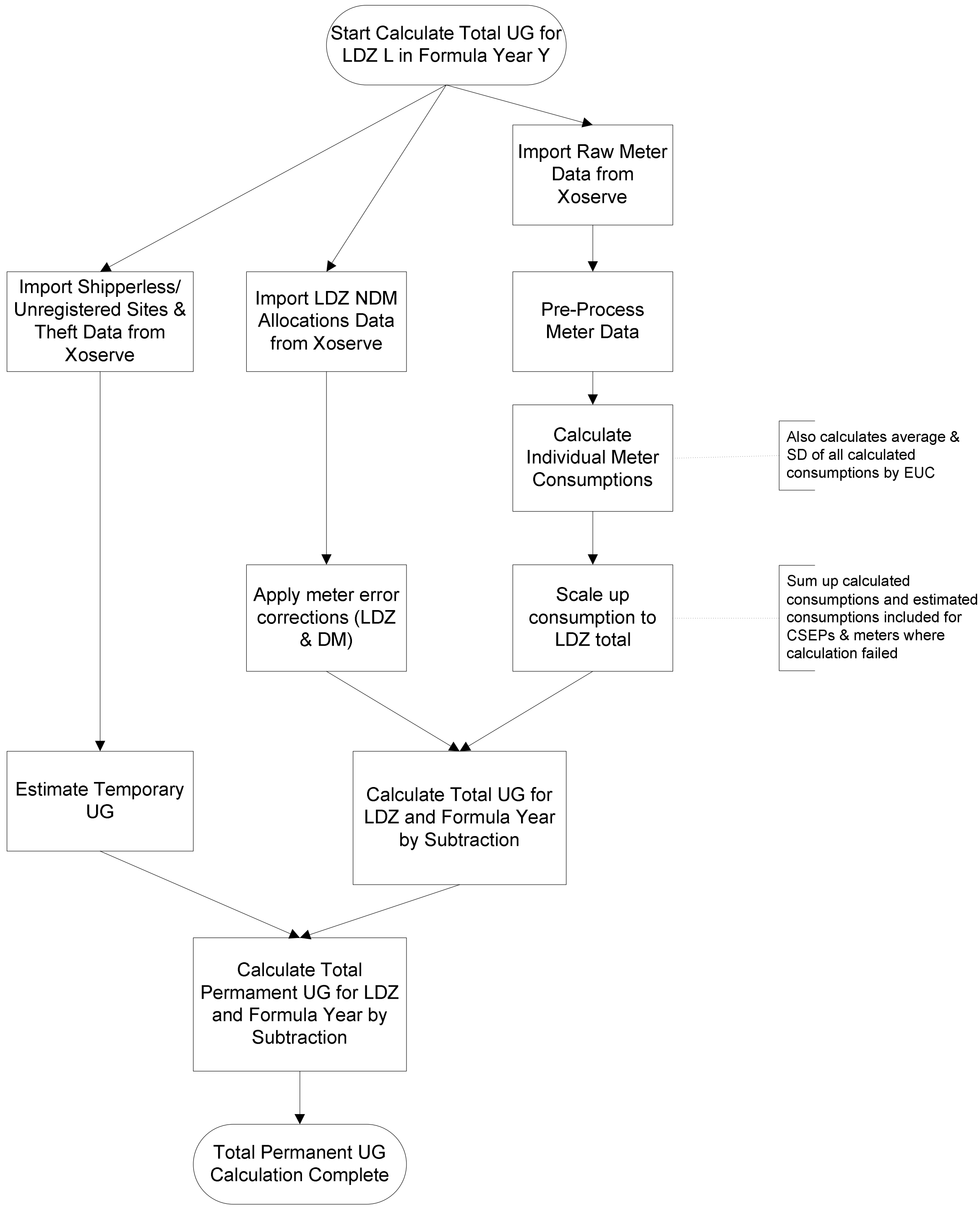


Figure 1: Overview of UG Calculation Methodology

Having obtained the total figure using the proposed 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 previous 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 [20], 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. The most recent data shows that this difference in AQ bias is not present for the most recent year 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 moving to 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, and other general background data on RbD values, AQs, allocation algorithm coefficients, etc. The 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 is therefore to first estimate the UG total for each LDZ, which can be defined as follows:

$$\text{Total UG} = \text{Gas into LDZ} - \text{Metered Gas Out} \qquad (3.1)$$

This can be expressed as follows:

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

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, the sum of DM metered volumes (including their errors) and Shrinkage, the total NDM demand plus UG plus any error in estimating Shrinkage can be obtained.

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 6.1.2.

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 3.4.

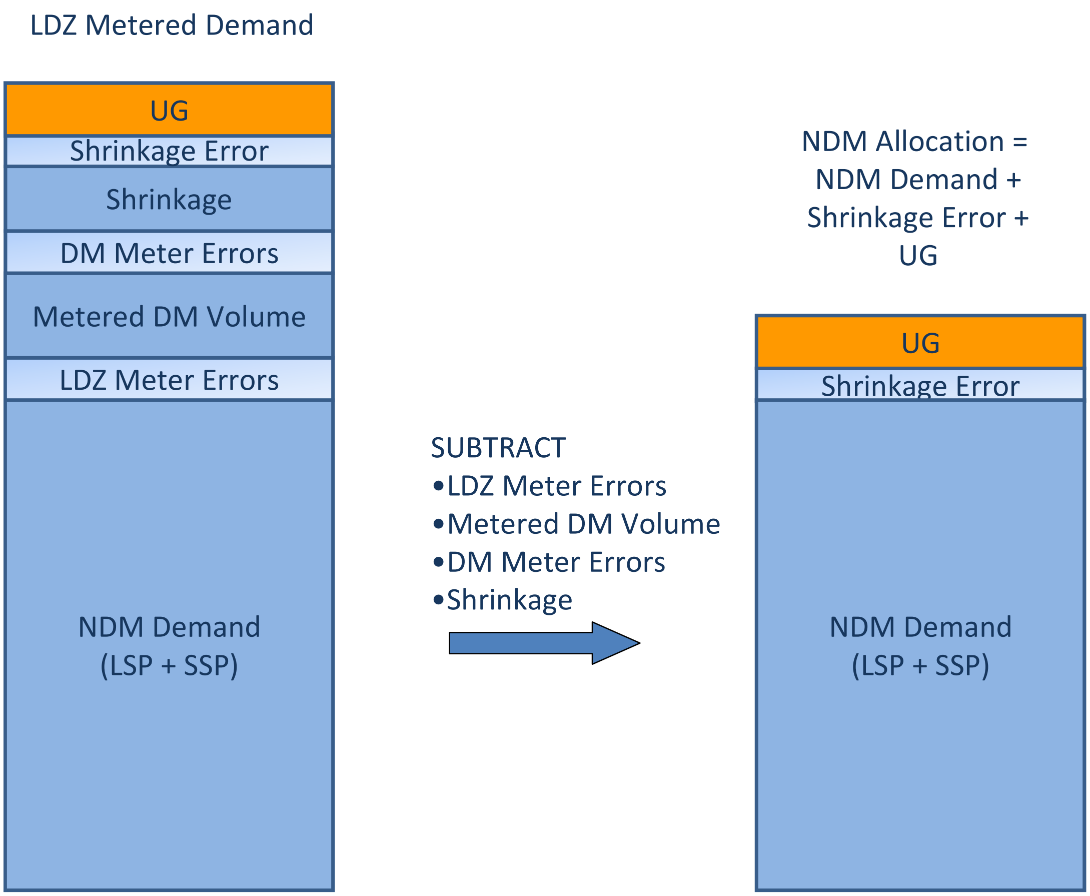


Figure 2: Derivation of Unidentified Gas

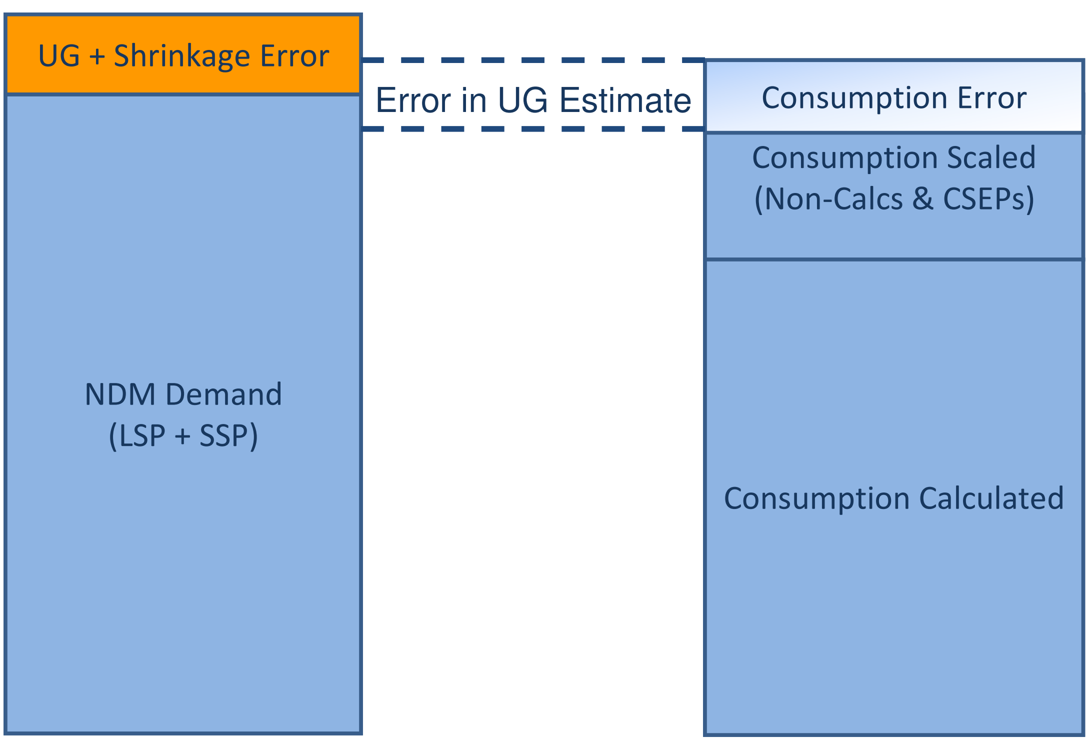


Figure 3: Calculation of Unidentified Gas from Consumptions and Allocations

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 Section 6, Appendices A-C, and in the 2011 AUGS for 2012/13 [20]. Brief descriptions of each LSP 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, the 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 proposes 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 theft 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 allocations, whilst LSP NDM and SSP metering errors have the potential to contribute directly to UG by affecting the metered total. The AUGE has assessed this area and propose that all of these elements of meter error are considered for inclusion in the analysis.

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 Alternative

At this time, in the event of the consumption method not being accepted by the industry for the calculation of total Unidentified Gas for the formula year 2013/14, the figures from the RbD-based method described in the 2011 AUGS for 2012/13 [20] would be used as per AUGE Guidelines [1].

3.4 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. RbD is credited as appropriate.

The consumption method for calculating the UG total includes both permanent and temporary UG. Therefore temporary UG has to be removed from this total prior to further processing. The RbD based method includes only permanent UG in its total, but the permanent/temporary split still affects the component parts of UG and hence influences the market sector split. Therefore the accurate identification of these two types of Unidentified Gas is vital to both methods.

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. *Permanent* for SSP sites on CSEPs.
Shipperless/Unregistered	
- Shipper Activity	*Temporary* if shipper carries out site works. *Temporary* if a third party carries out site works but asset meter read is the same as the shipper's opening meter read. *Permanent* otherwise.
- Orphaned	As for "Shipper Activity".
- Unregistered <12 Months	As for "Shipper Activity".
- Shipperless PTS	*Permanent*
- Shipperless SSrP	*Permanent*
- Without Shipper <12 Months	*Permanent*
Meter Errors	*Temporary* for detected errors that are corrected within the reconciliation period. *Permanent* otherwise.
Theft	*Temporary* for detected theft. *Permanent* for other theft.

4 Summary of Analyses

This section is a summary of the analysis work carried out during 2012. Section 4.1 provides an overview of analyses carried out prior to the AUGE process being introduced. Following Shipper comments on the 2011 AUGS for 2012/13 [20], a number of new areas were identified where potential improvements to the methodology could be made. A brief overview of these areas is given in Sections 4.2 to 4.5.

Sections 4.7 and 4.8 detail the analysis carried out to derive the new proposed approached to estimation of total UG (based on consumption data) and theft sector split.

4.1 Previous Analysis

This section summarises previous analyses and proposals for the Allocation of Unidentified Gas. This is not intended to repeat previous findings but recognise that a lot of work has been carried out previously to solve this problem.

Methodologies to apportion UG to the LSP/SSP markets have been proposed in a number of network code modifications, notably Mods 194 [3], 194A [4], 228 and 228A [5]. In addition Mods 115 and 115A [21] sought to correctly allocate NDM error.

Mod 194 proposed an RbD Allocation table which would apportion a percentage of UG to the SSP and Non-Daily Metered (NDM) LSP and Daily Metered (DM) LSP sectors.

Mod 194A was based on 194 and proposed assigning a fixed volume of UG to the NDM LSP and DM LSP sectors.

Neither proposal populated the tables, with the intention that this would be done via future modification amendments.

Mod 228 proposed to populate the RbD Allocation table defined in Mod 194 with a percentage of UG allocated to each market sector and a methodology to derive these values. Mod 228 also included a paper [18] from CEPA LLP reviewing the proposed methodology.

Mod 228A was based on Mod 228 and proposed fixed values instead of percentages, again with a methodology to derive these values.

None of the above modifications were approved and the rationale for this decision is documented in OFGEM's decision letter of 26th May 2010 [13].

In 2004 OFGEM carried out a study on theft in the GB Gas and Electricity Industry [15] followed up by a next steps document in April 2005 [8]. This showed quite a lot of variation year on year for alleged and proven theft cases. It was also noted that increases in allegations were partly attributed to increased detection activity by the Shippers. One common theme was lack of information regarding the levels of unknown theft and estimates of this vary significantly.

The 228/228A modification report [5] considered three options to calculate theft apportionment and proposed adoption of the third option:

- Estimates based on AQ proportions
- Corrected percentage of 'valid' theft energy
- Simple average between allegations and detected theft

However, it also attributed residual RbD error as being theft. The TPA Solutions report on Mod 228/228A [6] concluded that the hypothesis that reconciliation quantities comprise theft as proposed by Mod 228 did not reflect reality.

There have been several network code modifications considering theft. Mod 274 [10] proposed an independent agent to determine strategies to improve investigation/detection and prevention but was closed on 26th April 2011. Mods 231, 277 and 346 aimed to improve / consider issues with regard to incentives for detection of theft. On March 26th 2012 OFGEM published non-implementation letters [23] for all of them.

A further modification, Mod 399, which is designed to provide transparency of theft detections, is at the final report stage.

In 2011 the AUGE developed a methodology to estimate UG using an RbD based method utilising AQ bias, and a method to split theft based on year of occurrence rather than year of detection. Component parts of UG were identified and estimated for Shipperless Sites, Unregistered sites, Meter Error were included.

The AUGE's September 2012 Interim Report [25] defines the latest methodology for estimating UG and apportioning it between LSP and SSP sectors. This methodology, including assessments of the improvements it brings over previous versions, has been improved and is described in the rest of this document.

4.2 New and Lost Meter Points

The terms 'New Meter Point' and 'Lost Meter Point' have specific meanings in terms of the AUGS. A new meter point is a meter point that becomes live during the year for which calculations are being performed. For the original methodology based on RbD this is gas year, whilst for the proposed consumption based methodology (see Section 6), this is formula year. Similarly, a lost meter point is one that has been isolated during the calculation year. For a new meter point, the meter point is deemed to become effective and start taking gas from the first confirmation date. Xoserve confirmed that this is the date from which a meter point will start to be included in the allocation process regardless of whether there is a physical meter on site. For a lost meter point, it is deemed to have stopped taking gas only once it has been isolated and this is when it ceases to be included in the allocation process.

For the proposed consumption-based methodology, start and end dates (first confirmation and isolation dates) are used explicitly to ensure no consumption is calculated outside of the time window when the meter point was live. The remainder of this section is therefore only relevant to the RbD based methodology for estimating UG.

The deeming algorithm is used to estimate model bias (the component of RbD due to errors in AQ). The estimate is made by applying the deeming algorithm with two different sets of AQ values (see Section 6.3 of the 2011 AUGS for 2012/13 [20] for further details). In each case, however, the AQ values remain fixed for the whole gas year. In reality, the aggregate AQ used in the deeming algorithm changes from day to day as new meter points are introduced and some meter points are isolated. As it is the difference between the results of the deeming algorithm as applied with the two different sets of AQ values that is of interest, it was assumed that the day to day variation would cancel out to a large degree. An analysis has now been carried out to assess the error that this assumption introduces.

The approach taken to estimating the AQ bias is based on calculating the SSP and LSP allocations using the deeming algorithm and the data that was available at the time (AQ, WCF etc.). Repeating this allocation process using updated AQ values based on more recent meter reads (from the Mod81 report following AQ review) gives a more accurate view of what the allocations should have been. The difference in the allocations using the two sets of AQ values is an estimate of the allocation error which resulted purely from errors in the AQ values. For LSP meter points, we would expect these allocation errors to appear in RbD as they represent errors in the allocation process, which actual meter reads should subsequently account for.

Ideally, the above approach should use AQ values (at aggregate EUC level) which vary from day to day through the gas year as new meter points are confirmed and some meter points are isolated. This data was not available for the first year of the AUGS, however, so a single AQ value was used throughout the gas year at this stage. This was considered a reasonable approximation as the method relies on taking the difference between allocations based on two sets of AQ values. The day to day AQ difference will therefore cancel out to a large extent.

Additionally, it was recognised that as new meter points will have an estimated AQ value, there is scope for additional AQ bias as a result of inaccurate initial AQ estimates.

4.2.1 Daily Aggregate AQ Variation

Analysis using daily values of aggregate AQ applies only to the old RbD-based methodology.

Data regarding the AQ of new and lost (isolated) meter points was requested from Xoserve and has been received. In order to use the deeming algorithm with correct daily values, the approach was to start with the aggregate AQ as at 1st October and add/subtract the AQ for new/lost meter points each day. The model bias calculations have now been updated to include this day to day AQ variation due to new and lost meter points. The model bias calculations for 2007-2009 gas years have been repeated to assess the impact of this change and it was found to have an impact of less than 1% for each year, which confirms the validity of the previous assumption.

Analysis of the data regarding new and lost meter points has been carried out in order to understand the magnitude of any AQ variation within the gas year and assess the level of bias which may result. Figure 4 shows the total percentage AQ change from new and lost meter points within the gas year for each year from 2007-2010. This clearly shows that within the year AQ is increasing for both SSP and LSP as the aggregate additional AQ from new meter points is greater than the aggregate AQ from lost meter points. On average over the 4 gas years, new meter points increase the SSP AQ by 0.45% and increase the LSP AQ by 2.3%.

There is scope for new and lost meter points to contribute to AQ bias which is not currently accounted for in the AUGS methodology. A bias can occur if a new meter point starts taking gas on a different date to the first confirmation date or if a meter point stops taking gas before it is actually removed from the allocation process. No data is currently available to assess this further.

New meter points have the potential to contribute significantly to AQ bias as the initial estimated AQ is unlikely to be accurate (see Section 4.2.2 below) and depending when the meter point was first confirmed, there may be a significant time delay (at the very least 6 months and one day, the minimum period for AQ calculation plus another 2 to 6 months before it goes live) before the AQ is updated in the AQ review. Note that the first time that a new meter point is included in the AQ review; it will not be in the Mod81 report as this relies on the meter point being present for two consecutive AQ reviews. Any AQ bias prior to the first time the meter point is included in the AQ review is therefore not accounted for in the current AUGS process.

Lost (isolated) meter points have less potential to contribute to UG. Although their AQ may not be 100% correct, it will have been based on recent meter reads and should therefore be more accurate than the initial AQ estimate for new meter points. Any bias will also only be present for part of a gas year depending on the date at which the meter point is isolated.

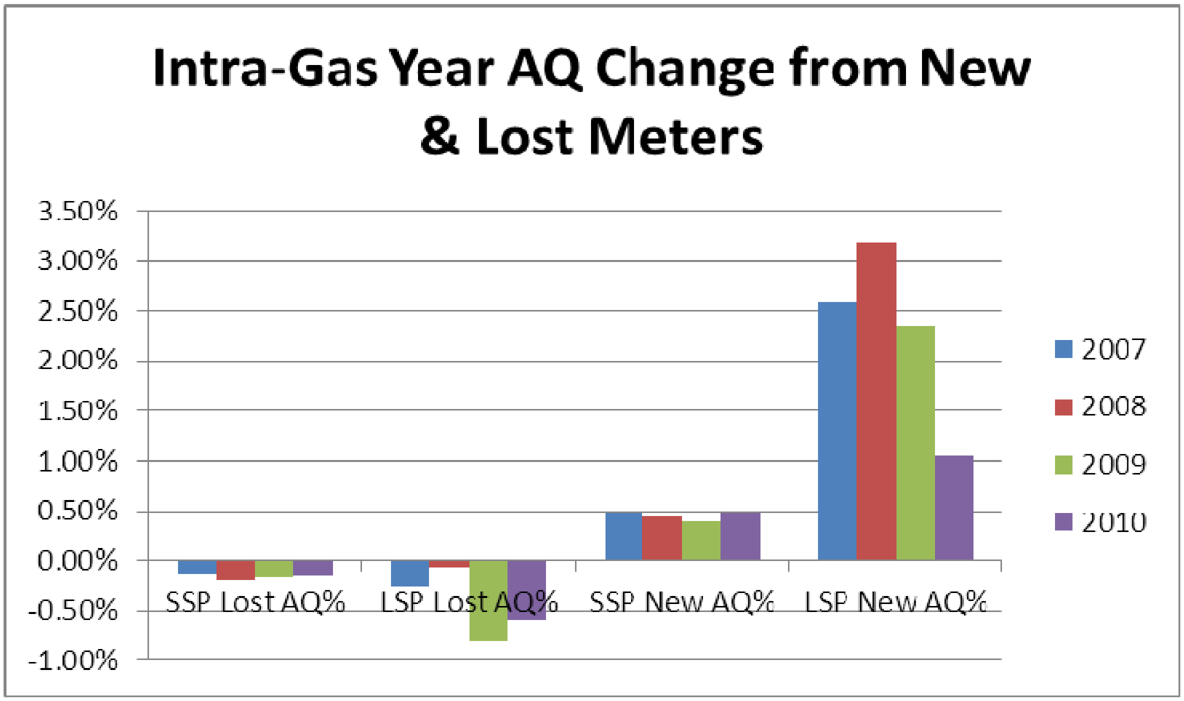


Figure 4: Total % AQ change within year 2007-2010

4.2.2 Estimated Initial AQ Values

A new meter point can contribute to AQ bias if the initial AQ estimate is inaccurate. Although the AQ will be updated following meter reads (which in turn will be captured in the Mod81 report and therefore the current AQ bias estimation process), due to the timing of the AQ review process and the need for a minimum period of meter reads it is likely that any bias will persist for some time before being accounted for. An analysis was therefore carried out to assess the accuracy of initial AQ estimates.

Xoserve provided the AQ history for a sample of sites including their first confirmed (estimated) AQ. The sample consisted of 966 SSP meter points and 223 LSP meter points (based on estimated AQ). Comparing the first AQ with the next different AQ (same AQs were ignored because this will be an AQ roll-over after recalculation failure) provided information about the overall AQ bias for new SSP and LSP meter points in the sample as well as the number of meter points initially in the incorrect EUC. The overall SSP AQ value was overstated by 16.5% whilst the overall LSP AQ was overstated by 48%. This includes the effect of two SSP meter points moving to LSP and 16 LSP meter points moving to SSP.

The bias observed was made up of a general tendency to overstate the AQ by a modest amount plus a subset of the sample where the AQ bias was very large. An example is shown in Figure 5, which shows the initial and subsequent AQ values for a sample of 50 meter points. Figure 6 shows a frequency histogram for the initial AQ biases across all meter points in the sample. From this it can be seen that most meter points have an AQ bias of 0-20% but there are a significant number with AQ bias >100%. The meter points with largest bias from our sample were queried with Xoserve to confirm that these were not the result of incorrect data or that these meter points were suspected of theft.

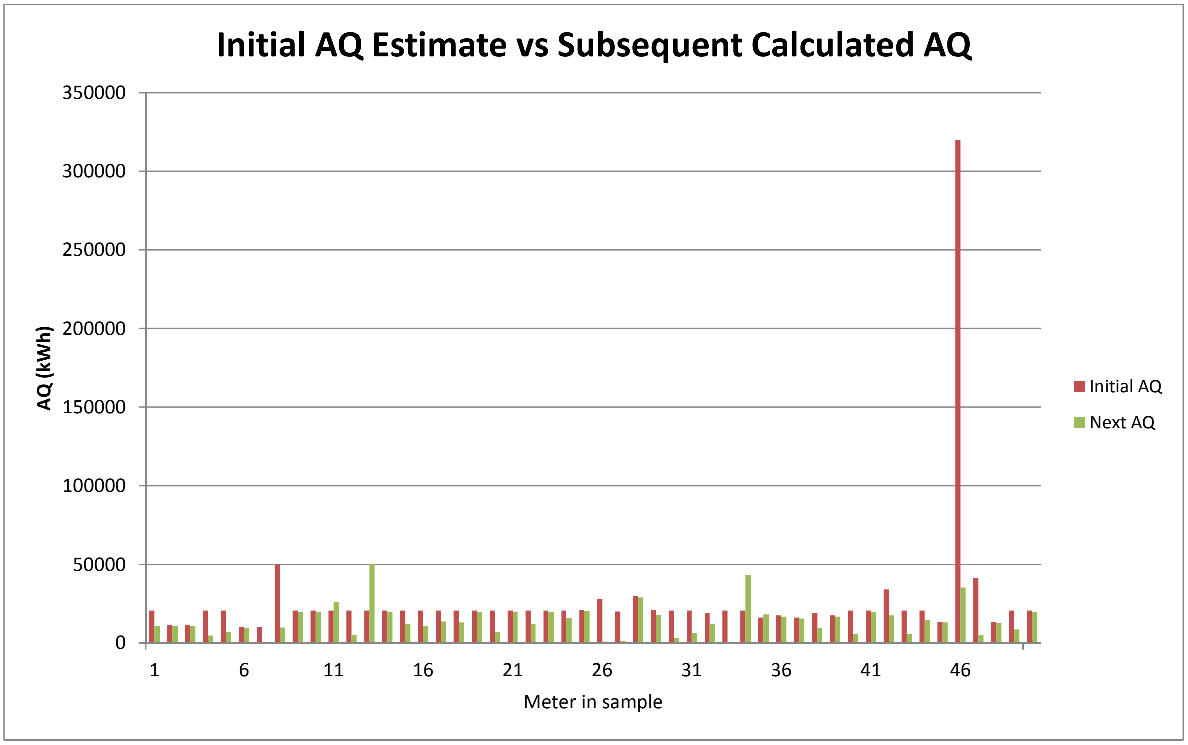


Figure 5: Initial AQ Estimate vs Subsequent Calculated AQ

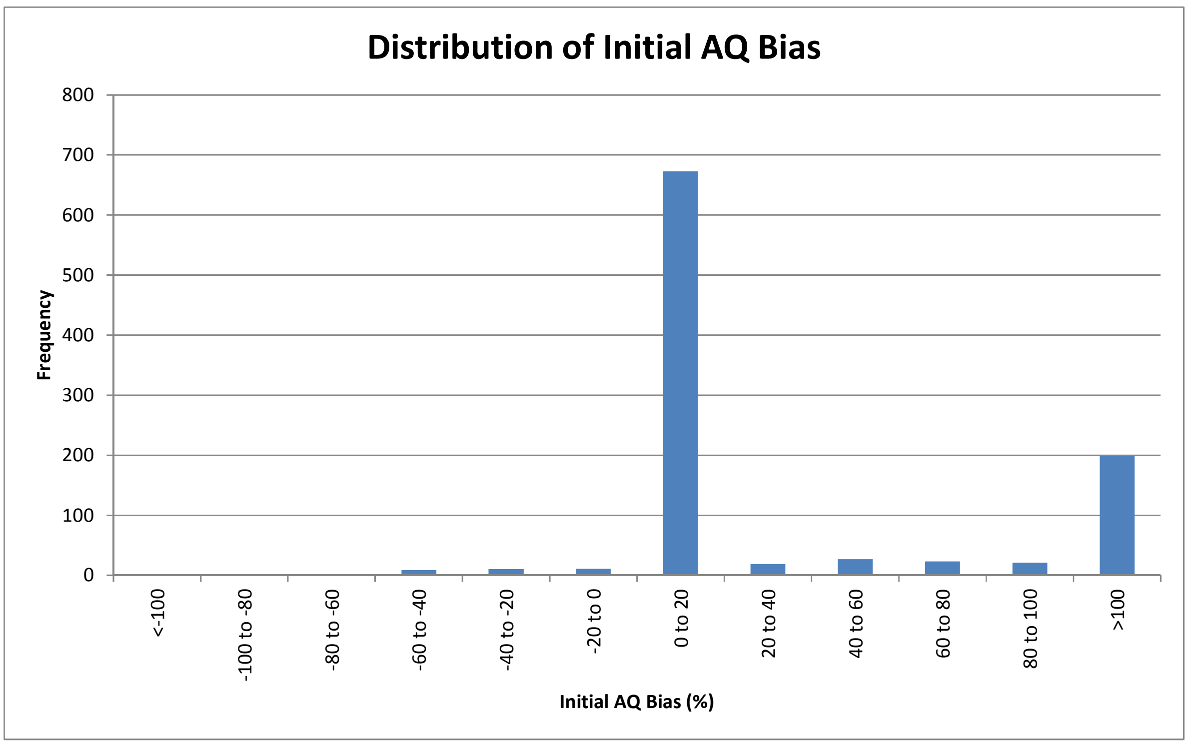


Figure 6: Distribution of Initial AQ Bias

4.2.3 Implications of Initial AQ Estimate Bias for UG Calculation

This bias in initial estimates of AQ will have a significant impact on the RbD-based UG estimation methodology because this uses AQ to estimate model bias. It is recommended that this AQ bias due to new meter points is added to the AQ bias calculated if the RbD based methodology were to be used. For each gas year, the additional bias from equations 4.2 and 4.3 below should be added.

*New Meter point AQ Bias = (Average percentage bias of new meter points * Total AQ of all new meter points added within the gas year)/2* (4.1)

This should be done separately for SSP and LSP. The factor of 2 is to account for the fact that on average new meter points will be taking gas for a half gas year. Using the figures from 4.2.1 and 4.2.2 above, the additional AQ bias from new SSP sites will be (16.5% * 0.45%) / 2 = 0.04% whilst the additional AQ bias from new LSP sites will be (48% * 2.3%) / 2 = 0.55%.

In terms of the proposed consumption based UG calculation, the impact of bias in initial AQ estimates should be negligible in the calculation of consumption as AQ values aren't used directly, only for validation purposes.

Any bias in initial AQ values will impact the calculation of UG for Unregistered sites. The AQ values held by Xoserve are used to estimate the UG for Unregistered sites for SSP and LSP separately. In order to account for the AQ bias, UG for Unregistered sites for the SSP and LSP market sectors should be reduced by the appropriate factor as follows.

$$UG_{SSP}^{unreg} = \frac{UG_{SSP}^{unreg}}{1.165} \quad (4.2)$$

$$UG_{LSP}^{unreg} = \frac{UG_{LSP}^{unreg}}{1.48} \quad (4.3)$$

4.3 Weather Correction Factor

Feedback from the 2011 AUGS for 2012/13 consultation suggested that it is important to understand how UG is allocated by the deeming algorithm, and in particular how it affects the value of the Weather Correction Factor (WCF). Further analysis has therefore been carried out in this area. To aid in this analysis, EWCF (Estimated Weather Correction Factor) data has been supplied by Xoserve. Note that this analysis affects the RbD based method of estimating total UG only, but the analysis may be of wider interest in understanding how UG is apportioned during allocation.

During consultation discussions there were some questions raised about how UG is allocated by the deeming algorithm. It was suggested that all UG is allocated via the Weather Correction Factor (WCF). If this were the case then UG would be split between SSP and LSP in the ratio

$$\frac{UG_{SSP}}{UG_{LSP}} \approx \frac{PSND_{SSP} \times DAF_{SSP}}{\sum_{euc=02B}^{09B} \left(PSND_{euc} \times DAF_{euc}\right)} \quad (4.4)$$

where UG is the Unidentified Gas

PSND is the Pseudo Seasonal Normal Demand

DAF is the Daily Adjustment Factor

If this were the case, it would result in significant quantities of UG being allocated to SSP. The AUGE did not believe this to be the case due to the AQ bias, which overstated LSP AQs by more than SSP AQs during the time period used for the 2011 AUGS for 2012/13, in which the RbD based method was used.

At the start of the 2008 gas year, Mod204 was implemented. This changed the definition of the WCF and significantly changed how UG is allocated by the deeming algorithm. Prior to Mod204, the WCF was given by equation 4.5 where D is the measured LDZ demand and SND is the transporter's estimate for the coming year incorporating historical demands.

$$WCF = (D-SND)/SND \qquad (4.5)$$

As measured demand into the LDZ will include any Unidentified Gas, both the D and SND terms contain quantities of UG. Although the amount of UG in these terms will not be identical, unless UG levels have changed significantly over time then the UG will largely cancel out and the WCF will be a good measurement of the difference in demand due to non-seasonal weather, which was the original intent of this term. Unfortunately, the WCF will also contain an element due to any overall change in average demand e.g. if demand is reducing then the SND calculated as an average over many years will be higher than the current year's demand.

Considering the other terms in the deeming algorithm, any bias in AQ will affect the PSND term (AQ*ALP/365). An overstatement of AQs will cause the right hand side of the deeming algorithm to over-allocate. However, the Scaling Factor is used to ensure the total allocation for the LDZ matches the measured demand, which includes UG. The SF is therefore accounting for the net effect of AQ bias and UG in addition to the random model error.

If UG and AQ bias (overstatement) are of similar magnitude, the SF would be expected to vary randomly around the value of 1. However, if UG is greater than any AQ overstatement then the SF will need to be greater than 1 in order to balance the deeming algorithm to LDZ measured demand. Conversely, where the AQs are overstated by more than the quantity of UG then the SF will need to be less than 1.

Figure 7 shows the SF for EA LDZ as an example (note that the SF has been weighted by the allocation such that the value is proportional to the quantity of gas that the SF will be accounting for). The figure highlights a noticeable change in behaviour at the start of gas year 2008 when Mod204 was introduced. Prior to this, the SF is much more variable as a result of the fact that it is adjusting not only for the model error but also for the net effect of UG and AQ bias. Prior to 2008, the average SF is less than 1, suggesting that the AQs have been overstated by more than the quantity of UG.

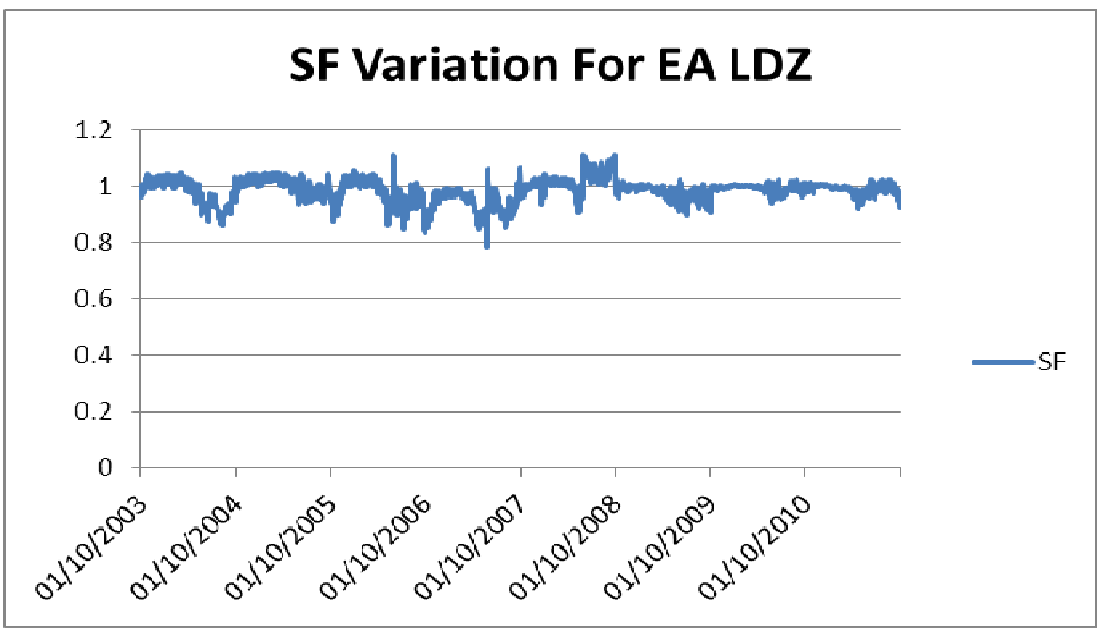


Figure 7: Scaling Factor Variation for EA LDZ

Post-Mod204, the WCF was redefined such that WCF = (D – PSND)/PSND. This changed the way in which the deeming algorithm works, as D contains UG but PSND doesn't. The PSND term is based on AQ, however, and so includes AQ bias. If the AQs are overstated, UG and AQ bias cancel each other out to some extent because they have opposite signs. The WCF is therefore no longer a measure of the effect of non-seasonal weather but includes a correction depending on UG and AQ bias.

An alternative way to look at the deeming algorithm with the new definition of the WCF is that the PSND terms largely cancel out such that the algorithm as a whole is relatively insensitive to AQ bias but includes UG through the LDZ Demand term. This is the reason why post-Mod204 the SF is less variable and closer to 1.

In summary, pre-Mod204 the WCF is determined by the difference in weather from seasonal normal plus any change in overall average demand level between the current gas year and the gas years used to estimate SND. The SF term is adjusting the result to compensate for UG, AQ Bias, change in overall demand level and random model error. Post-Mod204, the WCF contains a mix of weather correction, UG and AQ Bias, whilst the SF is predominantly adjusting for random model error.

Figure 8 shows the WCF and EWCF for EA LDZ (weighted by allocation). EWCF is taken to be a good indicator of the actual change in demand due to non-seasonal normal weather. For 2007 (pre-Mod204), WCF is less than EWCF suggesting that it has been affected by reducing average demand levels causing SND to be higher than recent demands. For 2008 and 2009 (post-Mod204), the WCF is still significantly less than the EWCF. This suggests that the AQ Bias is greater than the level of UG. For 2010, there is a switch with WCF now higher than EWCF, though the difference is much smaller. This suggests that the level of UG is now slightly higher than the AQ bias. This is consistent with the observed reduction in AQ bias (which can be seen in Figure 9 – note that the 2011 Mod81 report year corresponds to the AQ bias in gas year 2010). This suggests that the level of UG is somewhere between the levels of AQ bias observed in 2008/9 (5.6% combined SSP/LSP bias from Mod81) and those in 2010 (1.4% combined SSP/LSP bias from Mod81), but closer to the 2010 value because the difference between WCF and EWCF is smaller for 2010.

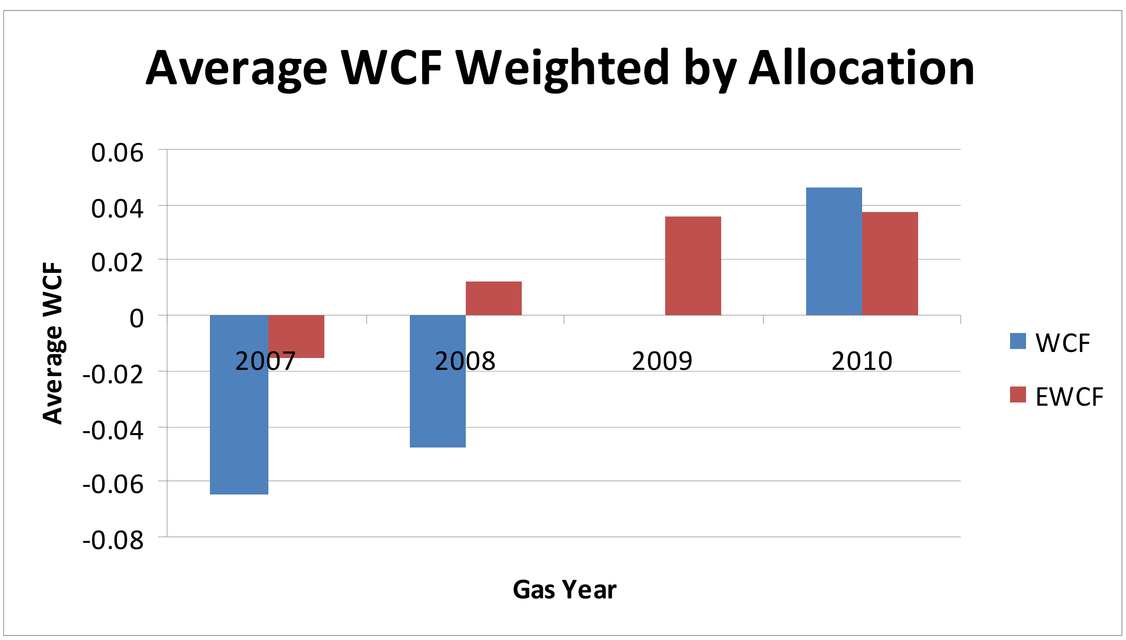


Figure 8: Average WCF Weighted by Allocation

4.4 Seasonal Normal CWV

The 2011 AUGS for 2012/13 [20] noted that the definition of SNCWV (Seasonal Normal Composite Weather Variable) was updated in 2010. As AQs are calculated by adjusting metered consumption back to seasonal normal using the SNCWV, any changes in SNCWV will have a knock-on effect on AQ even if the consumption remains unchanged. At the same time Mod254 [16] also came into force which further adjusts the SNCWV for future forecast values (allowing for the use of forecast data in the definition of SNCWV). These two effects can be assessed together as they both change the value of the SNCWV used to estimate AQ.

An analysis has been carried out to estimate the effect of the changes in AQ as a result of the updates to SNCWV.

To assess the impact of the change in definition of SNCWV, Xoserve provided factors for each EUC. These can be used to adjust the AQ calculated using the old SNCWV definition to the AQ which would result from the new definition. These factors are generally less than 1, i.e. some of the AQ reduction seen between 2009 and 2010 was due to the SNCWV change. It is important to have these factors split by EUC as some EUC bands are more temperature sensitive than others and therefore will be affected more by this change.

The methodology used in the 2011 AUGS for 2012/13 was based upon using the AQ change from the Mod81 report to estimate a trend (bias) in allocations resulting from errors in AQ. Prior to 2010, the Mod81 AQ values are all based upon the old SNCWV definition. Post 2010, the Mod81 AQ values will all be based upon the new SNCWV definition. However, in the 2010 Mod81 report, the 'previous AQ' values are based on the old SNCWV definition whilst the 'current AQ' values are based on the new SNCWV definition. This inconsistency means that some of the change in AQ observed in the 2010 Mod81 report is due solely to the SNCWV definition change (including Mod254 implementation).

Figure 9 below has been generated to show the size of the correction resulting from this change (denoted by the arrows). As expected, the effect on LSP is less than on SSP due to its lower temperature sensitivity. This effect would need to be removed using the factors provided by Xoserve if the RbD based methodology is used. This change does not affect the proposed consumption approach as it does not rely on the AQ values.

Figure 9 is derived directly from the Mod81 report data. For each year (Mod81 report), the sum of the previous AQs based on previous EUC band is calculated. Similarly, the sum of the current AQ based on current EUC band is calculated. The change is then the difference between these values (current AQ – previous AQ) i.e. negative values represent a reduction in AQ following AQ recalculation. The raw Mod81 values are shown by the points on the chart. The data adjusted for the change in SNCWV is shown by the lines. The only difference occurs in the 2010 Mod81 report year and the arrows show the correction.

Note that a manual correction was applied to the Mod81 data to remove a very large 'rogue' meter point in WS 09B (AQ=3,984GWh). Xoserve have confirmed that this meter point was created in error and does not form part of the allocation process.

It should be noted that the AQ changes shown in Figure 9 represent the change for a subset of meter points only (meter points which were live at consecutive AQ reviews and therefore included in Mod81). The change in AQ is the change which has occurred as a result of the AQ recalculation process. There will therefore be meter points which are included in these statistics which have not had their AQ updated for some reason (e.g. insufficient meter reads). In particular, new meter points will not be included in the Mod81 report as there will not be sufficient meters reads.

A process similar to the one proposed here should be used in any future years if changes are made to the SNCWV definition.

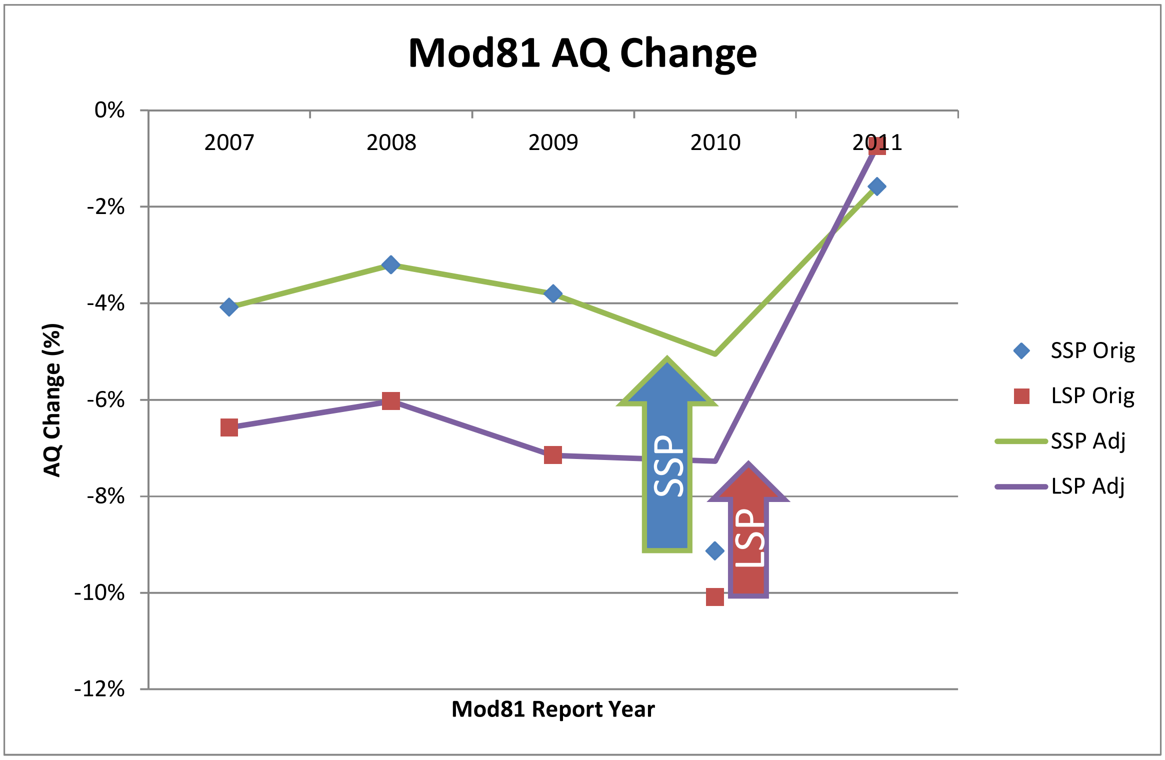


Figure 9: Mod254 and new Seasonal Normal CWV impact on AQ changes

4.5 Meter Points with AQ of 1

During the analysis of meter read and theft data, it became apparent that there are a significant number of meter points with an AQ value of 1. This affects theft estimates, the scaling of sample consumptions to the full population, and any other process which relies on the AQ (including the estimation of model bias for the RbD-based method). Further analysis has been performed to understand the issues of both AQ=1 and other very low AQ values, including

- Why and under what circumstances meter points are assigned an AQ of 1
- How prevalent are meter points with AQ=1 and how many of these will have a different AQ as a result of AQ review?
- Why are there meter points with very low AQ values (not equal to 1)?
- What is the impact on the estimation of UG which arises from meter points with AQ=1 and other very low AQs?
- How should meter points with AQ=1 or other very low AQ values be treated in the estimation of UG?

Based on discussions with Xoserve, the AUGE's understanding is that there is only one way in which a meter point can have its AQ set to the value of 1. This arises during AQ review when two valid meter reads which are identical are used to calculate AQ i.e. the meter reads suggest that no gas has been taken. This could happen for

1. New sites that are vacant and have been long enough to be in an AQ review with two reads at least 6 months and 1 day apart. Initially the new site will have an estimated AQ>1, but if no gas is taken, then at AQ review the meter point AQ can be set to 1 given sufficient reads

2. Vacant sites or sites which have stopped consuming but which have not been isolated or had the meter removed
3. Theft or Meter bypass
4. Faulty meter

Note that in all cases the meter must have been read and have the same value at each read.

For 1 and 2 above, the AQ value of 1 is correct as the meter point is not flowing gas. An AQ bias can result when the site becomes occupied and gas starts to be taken whilst the AQ is still 1. However, this will be corrected in subsequent meter reads and the AQ review process. The RbD based methodology will therefore account for this through the data in the Mod81 report. However, in the case of a new site with an estimated AQ but which does not immediately start taking gas, an AQ bias can result. This situation is covered by the analysis of new sites in Section 4.2.

In cases 3 and 4 above, the AQ value of 1 is not correct and will not be accounted for in the AQ review as the meter reads will not reflect the true consumption.

In terms of the RbD based UG calculation process, no further action is required for meter points with AQ of 1. Cases 1 and 2 are already accounted for by the methodology, whilst theft is handled separately. Faulty meters will eventually be identified and corrections applied through the RbD process. In terms of the recommended consumption based approach, meter points with AQ=1 are treated differently (see Section 6.1.2 for details).

It is also worth noting that there is already a process in place in order to minimise the effect of sites with AQ of 1 which start flowing significant quantities of gas. Xoserve produce a report giving details of any meter points whose AQ is 1 and is rolled over, but where there are meter reads suggesting that the meter point is flowing gas. This report only includes meter points where the meter reads suggest that they may fall into LSP, and is provided to shippers to allow them to update the AQs of these meter points during the AQ amendment process.

Xoserve also provided details of meter points with AQ=1 that were included in the Mod81 report for 2011. There were 299,640 meter points whose previous AQ was 1 and 216,628 whose current AQ was 1. Figure 10 and Figure 11 show how the AQs of these meter points have changed with the AQ review.

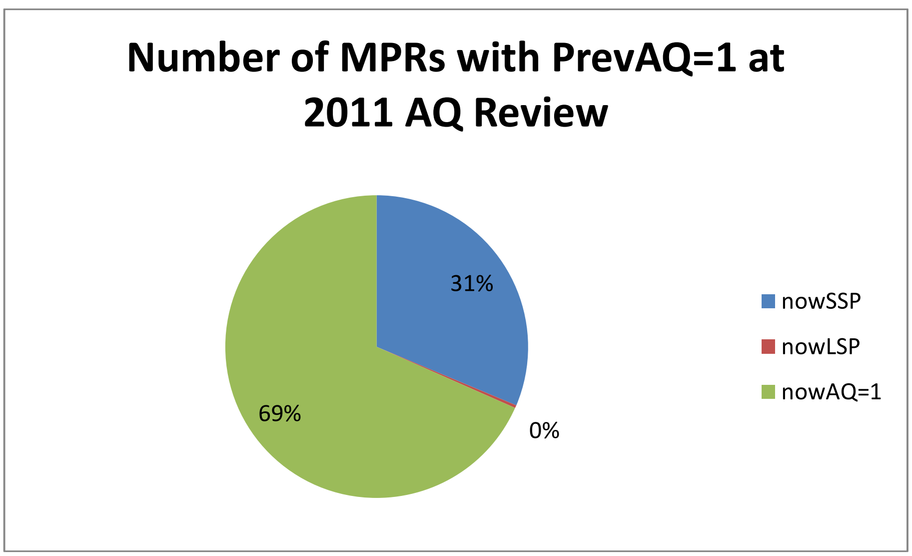


Figure 10: Number of MPRs with Previous AQ=1 at 2011 AQ Review

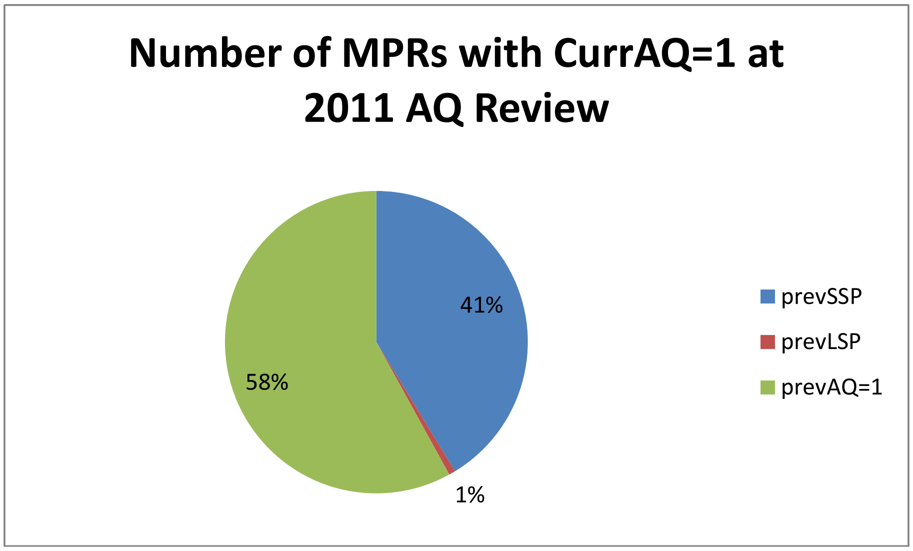


Figure 11: Number of MPRs with Current AQ=1 at 2011 AQ

Figure 10 shows that of the 299,640 meter points with previous AQ=1 (i.e. sites that were not consuming in the previous gas year), 69% kept an AQ of 1, 31% started consuming gas and became SSP whilst less than 1% became LSP. The increase in SSP AQ as a result was 810.7GWh, whilst the increase in LSP AQ was 320.8GWh. Although this shows a bias towards SSP, it should be noted that the AQ calculated could be based on a meter point consuming for a part year which may result in some meter points being SSP which if consuming for the full year would have been LSP.

Figure 11 shows the 216,628 meter points with current AQ=1 (i.e. sites that were not consuming in the most recent year). Of these, 58% were not consuming in the previous year (AQ=1), 41% were previously consuming at SSP levels and 1% were previously consuming at LSP levels. In terms of AQ, the reduction in SPP AQ resulting from meter points stopping consuming was 744.2GWh compared to 754.6GWh for LSP meter points.

In addition to meter points with AQ=1, the AUGE has noted there are a significant number of meter points with very low AQ values. Xoserve have confirmed that these are true values which have resulted from a successful AQ calculation i.e. the meter reads indicated that the consumption was very low. This may be expected in cases where a site has stopped consuming gas such that the meter reads reflect consumption over a part year only. Figure 12 below is a cumulative histogram showing the distribution of low AQ values for WM LDZ in 2010.

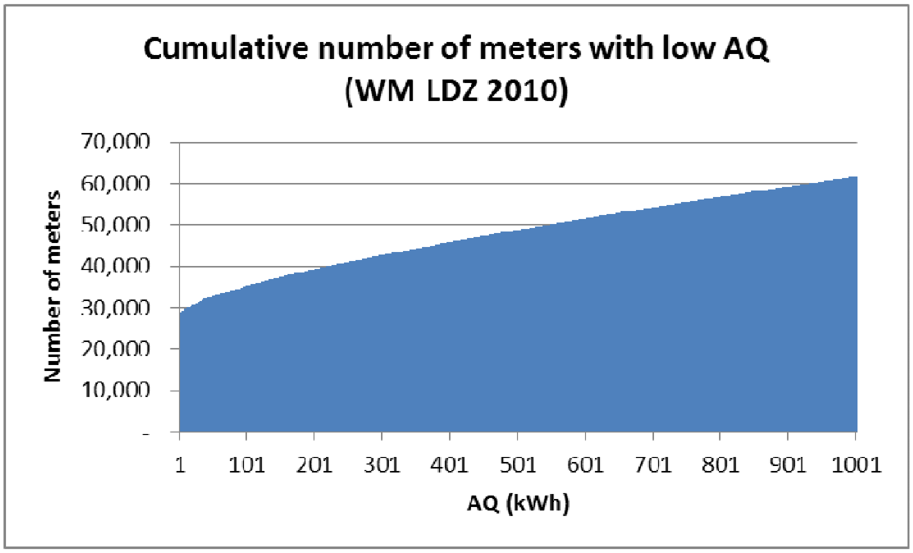


Figure 12: Cumulative Number of Meter Points with Low AQ (WM LDZ 2010)

It can be seen that the number of meter points with AQ=1 is approximately equal to the number of meter points with an AQ between 2 and 1000.

Currently the UG estimation based on RbD makes no specific allowance for meter points with AQ=1 or low AQ values. However, AQ changes are accounted for in the model bias calculation by using the Mod81 data. AQ=1 meter points are handled explicitly in the consumption methodology as described in Section 4 and 6.1.2.

4.6 First Draft AUGS Response Issues Update

4.6.1 iGT CSEPS

An issue was raised during the review of the 1st draft of the 2012 AUGS for 2013/14 regarding the use of CSEP reconciliations in the RbD based method. The consumption based method does not use CSEP reconciliations so this particular area of analysis is not required for the recommended UG calculation approach. It remains the case, however, that information about iGT CSEPS is limited, and this area could be improved if iGTs were willing to provide meter reads and metered volumes as per the rest of the industry. At this time the AUGE's focus has been on the larger issues of the total UG figure and theft split. There are a variety of smaller topics that we can then look at to refine the estimates further, of which this is one.

4.6.2 Asset Meter Inconsistencies

An issue was raised during the review of the 1st draft of the 2012 AUGS for 2013/14 that the scale of the Balancing Factor required further investigation to identify its component parts. One item highlighted was asset and meter inconsistencies. During this year's analysis the main focus has been to improve the method of calculating the total UG figure and the handling of theft split, believed to be the main component of the Balancing Factor.

During the consumption analysis inconsistencies have been identified in terms of read units, meter read entries, meter index issues and LDZ assignments. Where these have been identified the AUGE has corrected or removed from the sample as appropriate. The AUGE has also followed up with Xoserve where common issues are found in order to clarify and resolve them.

The AUGE considered that these issues have a greater potential to affect the UG estimate and hence these were prioritised and resolved. Once the methodology for the main elements of UG has been agreed, however, further analysis can be carried out to address specific issues that the industry may identify.

4.6.3 Meter Drift

No further analysis has been undertaken regarding meter drift. As noted in the responses to the 1st draft 2012 AUGS for 2013/14, a much larger sample of meters would need to be analysed than the 25 meter sample quoted in the query received. In addition, a considerable amount of additional data from each would have to be supplied in order for such an analysis to take place.

Where a check read occurs as also noted in the query this does not necessarily mean that the physical meter is drifting, as this can be due to the device that captures the readings not updating correctly. The meter continues to read correctly, but the received information can be out of step. A check read corrects this and the difference would get picked up in RbD. This aspect of the issue is not attributed therefore to meter measurement issues but synchronisation of data issues.

4.6.4 T/P Factors

Information has been requested from Xoserve regarding T/P factors, particularly for DM sites. Xoserve's priority has been to supply all consumption data up to this point, and therefore this data has not yet been received. Analysis will take place when the data is available.

4.7 Consumption Analysis

During 2012 a significant amount of analysis was performed on assessing the potential to use individual metered consumption data to directly estimate the total quantity of UG. If successful, this approach would replace the current indirect method of estimating total UG from RbD. Due to the complexity and sheer scale of the task using individual metered data, an initial feasibility study was carried out on a single LDZ (EA) with the following aims.

- Identify all required data and generate data specification document
- Estimate total data quantity and therefore data storage requirements
- Estimate time taken to obtain data from Xoserve (feeds into delivery plan)
- Estimate processing time required (feeds into delivery plan and together with data storage requirements determines hardware requirements)
- Identify any issues with the consumption based approach, document these issues and address where possible
- Compare consumption based approach with previous RbD-based approach
- Recommend best approach for future calculation of UG

The results of this feasibility study have already been reported in the September 2012 Interim Report [25]. In summary the findings were as follows.

- The methodology developed to estimate total UG based on metered consumption is a more robust approach to estimating total UG. The main reasons for this are
  - It does not rely on RbD
  - It includes SSP assigned UG
  - It does not rely on AQ values (except for validation) and AQs are not used to scale up the non-calculating consuming meter points
  - It less prone to changes in the deeming algorithm or bases of SND, CWV etc.

The results of the analysis for EA were presented to the UNCC at the AUGS technical meeting on 17th September 2012. Feedback was positive, but the industry wanted to see similar analysis on more LDZs, ideally all thirteen. This was felt necessary as there may be issues which affect some LDZs and not others e.g. metering errors in some LDZs. Subsequently, the AUGE requested data for all LDZs from Xoserve and have been running the UG estimation process as data arrives.

Currently the AUGE has processed data for 10 LDZs, the data for the remaining 3 have now arrived and will processed top provide a complete set of results. The methodology has also been updated as new issues have been identified, and to reflect feedback received from the AUGS technical meeting on 17th September 2012 and queries raised through the consultation process. The following sections provide an overview of the consumption analysis, concentrating on issues which arose and how they were overcome. The resulting proposed methodology is then described in Section 6.

The overall concept of calculating total UG using metered data is simple. Total UG is calculated 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. Note that the total UG is estimated for each LDZ and formula year separately.

4.7.1 Calculation of Total NDM Demand

In order to calculate total UG it is necessary to obtain a robust estimate of total NDM demand. UG is then calculated by subtracting the aggregate metered consumption for all meter points from this total. Figure 2 shows how the total NDM demand is calculated. The LDZ allocations as designed to sum to the total LDZ metered demand minus DM demand minus Shrinkage (where LDZ demand is measured at the offtakes). By adjusting the allocations for any meter errors, the total NDM demand (plus any error in the Shrinkage estimate) can be calculated. It is assumed that any significant meter errors (LDZ and DM) are identified and corrections estimated. These meter error corrections have been requested from Xoserve but not received as at the time of writing and so are not included in the results. The only exception to this are three examples where the LDZ meter errors were very large and estimates of their size were taken from the Joint Office website. These are shown in Table 3 below. The errors were split across formula years using the 01B ALPs for the relevant LDZ.

Table 3: Large LDZ Meter Errors (GWh)

LDZ	2009	2010
NT	62.4	10.9
SC	2,572.9	650.1
SO	932.1	228.9

The data requested from Xoserve regarding meter errors (LDZ and DM) comprises LDZ, meter point type, start date of error, end date of error and an estimate of the total value of the error. This will allow the AUGE to apportion the meter errors between formula years. The net effect of all meter errors for each LDZ and formula year can then be removed from the allocation to obtain the total NDM demand (plus Shrinkage error).

The Shrinkage error is unknown and cannot be separated from the total NDM demand. It will therefore be included in the calculated total UG figure and end up in the Balancing Factor.

4.7.2 Calculation of NDM Consumption at Meter Points

Having obtained a best estimate of total NDM demand (including Shrinkage error), it is necessary to subtract the sum of all metered NDM consumptions to obtain an estimate of total UG. This involves calculating the consumption in each formula year for each individual meter point.

4.7.2.1 Calculating Consumption using Meter Reads/Metered Volumes

The first choice for estimating consumption is to use the best available meter reads. The process developed for this by the AUGE is described in detail in Appendix C with a worked example in Appendix D.

One advantage of working at individual meter point level (rather than using an aggregate AQ) is that information specific to each meter point can be utilised. For example, a new meter point (first confirmation date falling part way through a formula year) can have its consumption estimated for the part year for which it was consuming, as the start date is available. Similarly, lost meter points can be accounted for correctly using their end date. The same is also true for larger NDM meter points that switch from being NDM to DM (or vice versa). This can be detected from the complete AQ records for a meter point. When consumption is calculated for a meter point it is only non-zero for periods when the meter point was a 'live' NDM meter point. This should correspond to the meter point's inclusion in the allocation process. This consistency between the period of a meter point having a non-zero consumption and it being included in the allocation process is important given that the UG is estimated by subtracting metered consumption from the allocations.

In order to convert from the metered period to a specific formula year is it necessary to scale based on an appropriate set of factors, i.e. ALP*(1+DAF*EWCF). In order to do this, the most appropriate EUC band for a meter point during a given formula year must be determined. This is complicated by a mismatch between the AQ/EUC process, which is by gas year, and the UG analysis, which is by formula year. Therefore the following decision was taken:

- Look up the first AQ estimate effective after the end of the formula year. If none exists use the latest value.
- This is used to determine the best ALPs to use to profile the demand
    - It is assumed that the AQ after the end of the year best reflects the site's behaviour during that year.
    - If the AQ was updated during the gas year it suggests an improved estimate was made, so we should use the latest estimate from the gas year after the formula year being calculated. The latest record for each year is determined on a meter point-by-meter point basis as part of the pre-processing of the data from Xoserve.
    - Multiple supply point sites have an aggregate EUC allocated to all the meter points. This may represent the aggregate behaviour but it is useful to know the true size of a meter point. This will be used instead of the stated EUC to determine how to estimate the meter point's consumption if the algorithm fails. The EUC band directly corresponding to each meter point AQ record is calculated as part of the pre-processing of the data from Xoserve.

When choosing the meter reads to use in the calculation, it is preferable to cover as much of the formula year in question as possible without covering an excessively long period outside the year during which behaviour may significantly change. The following approach was chosen:

- The meter reads should be at least 120 days apart, cover at least 60 days of the formula year and both lie within 540 days of the ends of the formula year.

Meter changes cause problems with the consumption calculation because closing/opening reads are generally not available. For this reason meters are rejected if meter replacement is detected. Meter replacement information was provided by Xoserve but it was unreliable and the decision was taken not to use it.

Metered volumes for LSP meter points are corrected as part of the reconciliation process and already contain meter unit and temperature/pressure corrections. Therefore it is desirable to make use of these figures where possible. The metered volumes for SSP meter points are not corrected, however, and a significant number of errors were found with this data field during the analysis. The main issue with these errors is not so much their frequency but the large impact they can have. In many cases Meter Index rollover recording issues result in very large positive and/or negative metered volumes.

For example, for the meter reads provided for EM LDZ, 0.88% of metered volumes for SSPs were negative compared to 0.62% for LSP. Also 1.64% of metered volumes for SSPs give metered volumes greater than 5 times the AQ compared to 0.60% for LSPs. Further examples of such errors can be found in Appendix E.

Therefore it was decided to use the meter reads directly to calculate consumption for SSP meter points. This has the advantage that only two good reads are required rather than there being a requirement for all reads between the start and end points to be correct. The temperature/pressure correction for all SSP meter points is the same, so the only other data needed are the Read Units. Xoserve have this data but it is not believed to be reliable, and the analysis showed that there are numerous errors in it. Therefore the decision was taken to estimate Read Units by comparing the meter reads and meter volumes. Whilst doing this, errors can be detected with meter reads which result in negative volumes. These are then classified as Meter Index roll-overs, meter replacements or “bad reads”.

Meter Index roll-over is another issue when calculating metered volumes from raw meter reads. Xoserve once again provided information with the meter reads but it was found to be unreliable (for example, some Meter Index roll-overs are recorded against the incorrect meter read, leading to incorrect calculation of metered volumes if not detected). Therefore the AUGE implemented its own check as described in Appendix C.

Errors in this calculation can arise from a number of sources, as follows:

- Incorrect meter reads
- Incorrect metered volumes
- Incorrectly estimated meter units
- Undetected meter replacement
- Incorrect meter index roll-over detection

As a result it is necessary to validate the calculated consumptions. Without this validation, some incorrect extremely large values can skew the results. Therefore the consumption estimate (where successfully calculated) was compared with an estimate based on the AQ selected as part of the algorithm for each meter point. If the consumption estimate was greater than 73,200kWh and greater than five times the AQ-based estimate it was rejected. This process protects against very large errors but does not affect small sites which may vary significantly without impacting the overall result.

4.7.2.2 Scaling Up

The validation checks performed as part of the consumption algorithm result in a number of meter points failing the consumption calculation process. Meter read information is also unavailable for meters in CSEPs so these cannot have calculated consumptions.

Figure 13 shows the percentage of meter points (excluding CSEPs) which fail the various checks for all data processed so far.

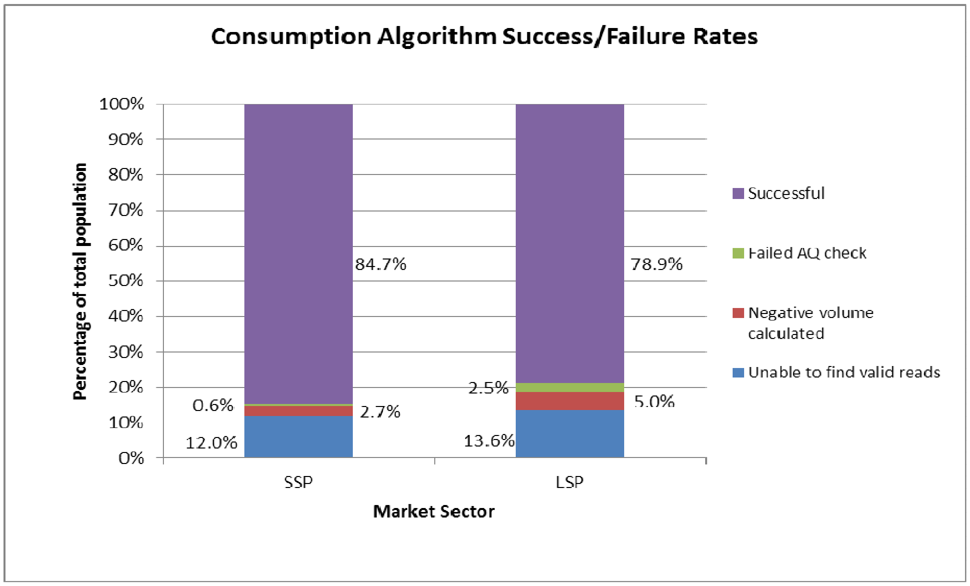


Figure 13: Consumption Algorithm Success/Failure Rates

In order to correctly calculate the sum of all metered NDM consumptions, an estimate of the consumption for non-calculating meter points (including those in CSEPs) needs to be made. As discussed in the interim report [25], the use of AQs to scale up the failed consumptions was rejected as the AQs do not necessarily represent he consumption for the period we are trying to estimate.

The approach chosen is therefore to set the consumption for each failed meter point to the average consumption for the relevant EUC by LDZ and formula year (as calculated from the successful meter points).

There are two issues to consider with this approach, which are described below

Non-Consuming Sites

Non-consuming sites can be divided into meter points with AQ=1 and those with AQ>1, which are expected to behave differently.

Table 4 below shows how the proportion of meter points with AQ=1 varies between those meter points in the sample and those outside it. This shows that there are 3-5 times as many meters with AQ=1 in the sample of non-calculating meters than in the sample of successfully calculated meters.

Table 4: AQ Split Comparison Between Successful and Unsuccessful Meter Points

	Successfully Calculated	Number AQ=1	Number AQ>1	Total	% AQ=1
2009	N	74,345	2,093,004	2,167,349	3.4%
2009	Y	152,442	13,269,417	13,421,859	1.1%
2010	N	64,943	1,623,885	1,688,828	3.8%
2010	Y	151,762	13,768,912	13,920,674	1.1%
2011	N	66,194	1,240,724	1,306,918	5.1%
2011	Y	143,384	14,184,039	14,327,423	1.0%

Table 5 shows the proportion of meter points in each of the two AQ groups that are actually consuming (according to their meter read data). By definition this is based on successfully calculated meters only.

Table 5: Proportion of Meter Points Consuming

	Consumption>0 and AQ=1	AQ=1	Consumption=0 and AQ>1	AQ>1	% AQ=1 Consuming	% AQ>1 not Consuming
2009	43,742	152,442	122,397	13,269,417	28.7%	0.9%
2010	43,378	151,762	122,681	13,768,912	28.6%	0.9%
2011	50,694	143,384	135,784	14,184,039	35.4%	1.0%

We can therefore apply these proportions to the failed meters (i.e. those for which the consumption calculation was unsuccessful), thereby taking into account the number and behaviour of sites with AQ=1 and those with AQ>1 outside the sample.

As well as AQ, another indicator of potential non-consumption is meter reads stopping for an extended period. This may indicate that the property is vacant and not accessible for meter reads.

The number of meters with no meter reads after the start of the formula year is higher at the end of the data set as infrequently metered sites are awaiting a new read. For the 10 LDZs currently processed the numbers of such sites are as follows:

Table 6: Number of Meter Points Without Reads

Year	Number of Meters
2009	10,704
2010	51,281
2011	280,182

We have no information about whether the meter points with no reads are consuming or not, but it is likely that the proportion not consuming is higher than the general population. We can quantify the possible effect by comparing the UG estimates under two scenarios: all such meters consuming vs all such meters non-consuming. A sensitivity analysis has therefore been carried out to determine the potential error that this could introduce into the calculated UG figure, and this is described in Section 4.7.4.

Sample used to Calculate each Average

Where the average for an EUC is calculated based on a very small sample it may be unreliable. Therefore the decision has been taken to use the national rather than LDZ average EUC consumption for those cases where the sample size is less than 30 meter points. This only affects EUC Groups 07B, 08B and 09B.

4.7.2.3 Summary

Figure 14 below summarises how the consumption value is obtained for each meter point.

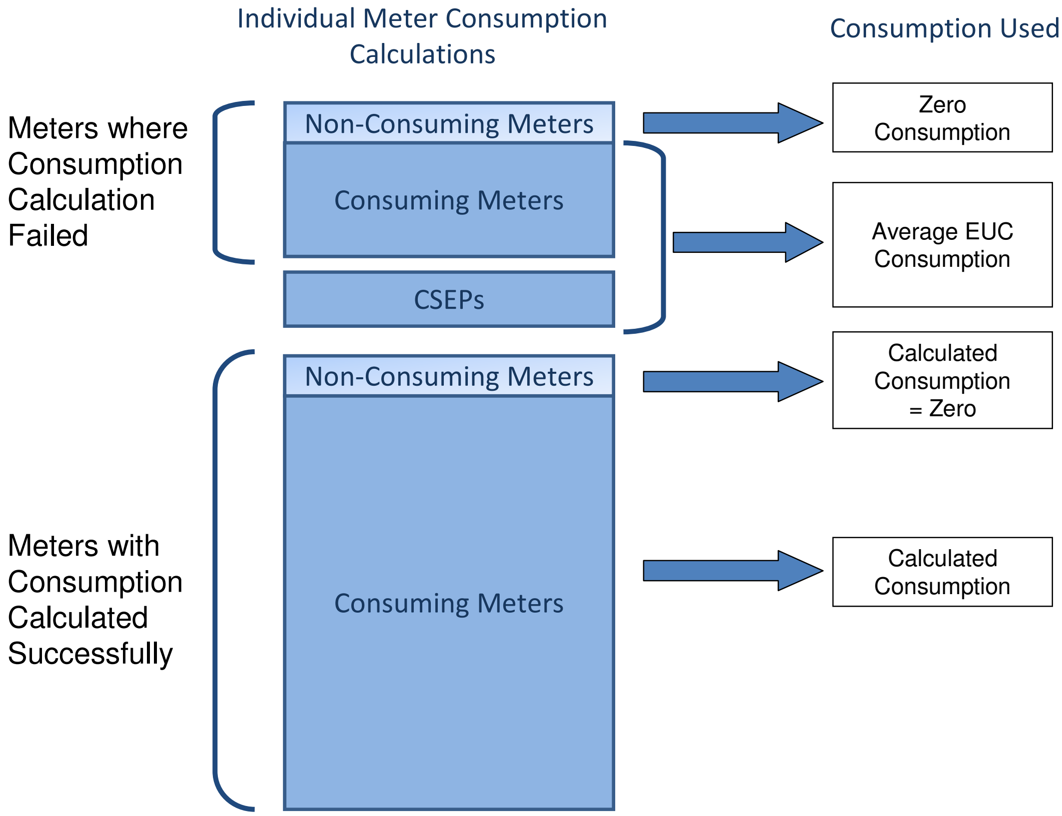


Figure 14: Building up Consumptions

4.7.2.4 Summary of Data Issues

During development of the consumption methodology a number of data-related issues have been identified. Some of these relate to raw data provided initially by Xoserve, e.g. incorrect data provided or misunderstanding of what the data represents. These have been clarified and resolved where possible. Table 7 is a summary of the issues which have been identified and raised with Xoserve along with the outcomes.

Table 7: Consumption Analysis Data Issues

Issue	Outcome
Multiple meter reads for same meter point and same day with different values	These are usually amendments. The AUGE requires only a single “best value”. However, no timestamp is currently available for each record so it is difficult for Xoserve to determine the most recent value. Xoserve have developed a process to determine the best read based on read type code. This removes most cases of multiple reads. The remaining cases are kept separate and assessed by the AUGE. In most cases these meter points are excluded but if in a high EUC band then manual selection of most appropriate meter read may occur. The Read Type code was added to data

Issue	Outcome
	specification for AUGE information. As a general point, it has been very difficult and in some cases impossible to work out what happened when, as the timestamps of when data was entered or changed may not always be captured or are not directly accessible. If timestamps are captured or provided of what changed when (and even by whom) it would greatly improve the tracking and extraction of data pertinent to the estimation of UG. For example to pick up metered volume corrections it may be necessary to obtain the last 4-5 years of consumption data every year which would be a significant task. This is an area that will require further investigations going forward with Xoserve.
EUC group not consistent with AQ	The data was found to be correct. Where multiple meter points are part of the same supply point, the EUC is calculated for the supply point i.e. based on aggregate AQ across all meter points. This may not reflect the EUC group for all meter points that are part of that supply point as some may have different behaviours. A Meter point specific EUC band is calculated by the AUGE as part of the data pre-processing.
Some high consumption meter points affecting consumption estimates	Some DM meter points and unique sites were incorrectly present in the dataset. The data specification was updated to include a site type flag so meter points that switch between NDM/DM/Unique can be handled correctly. The site type flag is used to identify sites which change and dates are manually merged with the new lost sites data.
Unreliable 'Round the Clock' Meter Index Indicator	The AUGE was expecting a Y/N indicator but received a value. Xoserve confirmed that this should represent the number of times the meter index has gone through zero since the last meter read. This has been found to be unreliable, with numerous cases where it is incorrect, including negative values and values suggesting the meter index has passed through zero several times. For example, in EA LDZ this varies from -1 to +5. This is the correct data from the shipper held by Xoserve. This value is therefore not used in consumption methodology
Inconsistent meter asset information	This includes inconsistent numbers of dials, imperial indicator and units for the same meter. This is the data provided by Shippers and held by Xoserve. As a result this information is not used and: • Read Units are calculated by the AUGE based on meter reads and metered volume. Where the Read Units are incorrect, this will be by a factor of 0.1, 10, 100 etc. so the calculated consumption will usually fail validation checks • The Imperial indicator is taken from the associated meter read rather than asset information. This may be incorrect but it is the best data available

Issue	Outcome
Inconsistent AQ values between AQ data table (AQ by formula year) and AQ column against individual meter reads (AQ at time of read)	The data specification was updated so that the AUGE now receives a full history of all AQ changes for all meter points.
Metered volumes inconsistent with meter reads	Xoserve advised that metered volumes are corrected if they are wrong but that the actual meter reads are never updated. This is the case for LSPs only (this is part of the reconciliation process). The consumption calculation method was therefore updated for LSPs to use metered volume rather than meter reads in order to include these corrections. However, there are instances where these metered volumes are clearly incorrect and these are rejected accordingly. For SSP meter points, the metered volumes do not get corrected and as described in this document the meter reads are used in preference to the metered volumes as this was found to be more reliable.
Meter points changing LDZ	Some meter points are mis-classified and later moved to the correct LDZ i.e. some meter reads are associated with one LDZ and then later meter reads are associated with a different LDZ. Sites are assigned to the LDZ they were allocated to at the time as they will have formed part of the allocation process for that LDZ.
Missing meter points	Initially, meter points with no meter reads and meter points which are no longer active were not included in the data set. The data collation process has been updated to include these.

In addition to the above there are many issues relating to data quality which cannot be resolved and which are simply the best data held in Xoserve's systems. Some examples of these data quality issues are given in Appendix E.

4.7.2.5 Practical Considerations

The consumption-based approach to estimating total UG is a complex task using very large quantities of data. Xoserve have so far provided ~170 million meter read records, and the consumption database takes up ~60GB of Oracle tablespace. There are a number of challenges associated with this. In order to ensure that this approach is practical, significant effort has been focussed on developing an optimal process.

Detailed data specifications have been developed in conjunction with Xoserve to ensure that the correct data is provided in the correct and most suitable format for the AUGE. This includes some pre-processing and data validation by Xoserve. Given the size of the datasets involved it is vital to minimise any need for data resends. Having received the data, a process has been developed to import the raw data into an Oracle database and perform some pre-processing and validation on the data. This requires approximately 24 hours per LDZ.

Once the data has been imported and pre-processed, the consumption calculations can be run. Initially this was taking too long to be practical, but by optimising the code and database, the processing time has been reduced to around 10 hours per formula year for an average LDZ.

Overall, the whole process now takes between 2 to 4 days per LDZ and the overall storage requirements are expected to be approximately 90GB. An export of the database is expected to be 25Gb in size once all the LDZ data has been processed which in turn can be compressed to about 6Gb. The AUGE intends to prepare smaller multiple exports such that Code Parties can examine a subset of LDZs without having to import all the data at once.

4.7.3 Results

Table 8 below shows the interim total UG figures calculated for 10 LDZs using the consumption method. These figures are **not** final estimates of UG and are subject to a number of conditions, as follows:

- Adjustments for meter errors have been made only for the known large issues in SC, NT and SO LDZs, data for which was retrieved from the Joint Office website.
- An approximate adjustment for temporary UG has been made based on calculated temporary UG from the 2011 analysis.
- The figures have been adjusted for detected theft.
- SC LDZ includes the Scottish Independents.

UG for each LDZ is expressed both as an energy value in GWh and as a percentage of throughput for the LDZ in question.

Table 8: Initial Total UG Estimates Using Consumption Method

LDZ	Total UG (GWh)			Total UG as Percentage of NDM Allocations		
	2009	2010	2011	2009	2010	2011
EA	768	1,576	149	1.93%	3.84%	0.44%
EM	1,028	1,414	-620	2.05%	2.74%	-1.44%
NE						
NO	510	839	-30	1.93%	3.14%	-0.13%
NT						
NW	1,077	1,134	-985	1.78%	1.87%	-1.93%
SC	1,343	1,602	-193	2.77%	3.44%	-0.48%
SE						
SO	596	955	186	1.68%	2.65%	0.62%
SW	490	571	-105	1.70%	1.96%	-0.43%
WM	243	811	-372	0.54%	1.78%	-0.98%
WN	211	225	-29	3.95%	4.18%	-0.64%
WS	677	633	-169	3.70%	3.50%	-1.13%

The following observations are made about these interim figures:

- For 2009 and 2010, the interim UG figures are consistent and universally positive.
- The figures for 2009 and 2010 are of a similar magnitude to those presented in the 2011 AUGS for 2012/13, but are in general somewhat higher. This is not a strict like-for-like comparison, however, due to the fact that meter errors and temporary UG have not been fully accounted for in the interim figures in Table 8. The effect of removing temporary UG will always be to reduce the interim figures, whilst meter errors could act in either direction depending on the nature of the error.
- There is a step change in the UG figures from 2009 and 2010, which are generally consistent, to 2011, which is often negative. Only two LDZs of the 10 that have been analysed to date returned a positive UG figure for this year.

This step change in 2011 is likely to be due to a lack of corrections in the LSP consumption dataset for this year. There are significantly fewer corrections in the 2011 dataset at this stage due to the smaller time that has elapsed since the end of this year ending 31st March 2012. This means that a number of errors that will be detected and corrected in the future have not yet been applied, and these errors are feeding into the current interim 2011 figure.

As described in Appendix C, the consumption method uses raw meter reads for EUC 01B and Xoserve's recorded metered volumes for all other EUC groups. The reason for this is that when corrections are made by the Shippers to the data held by Xoserve, these are made to the metered volumes but not the meter reads, and hence this field must be used if the corrections are to be picked up in the analysis. Corrections are not made to SSP metered volumes, however.

It is also possible that undetected LDZ meter errors have resulted in a reduction in allocations – however, the reduction in allocations observed is consistent across all LDZs so this is unlikely as the LDZ meter errors would have to have had a similar effect on almost every LDZ.

It is therefore proposed that in order to allow sufficient time to elapse for errors to be corrected before the figures are used to calculate UG, 2011 is dropped from the current analysis and the estimate of UG is based on the figures from 2009 and 2010.

4.7.4 Sensitivity Analysis

The estimation of UG based on meter reads and metered volumes is subject to a number of sources of error. In order to understand the potential impact on accuracy of the UG estimate, a number of sensitivity analyses were carried out.

4.7.4.1 Sample Size

The accuracy of the total UG figure will be dependent on the size of the sample used to estimate it. Rather than selecting a fixed-size sample of meter points and scaling up, the AUGE has taken the approach of using as many meter points as possible in order to minimise the error due to sampling. The sample size is still not 100% of the population, however, because consumption cannot be calculated for all meter points.

As Figure 13 shows, the overall success rate is currently 84.6%. This is the maximum sample size that can be achieved given the quality of meter read information available at present.

In order to understand the sensitivity to the sample size chosen, the AUGE has carried out a small Monte-Carlo analysis based on taking random samples of different sizes from the maximum sample set for one LDZ. Sample sizes of 25%, 50% and 75% were taken and compared to the 100% case (note that this is the maximum sample size of the 84.6% of meter points for the sample LDZ, not 100% of all meter points). Thirty sets of random samples were taken in each case and the average, minimum and maximum values are shown in Figure 15.

The average value of the total UG calculated using each sample size remains relatively constant which provides confidence in the estimate and the method used to calculate it. However, the minimum and maximum values from each sample vary significantly with sample size, with a wide variation for the 25% sample that decreases in magnitude with sample size. Given the size of this variation compared to the level of UG the AUGE recommends using the full dataset rather than a fixed sample.

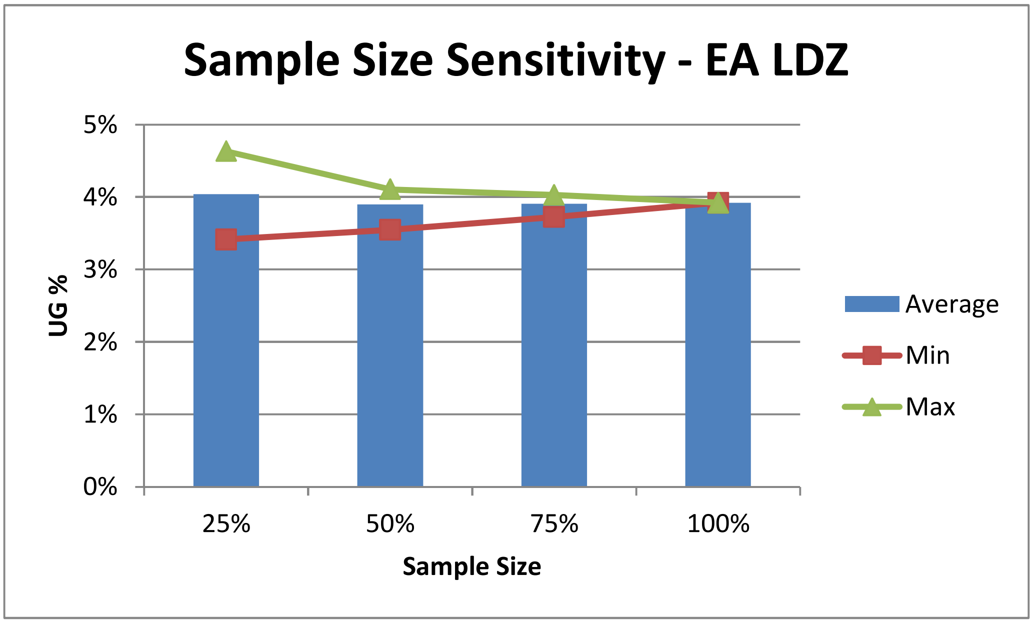


Figure 15: Sample Size Sensitivity for EA LDZ for Formula Year 2010-11

4.7.4.2 Assumptions Regarding Non-Consuming Meter Points

The sensitivity analysis based on sample size looked at the variability caused by using only a subset of all available meter points in the UG calculation and scaling up to cover the full population. Within this scaling up process there is an assumption about the proportion of meter points in the population outside the sample that are not consuming. This proportion is based on the AUGEs best estimate as described in Section 4.7.2.2.

Table 9 below shows the top level UG (for EA LDZ in 2009) comparing the AUGEs best estimate along with high and low scenarios. They show a spread of ±0.2% in the UG figure.

Table 9 Sensitivity to Assumptions Regarding Non-Consuming Meter Points

	% of failed population with AQ=1 who are consuming	% of failed population with AQ>1 who are consuming	Overall % of failed population who are consuming	UG (GWh)	UG (%)
Low Estimate	57.4%	100.0%	98.9%	750.70	1.9%
Best Estimate	28.7%	99.1%	97.3%	825.55	2.1%
High Estimate	0.0%	98.2%	95.6%	900.39	2.3%

4.7.4.3 Sample Size for High EUCs

The small sample sizes in the larger EUC groups leads to these requiring specific attention. For example, if an EUC band has only two meter points for a given LDZ and one is unable to have a consumption calculated, then the other will be used to give the same consumption (i.e. the average consumption of the EUC band will be based on the one available meter point). This can lead to unreliable results, and the rule of thumb for these situations is that a minimum sample size of 30 is required. The methodology therefore calculates average consumption by EUC band across all LDZs and uses this where there are less than 30 meter points in an EUC band for any given LDZ. The variation in averages across LDZs is shown in Appendix F, which shows the average consumption by EUC for individual LDZs compared to the national values. Choosing a sample size of 30 means that only EUCs 07B, 08B and 09B are affected. The impact on the UG estimate for any LDZ is up to ±0.2% of the total allocation.

4.7.4.4 AQ Check

Figure 16 below shows for (10 LDZs) meter points which fail the AQ check, the distribution of their consumptions as a multiple of their AQ. The long tail on the chart has been excluded, but a future 24,320 meter points have calculated consumptions greater than 200 times their AQ. So the majority of calculated meter point consumptions that fail the AQ check are very much larger than the meter point's AQ. A small change in the rejection criteria to 10*AQ would result in the failure rate dropping to 0.56% instead of 0.66%.

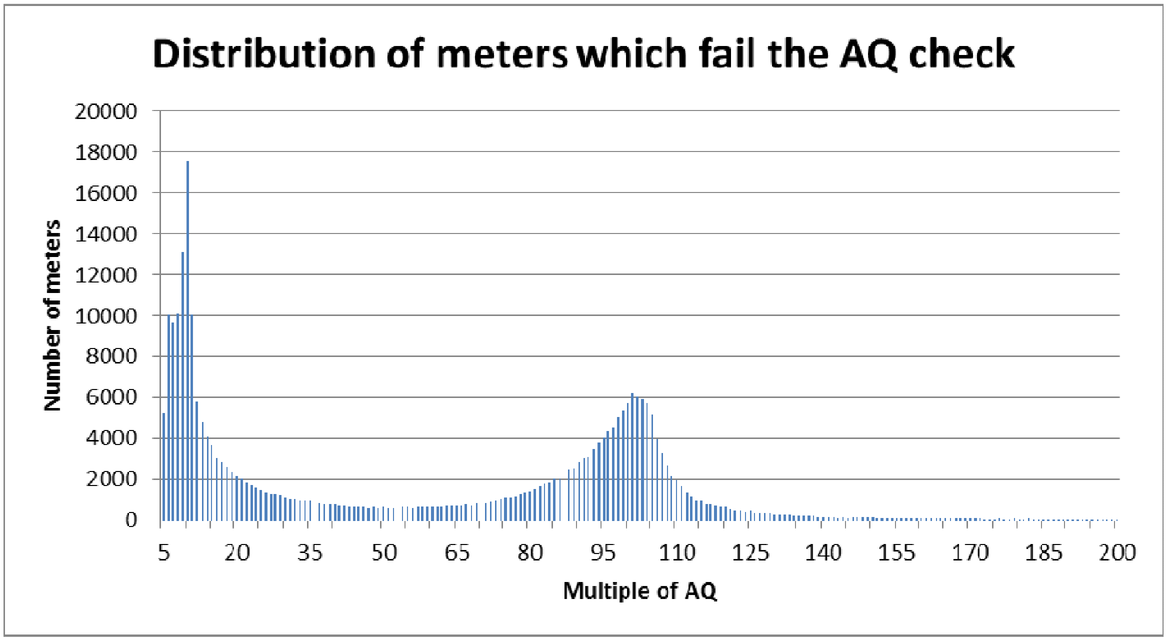


Figure 16: Distribution of Meter Points which Fail the AQ Check

4.7.4.5 Read Units

There is uncertainty in the Read Units provided by Xoserve (who indicated that the data was of poor quality). The AUGE therefore used a process to infer the units from the meter reads and metered volumes (assumes that the metered volumes are correct even if the asset information is incorrect). Comparing the Read Units for EA LDZ provided by Xoserve and those calculated by the AUGE during data pre-processing agreed for 96.1% of meter points.

Meter points with very large calculated consumptions due to incorrect meter units will fail the 5*AQ check, but some meter points will have very small calculated consumptions that are not rejected.

From the bunching around 10 and 100 that can be seen in Figure 16 above, it appears that a large proportion of AQ check failures are due to incorrect meter units.

4.7.5 Method Comparison

In order to compare the old UG method used in the 2011 AUGS for 2012/13 with the consumption method, two different methods are used:

- A comparison of the impact of different error sources on each method
- An assessment of the relative variability (i.e. Confidence Interval width) within each method

These analyses are presented in sections 4.7.5.2 and 4.7.5.3 respectively. Section 4.7.5.1 is a comparison of calculated UG values compared to the final UG figures used last year.

4.7.5.1 Results Comparison

Table 10 provides a comparison of the estimate of total UG from the consumption method against the top level average UG from the 2011 AUGS for 2012/13. This includes corresponding UG as a percentage of throughput.

There are differences between the results which are explained in the following sections and this should be expected especially where improvements have been made over the previous RbD based method.
Once the remaining LDZs have been processed a total estimate of UG can be calculated.

Table 10: Unidentified Gas Comparison

LDZ	Total UG (GWh)				Total UG as Percentage of NDM Allocations			
	2009	2010	2011	AUGS 2011	2009	2010	2011	AUGS 2011
EA	768	1,576	149	841	1.93%	3.84%	0.44%	2.47%
EM	1,028	1,414	-620	587	2.05%	2.74%	-1.44%	1.36%
NE				467				1.71%
NO	510	839	-30	477	1.93%	3.14%	-0.13%	2.10%
NT				779				1.68%
NW	1,077	1,134	-985	272	1.78%	1.87%	-1.93%	0.53%
SC	1,343	1,602	-193	216	2.77%	3.44%	-0.48%	0.54%
SE				521				1.19%
SO	596	955	186	484	1.68%	2.65%	0.62%	1.62%
SW	490	571	-105	393	1.70%	1.96%	-0.43%	1.63%
WM	243	811	-372	649	0.54%	1.78%	-0.98%	1.71%
WN	211	225	-29	53	3.95%	4.18%	-0.64%	1.17%
WS	677	633	-169	291	3.70%	3.50%	-1.13%	1.93%

4.7.5.2 Impact of Error Sources on each Method

Table 11 contains a list of the main sources of error which affect the calculation of UG and provides an overview of the impact of each error source on the previous RbD based methodology and the proposed consumption based approach. Two key weaknesses of the RbD based approach are that it does not estimate SSP-assigned UG and that the AQ values used are not optimal for the period of calculation of UG (referred to as “Lagged AQ” in the table). More detail on the AQ timing issues follows.

One benefit of the proposed consumption-based methodology over the previous RbD based approach is that the meter read data used is more aligned to the UG calculation period. This is shown in Figure 17.

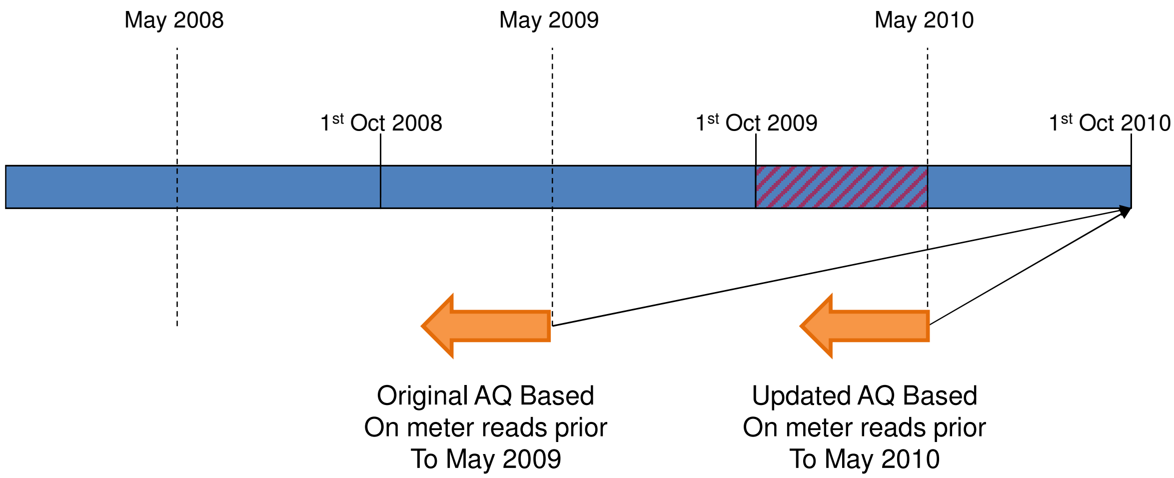


Figure 17: Comparison of Meter Read and UG Calculation Period Alignment

The allocation process for gas year 2009 uses AQs calculated based on meter reads covering a period of time up to May 2009. The exact period is determined by the dates of meter reads. As shown in Figure 17, these AQs are not representative of the period over which they are used (Oct 2009 – Oct 2010). The RbD based approach tried to account for the bias resulting from the lagged nature of AQ values by using the AQ resulting from the following AQ review. This improves the situation as there is now the potential for some degree of overlap between the meter reads and the period for which we have allocations and are attempting to calculate UG (shaded area in Figure 17 shows maximum potential overlap). The proposed consumption-based approach actually attempts to use meter reads as close as possible to the start and end of the period for which UG is being calculated, and should therefore give a more representative estimate of consumption than the AQs.

Figure 18 shows an example comparison of AQ values and consumption values (adjusted back to seasonal normal for comparison with AQs). Three example LDZs are shown. In all three cases, the original AQ used in the 2009 allocations is much higher than the AQ calculated at the subsequent AQ review (2010 AQ). The AQ calculated from the consumption data (Cons AQ) is based on meter reads spanning the period in question and lies between the other two AQ values. This suggests that the updated AQ values used in the RbD based methodology were too low, which would result in an over-estimate of model bias and a consequent under-estimate in UG. Note that the AQ is not a total for the whole LDZ. Because consumption can only be calculated for a sample (albeit a large one) of all meters, the AQs are based on this subset of meter points only.

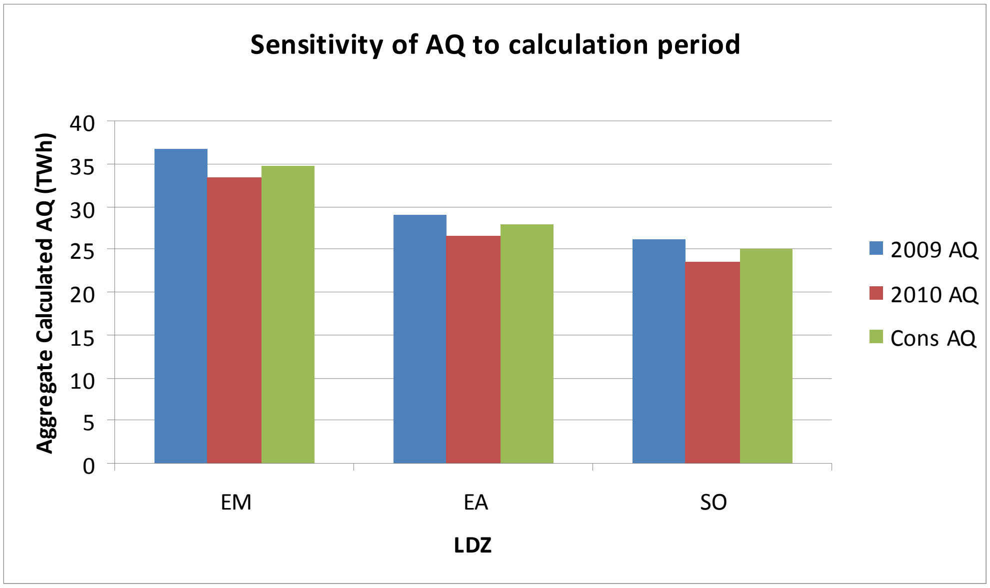


Figure 18: Sensitivity of AQ to calculation period

Table 11: Error Source Impact

Error Source	Description	RbD Based Method Impact	Consumption Method Impact
Meter Read Errors	Incorrect meter reads on Xoserve system	Incorrect AQs or AQ roll-over. AQ may be subject to appeal/amendment.	Error in calculated consumption or calculation failure if error is large.
Metered Volume Errors	Incorrect metered volumes on Xoserve system	Incorrect AQs or AQ roll-over. AQ may be subject to appeal/amendment.	Error in calculated consumption or calculation failure for LSP if error is large (SSP do not use metered volume).
Incorrect AQ (including rollover)	AQ calculated based on poor meter reads but passes AQ validation process	Error in model bias estimate & therefore UG.	The Consumption method does not use AQ, so AQs that have rolled over and may not represent the meters consumption in the RbD based method do not affect the consumption calculation method. There may be a minor impact due to use of AQ for validation but this will simply result in more calculation failures. Issues due to meter read errors are covered in the first line of this table which of course affects both AQs and consumptions.
Lagged AQ	AQ is based on historical meter reads which are not representative of year in question. Note that even currAQ from Mod81 report will be based on lagged meter reads	Error in model bias estimate & therefore UG.	Consumption calculated over period of interest using appropriate meter reads so no impact on consumption estimates. May be minor impact due to use of AQ for validation but this will simply result in more calculation failures.
Poor Initial	New meter points have	AQ bias will affect calculation of	No impact on overall UG

Error Source	Description	RbD Based Method Impact	Consumption Method Impact
AQ Estimates	estimated AQ which is biased	model bias and therefore UG. Split of UG by market sector also affected due to effect on Unregistered sites.	estimate as AQ not used. Split of UG by market sector is affected due to effect on Unregistered sites.
RbD 5 year window	RbD in gas year not consistent with calculated allocations for that gas year as corrections at any point in time could be for somewhere in previous 5 year window. Assume cancelling out over a number of years.	If assumption of averaging out over years is valid then this adds noise to each year's estimate. Unlikely that assumption is 100% correct so some error will be present in RbD average which results in an error in UG estimate.	Does not rely on RbD. However, corrected metered volumes will not be picked up unless the data is updated to include these corrections.
Non-consuming meter points	This covers meter points which have AQ > 1 but which are not consuming and meter points which have AQ = 1 but which are consuming	Model bias based on AQ so will not account for this unless AQ changes in Mod81 report. For LSP meter points, this will be accounted for in RbD. Therefore there will be an inconsistency between model bias and RbD which will affect UG calculation.	Non-consuming meter points are accounted for in the subset of meter points where calculation of consumption is successful. Estimate of non-consuming sites made from calculated data and applied to meter points which failed to calculate. Some error in consumption of non-calculated meter points will be present.
Low AQ (incl AQ=1)	Many meter points with low AQ values which may or may not be consuming	Model bias based on AQ so assumes that these meter points are consuming very small amounts. This may or may not be the case.	Analysis of meter points with low AQ has determined what proportion of these are consuming.
Limited CSEPs data	Only have a single snapshot of CSEPs data for whole formula year. This includes number of meter points and aggregate AQ only. It is therefore assumed that all of these meter points are active and consuming for the whole year	AQ of CSEPs is included in model bias calculation. AQ is scaled using Mod81 data on the assumption that CSEPs behave in the same way as the rest of the LDZ.	CSEPs are treated a 'non-calcs' so AQ are not used. Each meter point is given average consumption for that EUC band.
SSP Assigned UG	UG is initially allocated by the deeming algorithm. Previously the AUGE assumed no UG was allocated to SSP in this way as a result of larger AQ biases for LSP meter points than SSP meter points. There is now evidence that this bias no longer exists.	SSP-assigned UG assumed zero previously. This is no longer the case and will lead to a significant under-estimate of UG if ignored.	Method inherently includes all UG regardless of how it was initially allocated.
LDZ Meter Errors	LDZ metering errors which have been identified and investigated	Once meter error investigation is closed, correction applied through RbD. Correction applied at time of closure but covers historical period so corrections do not align with when error occurred. Method assumes this lag effect averages out over time.	Corrections applied to allocations for all errors which have been identified and intersect the formula year in question. This includes open investigations where a suitable estimate of the error is available. Error is

Error Source	Description	RbD Based Method Impact	Consumption Method Impact
		Correction applied via RbD will only be for part of error deemed to have occurred within reconciliation window so may not include whole error	apportioned to correct formula year (or part year) based on start/end dates and deeming algorithm parameters
DM Meter Errors	DM metering errors which have been identified and investigated	Treated in same manner as LDZ Meter Errors (see above)	Treated in same manner as LDZ Meter Errors (see above)
Profiling to year	Deeming algorithm parameters are used to profile data to align with year boundaries	Model bias calculation relies on AQ values which have been derived using WAALPs. Algorithm parameters are used to calculate allocation values.	Consumption calculation uses deeming algorithm parameters to derive consumption by formula year.
LDZ Switches	Meter points that are assigned to the wrong LDZ in one year are correct in the subsequent year	Mod81 prevAQ will be incorrect, resulting in incorrect model bias calculation. This should balance out across all LDZs.	Calculated consumption will be in wrong LDZ and will affect UG calculation. This should balance out across all LDZs.

4.7.5.3 Confidence Interval Comparison

For each estimate of Unidentified Gas made using the consumption method, an accompanying Confidence Interval can be calculated directly. This is because consumptions for each individual site making up the sample are used in this analysis, which means the sample Standard Deviation can be calculated in each case, which in turn allows the Standard Error of the final UG estimate to also be produced. The Standard Error calculation is complicated by the stratified nature of the sample and the requirement for a Finite Population Correction, and the procedure is described in full in Section 6.1.5 later in this document.

It is not possible to calculate Confidence Intervals for the RbD based method directly because this method is based around RbD and algorithm bias and would require the algorithm bias calculations to be carried out on a meter point-by-meter point basis. The Confidence Interval for this method can be estimated, however, based on the CI for the consumption method and the relative sample sizes used by each. This allows an approximate comparison of the CIs from the two methods to be made.

This approach relies on the fact that the RbD based method and the consumption method both use consumptions based on meter reads in one form or another.

RbD based method
Meter Reads → AQs → Algorithm Bias (LSP and SSP)

Consumption method
Meter Reads → Calculated consumptions

As stated in the 2011 AUGS for 2012/13 [20], the RbD based method uses approximately 76% of the LSP population and 85% of the SSP population in the algorithm bias calculations. Precise year-on-year figures across the combined population are given in Table 12 below.

Given that we cannot calculate algorithm bias on a meter point-by-meter point basis, which we would need do to in order to directly calculate the CI for that method, an approximation is required. This approximation is based on the fact that the two methods use different sizes of samples drawn from the same population. This can be used to calculate the relative size of the CIs that result from both the RbD based and the consumption methods using the formula for the Finite Population Correction. The Finite Population Correction is a multiplicative factor that is applied in the calculation of each CI limit, and therefore the ratio of the Finite Population Corrections from each method give an approximation of the ratio of the widths of the CIs produced by the two different UG calculation methods.

Table 12: Comparison of Sample Size and Finite Population Correction (FPC)

LDZ	2009/10 Sampling %	2009/10 FPC	2010/11 Sampling %	2010/11 FPC	2011/12 Sampling %	2011/12 FPC
EA	87.9%	0.35	92.6%	0.27	93.4%	0.26
EM	86.5%	0.37	90.9%	0.30	91.9%	0.28
NE						
NO	85.8%	0.38	90.7%	0.30	91.7%	0.29
NT						
NW	85.6%	0.38	81.0%	0.44	90.8%	0.30
SC	86.2%	0.37	90.4%	0.31	91.0%	0.30
SE						
SO	87.2%	0.36	90.8%	0.30	92.4%	0.28
SW	85.9%	0.38	91.8%	0.29	92.3%	0.28
WM	86.1%	0.37	90.8%	0.30	91.6%	0.29
WN	86.6%	0.37	90.8%	0.30	92.5%	0.27
WS	82.8%	0.41	89.0%	0.33	91.0%	0.30
Average (Consumption Method)	**86.1%**	**0.37**	**89.9%**	**0.32**	**91.9%**	**0.29**
RbD Bias Method	**84.9%**	**0.39**	**86.4%**	**0.37**	**88.3%**	**0.34**

Note that in this table, the sampling proportions quoted for the consumption method include non-consuming sites. This is to ensure a like-for-like comparison with the proportion of sites used in algorithm bias calculations for the RbD based method.

This table shows that for each year under consideration, the Finite Population Corrections from both methods are similar and will lead to CIs of comparable width. In each case the Confidence Interval from the consumption method will be slightly smaller.

As noted in Section 6.1.5, these Confidence Intervals represent the variability introduced by the sample size only. Other sources of variation, particularly the uncertainty in the meter reads themselves, are not accounted for. This provides a fair comparison because it is the relative capability of the *methods* that is critical here, and this is the variability cause by the sample size. Uncertainty in the meter readings affects both methods (directly for the consumption method and via AQ for the RbD based method) and hence can be discounted from a like-for-like comparison of the two techniques.

4.7.6 Conclusions

The decision between the two methods is based on an assessment of the likely accuracy of those methods (including susceptibility to the error sources listed in Table 10 and Table 7 above) and the variability inherent within the estimates produced.

It is clear from Table 11 that whilst the consumption method is still affected by some potential data issues, these are minimised by working with directly-calculated consumption values for all market sectors wherever possible. By contrast, the RbD based method makes many more assumptions and hence contains more potential error sources. These error sources, which have either no impact or a limited impact on the consumption method include:

1. Inconsistencies in RbD due to retrospective corrections and close-out period end.
2. AQ issues.
3. Non-consuming meter points.

In addition to these, the single biggest issue with the RbD based method is its inherent assumption that SSP-assigned UG is negligible.

Each of these elements can and will introduce systematic errors into the RbD based method without necessarily increasing the variability of the estimate, whilst having little impact on the consumption method.

Based on the above analysis, the AUGE believes that the consumption method is more accurate by design. Having considered all the information and results of the analyses carried out, the AUGE recommends that the consumption method should be used to calculate total UG, and this is the approach used throughout the remainder of this document.

4.8 Theft Analysis

In the 2011 AUGS for 2012/13 [20] the split of theft between the LSP and SSP market sectors was calculated using pre-theft AQ as the best estimate of the true AQ of each site during the period of theft. This was chosen in preference to other methods (such as using current AQ or post-theft AQ with a theft estimate added in) because post-theft AQ may or may not reflect theft-affected consumption, and current AQ does not necessarily reflect consumption levels for thefts that occurred several years ago. The provision of consumption data, in conjunction with the existing theft estimates, allows the AUGE to make a direct and independent calculation of the AQ during the period of theft for each site where valid meter reads are present, which could lead to a more accurate assignment of each site to market sector.

The full analysis of potential alternative methods for splitting theft by market sector is given in the September 2012 Interim Report [25]. Updates to the analysis and key points are provided in this document.

4.8.1 Inclusions and Exclusions

It is important that only those sites that satisfy the conditions for contributing to Unidentified Gas are included in the theft analysis. The UNC contains definitions of what types of theft are defined as Shrinkage and which are not, and only those thefts that do not contribute to Shrinkage fall into the Balancing Factor element of UG.

The broad brush classification of theft into UG or Shrinkage is that Shipper-responsible theft (i.e. theft downstream of the emergency control valve) falls into UG, whilst Transporter-responsible theft (i.e. theft upstream of the emergency control valve) falls into Shrinkage. There are more detailed rules concerning Unregistered and Shipperless sites, however, which mean that thefts from such sites are defined as Shrinkage no matter where they occur.

The relevant sections that create this definition come from Section N of the Uniform Network Code [Ref 32] and the Gas Act 1986 [Ref 33].

Paragraph 1.3.2 of Section N of the UNC states the following:

Shrinkage in a System shall:

a. *include gas offtaken from the System which has been illegally taken:*
    i) *upstream of the point of offtake (in accordance with Section J3.7) at any System Exit Point (it being recognised the effect of Standard Condition 7(3) of the Transporter's Licence is that the rates of Transportation Charges may reflect the taking of such gas); and*
    ii) *subject to paragraph (b)(ii), at or at a point downstream of the point of offtake at a System Exit Point, in a case in which the Transporter is (pursuant to paragraph 9(2) of the Gas Code) entitled to recover the value of the gas;*
b. *not include gas offtaken from the System:*

i) *except as provided in paragraph (a)(ii), illegally taken at or downstream of the point of offtake at any System Exit Point (but without prejudice to Section E3.5.2 or to any reduction of Transportation Charges pursuant to Standard Condition 7(3) of the Transporter's Licence); and*

ii) *taken at (or at a point downstream of) the point of offtake, at a Supply Meter Point of which the Registered User has ceased to be a User pursuant to Section V4.3, except in a case where, after the Supply Meter Point has been Isolated, the Transporter becomes (pursuant to paragraph 9(2) of the Gas Code) entitled to recover the value of the gas.*

Paragraph 9(2) of the Gas Code states:

Where—

a) *any person at premises which have been reconnected in contravention of paragraph 11(1) below takes a supply of gas which has been conveyed to those premises by the gas transporter; and*

b) *the supply is taken otherwise than in pursuance of a contract made with a gas supplier, or deemed to have been made with such a supplier by virtue of paragraph 8 above or paragraph 19 of Schedule 5 to the Gas Act 1995,*

the transporter shall be entitled to recover from that person the value of the gas so taken.

The practical result of these definitions is that where a site is Unregistered or Shipperless during the period of theft, that theft falls into Shrinkage rather than Unidentified Gas. Recovery of transportation or consumption charges is a separate issue and not the driving factor for what classification is used.

A list of Unregistered and Shipperless sites where theft has occurred is provided in Appendix J, and these are excluded from the theft analysis.

4.8.2 AQ Estimate using Metered Consumption plus Theft Estimate

The provision of meter read/consumption data allows the market sector of each theft-affected site to be calculated more accurately than simply using the pre-theft AQ. The new calculation, described in detail in the Interim Report [25], uses meter read data and the theft estimate to calculate both metered and unmetered consumption at the site in question during the period of theft, which in turn is used to derive an AQ for market sector classification. The key features of this method are as follows:

- Theft is applied to the formula year of occurrence.
- Theft that crosses year-to-year boundaries is split into different formula years using ALPs rather than a flat profile.
- Meter reads are used to calculate annual consumption during the period of theft and the AQ for the site is calculated using the standard formula (i.e. as used by Xoserve) based on this figure
- The theft estimate is adjusted for seasonal normal conditions and added to the calculated AQ to give the overall theft-inclusive AQ.
- If consumption cannot be calculated, then an alternative method of assigning sector is required. In this case pre-theft AQ was used where possible, and where this was not available, post theft AQ plus a seasonally normalised theft figure for that year added was used. Ultimately if no AQ can be found, the current AQ associated with the theft record is used.
- The new AQ as calculated using these methods is used to set the market sector for each site for each year in which the theft occurs. Market sector can potentially change from year to year for any site.
- If theft apportioned to a given year is greater than 73,200 kWh this overrides the AQ in cases where the consumption calculation fails, and hence in such cases the site will always be classified as LSP.

Using this method as a base, there are three potential methods of calculating the market sector split of theft for successive years:

1. Calculate the theft split directly using the above method for each year.
   This method allows a dynamic theft split over time.
2. Calculate the theft split at a point in time (i.e. now) and fix for future years.
   In this case the theft split is not dynamic but is protected from external influence.
3. Calculate the theft split at a point in time (i.e. now) and define as a percentage of throughput. Use the throughput percentage to calculate theft split for future years.
   This approach protects against external influence and allows for changing market sector size, but does not allow for changes in theft rates within market sectors.

4.8.3 AQ Over-Ride when Theft is Greater than 73,200 kWh per Annum

The original theft split method from the 2011 AUGS for 2012/13 did not account for situations where the amount of theft detected in a given year was greater than the 73,200kWh market sector threshold. This was discussed at the UNCC meeting in May 2012 as part of the clarification of the original method.

This method produced a single AQ to be used across all years of theft and did not support any comparison with the amount of theft that occurred in a given year to potentially adjust it.

The AUGE has investigated this and identified a number of sites that are defined as SSP by their AQ, but where the estimate of theft that occurred in the year exceeds the 73,200 kWh threshold. The method described above therefore incorporates the facility to override the AQ classification of a site for a given formula year if the theft occurring in that year exceeds 73,200kWh. Note that this exception is only relevant where the consumption calculation fails: where the consumption method is used, any site with theft of more than 73,200kWh in an individual year will automatically be assigned to the LSP market sector for that year. A list of meter points that would have fallen into SPP using the previous method but are classified as LSP with the new method is provided in Appendix K.

When applied to the theft method used in the 2011 AUGS for 2012/13, this amendment results in an increase in the LSP theft split of 8-12% depending on the year, occurrence and size of the thefts involved. It is recognised that this is a step change in the value of LSP theft, but the lower values quoted in the 2011 AUGS for 2012/13 were due the limitations of the method and in particular the data that was available at the time. The new values represent an improved estimate of the SSP/LSP market sector split based on more comprehensive data.

4.8.4 Throughput Method

Whilst the Consumption plus Theft Method described in Section 4.8.2 represents a significant improvement in the theft estimation procedure, it still contains a number of risks and drawbacks, as follows:

- The method carries an inherent assumption that the split of unknown theft between market sectors is the same as the detected theft split. This may not be the case if the levels of effort to detect theft in each sector are different.
- The calculation relies on the accuracy of the theft estimate and the estimate of the period of theft in two places: firstly the estimation of the MPR AQ (leading to market sector assignment) and secondly when calculating aggregate theft for each market sector.
- The potential exists for external influence on the theft split if mixed shippers focus on detecting theft in one sector over another.
- There is a high rate of consumption calculation failures. In all, approximately 50% of all consumption calculations fail.

- When using AQs from periods before and after theft (in the case of consumption failures) the consumption may not be representative of the consumption during theft if the customer has changed. Some information has been received from Shippers regarding this, which shows that the majority of customer changes are in the SSP sector as expected. The data available is insufficient to allow any further conclusions to be drawn about any potential impact on the market sector split for theft.

As a result of these concerns, the Throughput method was developed, which has no reliance on AQs, consumptions for theft affected meter points, or theft estimates. Under this method, theft is split between the SSP and LSP market sectors based purely on throughput. The advantages and disadvantages of this approach are as follows:

Advantages:

- 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.
- Other elements of the Balancing Factor (i.e. those elements bundled in with theft) would be apportioned by throughput.
- Issues concerning treatment of unregistered theft-affected sites and use of pre/post theft AQ are removed.

Issues:

- The throughput method carries a fundamental assumption that the rates of theft in each market sector and the volumes stolen as a percentage of the market sector total are similar – in other words, the prevalence of theft does not differ by market sector and so throughput can be used as an effective method of splitting total theft.

4.8.5 Analysis and Conclusions

The performance of the alternative theft apportionment techniques are compared using two criteria:

- The results produced.
- The integrity of the method. This includes the robustness of the method to data issues and external influence.

There were three variants of the Consumption plus Theft Method described in the Interim Report [25] which at this moment in time are equivalent and return the same results, and so these are compared against those from the Throughput Method. The results of this comparison published in the Interim Report have since been updated following improvements to the consumption method (for the total UG figure), which were then introduced to the theft consumption method. Details of these can be found in Appendix L. The full set of year-on-year results from The Consumption plus Theft Method is shown in Table 13 below. The consistency of the yearly values returned by this method can be seen, with the final value being an LSP theft percentage of 21.5% based on the average of results for 2008-2010.

Table 13: Consumption plus Theft Method LSP Percentage

	LSP (GWh)	SSP (GWh)	Total (GWh)	% Split
2007	4.01	16.82	20.83	19.2
2008	5.84	21.55	27.40	21.3
2009	5.05	17.27	22.32	22.6
2010	3.24	12.52	15.76	20.6
2008-2010				**21.5%**

Results from the Throughput Method are shown in Table 14.

Table 14: Market Sector Split Based on Throughput in TWh, All LDZs by Formula Year

Sector	2006	2007	2008	2009	2010	Final
NDM LSP	135.1	134.4	134.5	125.0	124.1	
SSP	348.7	368.4	379.4	363.2	376.2	
Total	483.7	502.9	514.0	488.2	500.3	
LSP %	27.9%	26.7%	26.2%	25.6%	24.8%	**23.3%**

It can be seen from Figure 19 that the LSP percentage of throughput has dropped consistently across the years under consideration.

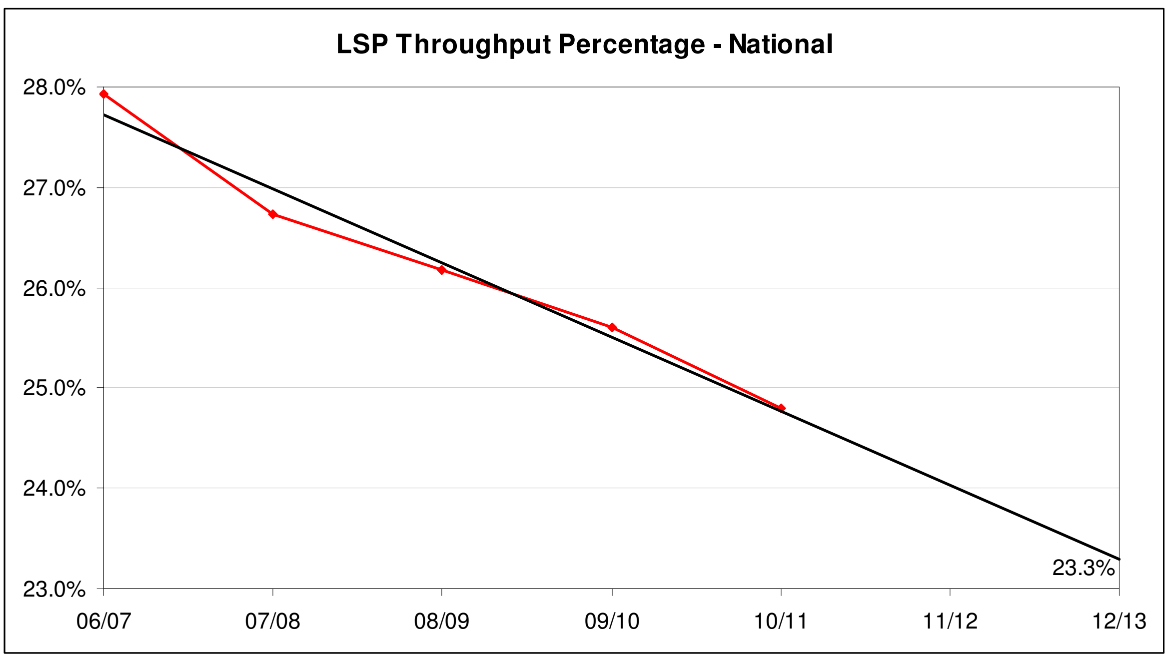


Figure 19: LSP Throughput Percentage

The data used to derive this trend in LSP market share runs from April 2006 to March 2011. As such it covers periods of strong growth in the UK economy (2006-07), the recession of 2008-09 and the current partial recovery where the economy has remained relatively stable. Throughout all of these different periods with contrasting economic background conditions, the trend in LSP market share has remained consistent. Therefore, the AUGE believes it is reasonable to extrapolate this trend to give the best estimate figure for the end of the 2012/13 financial year. Given that in this case the data finishes in 2010/11, it is necessary to extrapolate two years forward, giving a final figure of 23.3% from this method.

The AUGE notes that the allocations for formula year 2011/12 indicate a higher percentage of LSP throughput, which is entirely out of step with this trend. As seen in the consumption analysis, however, there are likely to be corrections to LSP metered volumes that have yet to be taken into account and will affect this figure. This is likely to have the effect of reducing the LSP allocation and percentage of throughput for 2011/12. Therefore only data to end March 2011 will be used at this time.

The figures returned by the two methods are therefore similar in magnitude, and the final stage is to test them statistically to ascertain whether any significant differences exist. The appropriate statistical test is a 1-sample t-test based on the Consumption plus Theft Method results tested against a hypothesised mean of 23.3%, the value from the Throughput Method. This approach is necessary due to the fact that data from the Throughput Method does not vary randomly. This comparison shows no significant difference between the two (although it is acknowledged that the power of the test is limited because of the very small sample size, i.e. the 3-year period used for the calculation of the results).

It has therefore been shown that the two methods produce results with no significant differences between them. The remainder of the comparison is an assessment of the integrity of the methods.

The variants of the Consumption plus Theft Method share a number of issues:

- The calculations are 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.
- Accuracy of the metered consumption calculation.
- Potential effect of customer changes on pre/post theft AQs.
- Site classification issues – e.g. Unregistered sites.
- An assumption that the market sector split of unknown theft is the same as that of detected theft.

The impact of these assumptions and issues must be weighed up against the single but fundamental assumption of the Throughput Method, which is that both the LSP and SSP market sectors steal gas in a similar pattern, so that the theft split is the same as the market sector throughput split.

The fact that the two methods return results with no statistically significant differences between them suggests that this underlying assumption is reasonable.

The Consumption plus Theft Method is highly sensitive to the quality of the input data (i.e. theft amount, theft duration, AQ, meter reads). The rules used in data cleansing/verification (for example, what constitutes an acceptable meter read) can have a large impact on the final results, and changes in these rules or the interpretation of the rules can result in significant changes to the final estimate. This variability and potential for manipulation is undesirable for such a key part of the overall methodology and this suggests that the Throughput Method, which is more robust to external influence, would provide a more stable basis for the calculations.

Therefore, based on the detailed data now available and the comparison between techniques presented above, the AUGE recommends that for the 2012 AUGS for 2013/14 and for AUGS in future years, the theft component of Unidentified Gas is split by market sector in the same proportion as throughput. This is the therefore the technique referred to in the remainder of this document.

4.8.5.1 Potential Amendments to the Throughput Method

Read Units:

The original analysis did not use Read Units when calculating consumptions and therefore the potential exists for calculated consumptions for some of the meter points to be out by various factors of 10. As there was a then comparison with AQ to validate the consumption, however, those that failed would have then defaulted to use of AQ. The failure rates are shown below.

Table 15: Consumption Failure Rate without using Read Units

Year of Occurrence	Number of Thefts/Part Year Thefts Occurring in Year	Consumption Calculation Failures	Failure Rate %
2007	2001	908	45
2008	2595	1119	43
2009	2774	1640	59
2010	1583	987	62

The AUGE understands from discussions with Xoserve that Read Units data held by Xoserve are of poor quality and the factors used may not necessarily be correct.

In order to verify this assertion, the AUGE applied the Read Units data subsequently provided by Xoserve to the theft estimate calculations and validated then against AQs. The results are shown in Table 16. A significantly higher number of sites now fail their consumption calculation. Given that part of the validation is against AQ it is possible that the AQs are theft affected, but it is likely that the Read Units used are also incorrect, resulting in such a large number of failures. As AQs are confirmed by Shippers it is more likely that the Read Units used are incorrect in this case.

This further highlights the issue of using detected theft and data associated with theft affected sites to estimate the market sector split.

Table 16 Consumption Failure Rate using Read Units

Year of Occurrence	Number of Thefts/Part Year Thefts Occurring in Year	Consumption Calculation Failures	Failure Rate %
2007	2001	1645	82
2008	2595	2083	80
2009	2774	2332	84
2010	1583	1310	83

Exclusion of EUCs

One potential extension of the Throughput Method is the exclusion of EUCs where theft can be shown not to occur. This is typically regarded as being for the higher EUC categories e.g. 07B-09B.

Table 17 below shows both the number of recorded thefts (during the time period used for the UG analysis) split by EUC and the population of those EUCs (for those 10 LDZs for which consumption data has been received). This table shows that there are indeed no detected thefts for any EUCs higher than 06B during the time period in question (for thefts that occurred within the period 01/04/2007 to 31/03/2011).

The population column also shows the number of sites in each EUC reduces as the EUC groups increase, however. A simple calculation of the expected number of thefts in relation to EUC size shows that for high EUC groups this is small. Even if theft did occur in these EUCs with the same frequency as for other EUCs, still none would have been expected in this timeframe: in other words, a much larger population of thefts would be required before any in the EUCs 07B-09B would be expected. Although there are no thefts for the higher categories in the detected population, it doesn't necessarily mean that there are not any in the remaining population that are yet to be detected.

Table 17: Theft Occurrences Split by EUC

EUC Group	Population	Thefts
01B	16,487,878	4150
02B	136,382	208
03B	32,717	6
04B	13,437	4
05B	3,391	2
06B	1,121	1
07B	330	0
08B	99	0
09B	16	0

There is therefore no evidence at this stage to suggest that sites in the higher EUCs steal gas with a different frequency to those in lower EUCs. As time goes on and more theft data is collected, it may be possible to draw more conclusions, and this area will be revisited as appropriate in the future. For the current AUGS, however, we cannot assume that the higher EUC groups do not steal gas and so this technique will be applied using all EUCs.

Meter Read Frequency:

In a similar manner to the above analysis for EUCs, the issue of whether meter read frequency affects tendency towards theft has been considered. Table 18 shows the frequency of thefts split by EUC and meter read frequency.

Table 18: Theft Occurrences Split by Meter Read Frequency

EUC Group	Meter Read Frequency	Number of occurrences
01B	6	2386
01B	A	1742
01B	M	22
02B	6	164
02B	A	29
02B	M	15
03B	A	1
03B	D	1
03B	M	4
04B	M	4
05B	M	2
06B	M	1

The data in this table shows that theft does occur for meter points with all read frequencies: daily, monthly, 6-monthly and annual. In order to analyse this area further, population data split by meter read frequency will also be required in order for theft rates to be calculated and compared. This data is not currently available to the AUGE and so will be requested and used in future analysis when available. The AUGE also notes that the meter read frequencies held by Xoserve may not necessarily be at the same level as meter read frequencies held by the Shippers. For example, some sites may be metered monthly, but the Shipper reports this on a 6 monthly or annual basis. This mismatch will also potentially affect the analysis, as it is the Shipper read frequency that is of relevance in this case.

There is insufficient evidence from the data examined to date to remove any data from the Throughput Method based on meter read frequency and so this technique will be applied using data for all meter read frequencies.

4.9 Industry Initiatives under Review

In the 2011 AUGS for 2012/13 [20] the AUGE identified a number of industry initiatives that may have an impact on UG going forward. An update on the status of these is given below. New modifications have also been raised that, if implemented, may also have an impact on the AUGE processes. These are summarised briefly below.

Mod 369: Re-establishment of Supply Meter Points – Measures to Address Shipperless Sites

This Modification Proposal [2] sought to modify the existing provisions of the Uniform Network Code regarding Re-establishment of Supply Meter Points to ensure Supply Point Registration where gas is consumed at a Supply Point which has been subject to Effective Supply Point Withdrawal but the original Supply Meter remains connected (or has been reconnected) and is capable of flowing gas. If adopted, this Mod would have resulted in the removal of the "Shipperless Sites (Passed To Shipper)" category from the Unregistered/Shipperless element of the UG calculation. It does not apply to sites where a new meter has been installed and hence the remainder of the calculation would remain the same and as described in this document.

This modification panel voted in favour of this modification on 16th February 2012, but OFGEM decided not to implement the modification and a non-implementation decision letter was published on 26th March 2012 [23]. This is unfortunate as the modification would have resulted in a significant reduction in UG attributed to Shipperless sites. These sites will therefore remain a part of permanent UG for this draft of the AUGS. Two new modifications (0424 and 0425) have since been raised to address Shipperless sites and are noted below.

Mod 254 (implemented in TPD [14] section H 1.5.2): Facilitating the Use of Forecast Data in the UNC

This modification proposal [16] has already been implemented and relates to the basis for calculating seasonal normal CWV.

The definition for Seasonal Normal Composite Weather Variable (SNCWV) was updated in 2010. As AQs are calculated by adjusting metered consumption back to seasonal normal using the SNCWV, this will have a knock-on effect on AQ values. At the same time, Mod254 also came into force which further adjusts the SNCWV for future forecast values. The effect of these changes has been assessed and data updated as required. See Section 4.4 for more details.

Mod 0429 Customer Settlement Error Claims Process

Modification 0429 [28] seeks to address the mismatch between the current reconciliation window (4-5yrs) 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 proposed solution requires the AUGE to assess the amount of energy that would have been corrected (and class this as temporary Unidentified Gas) over the full period. The Modification Workgroup report also indicates costs for the AUGE should be minor.

The proposed methodology for 2013/14 is based on meter reads and consumptions rather than the RbD based method as used for 2012/13. RbD is not used in the proposed method for 2013/14.

The proposed methodology based on meter reads and metered volumes may benefit from any historical corrections to those read/volumes in the same way as RbD does (assuming that such change succeeds in passing the validation processes for the methodology). However, there is no mechanism within UNC to reconcile Unidentified Gas for years prior to the reconciliation period, which this modification effectively requires.

In order to achieve the required back-correction, the AUGE would require meter point level details of exactly what corrections were applied and when. For the meter read/consumption method this would be the corrected meter reads/volumes as appropriate, i.e. anything that had changed since the point at which Xoserve provided the data to the AUGE. From our experience of obtaining and processing such data this is likely to be a significant undertaking even on the assumption that the data is available with the tracking of dates as required. Furthermore, the AUGE would rely on corrections being captured for the full Limitation Act period regardless of the current reconciliation window. The AUGE believes this information may not be available and thus estimating this mismatch may not be feasible.

The AUGE has not been involved in the preparation of the Mod 0429 report, and this lack of involvement is necessary to ensure neutrality. It would be worth discussing the implications of Mod 0429 before the final report is produced, however, to ensure the proposed changes are actually feasible.

This is a live modification and has not yet been implemented so has no impact on the proposed AUG methodology at this time.

Mod 410/410A Responsibility for Gas Off-taken at Unregistered Sites following New Network Connections

Modification 410/410A [29] 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.

Mod 0410 is a live modification and has not yet been implemented so has no impact on the proposed AUG methodology at this time.

Mod 424 Re-establishment of Supply Meter Points – Prospective Measures to address Shipperless Sites

Modification 0424 [30] seeks to reduce the impact of Shipperless sites on Unidentified Gas. This is currently a live modification that, if implemented would have an impact on the level of Unidentified Gas attributable to Shipperless sites. If implemented, the AUGE would pick up the effects of this through the information currently provided by Xoserve.

Mod 0424 is a live modification and due to submit its workgroup report by January 2013 and has no impact on the proposed AUG methodology at this time.

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

Modification 0425 [31] 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.

Mod 0425 is a live modification and due to submit its workgroup report by February 2013. It has no impact on the proposed AUG methodology at this time.

5 Data Used

This section describes the data requested, received and used to derive the methodology to calculate UG. As a general point it should be noted that during analysis it became apparent that the data available was not always on a comparable basis. The AUGE has therefore taken care to ensure that all datasets include all components of NDM consumption, i.e. CSEPs and Scottish Independents are included throughout.

There were a variety of issues with obtaining data in 2011. 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 have been issues obtaining meter reads and metered volume data and this has been covered in previous sections. There are areas for improvement in the way the industry holds data that would greatly assist the AUGE process.

Section 5.1 below gives a summary of the data items requested for the 2013/14 analysis and their current status. The subsequent sections give more details about the data items for each individual element of the analysis.

5.1 Summary

Table 19: Data Status Summary

Analysis Area	**Dataset Requested**	**Status**
Direct Total UG Calculation	Allocated SSP and LSP loads	Received
	Metered SSP and LSP loads	Received for 13 out of 13 LDZs
	LDZ, DM and Unique Sites Metering Errors	Received
	Meter Asset Information	Received
Long Term RbD Bias	RbD quantities	Updated data received
	CSEP RbD quantities	Updated data received
Allocation Algorithm Error	Mod81 data	Updated data received
	Algorithm data (ALPs, DAFs, WCFs, SFs)	Updated data received
	CSEP AQ data	Updated data received
	Non-CSEP AQ data	Updated data received
	Proportions of SSP and LSP sites successfully recalculated in AQ review	Updated data received
Unregistered and Shipperless Sites	Asset and Shipper meter reads for new LSP sites	Updated data outstanding
	Asset meter reads for orphaned sites	Updated data outstanding
	Gas Safety Visit data	Updated data outstanding
	Snapshot files	Received on an ongoing basis
iGT CSEPs	Known CSEP data	Updated data received
	Snapshot files	Received on an ongoing basis
Theft	Detected and alleged theft for 2011-12	Updated data received
	AQs before, during and after theft	Updated data received
	Metered volumes and meter reads, Read Units and T/P factors for theft detected sites	Received
	EUC groups and meter read frequencies for theft affected sites	Received

Meter Error	Meter capacity report	Updated data received
New Analysis	New and lost sites	Received
	EWCF	Received
	SNCWV Adjustment Factors by EUC	Received
	Updated AQ=1 information	Received

5.2 Direct Total UG Calculation (Consumption Method)

As described in the September 2012 Interim Report [25], using the consumption method the total UG is estimated as follows:

Total UG = (Alloc SSP + Alloc LSP) – (Metered SSP + Metered LSP) (5.1)

This can be alternatively stated as:

Total UG = Aggregate LDZ Load – DM Load – Shrinkage – (Metered SSP + Metered LSP) (5.2)

This is the case because the aggregate allocations are scaled (using SF) to total LDZ load with DM and Shrinkage removed.

Using the first version of this equation (5.1), this creates a requirement for the following data for each LDZ for a full UG calculation based on meter reads:

1. Allocated LSP loads.
2. Allocated SSP loads.
3. Metered LSP loads plus the number of MPRNs for which metered data is available.
4. Metered SSP loads plus the number of MPRNs for which metered data is available.
5. Total number of MPRNs in the LDZ (including those in CSEPs) for the LSP sector.
6. Total number of MPRNs in the LDZ (including those in CSEPs) for the SSP sector.

Calculations are then carried out at the formula year level of granularity.

Note that whilst the equation intrinsically uses *LSP Alloc – LSP Metered* (which is what RbD is defined as), it is not possible to use RbD figures in place of the raw LSP information. This is because RbD also contains a significant proportion of retrospective corrections, and so each month or year's figures do not represent the true difference between allocated and metered LSP load in that time period. Therefore the raw figures for LSP allocations and LSP meter point consumptions must be used in the UG consumption method calculations.

Data has been requested from Xoserve in the following formats. In all cases, data was provided for the time period 01/04/2008 to 31/03/2012.

- 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. In addition to meter reads, the EUC and the AQ have been provided for each MPRN so that calculated consumptions can be reconciled against allocations on an EUC-by-EUC basis.
- A list of MPRNs for which no meter reads were recorded in the analysis time period. This list also includes both EUC and AQ. Therefore, the total number of MPRNs in each EUC can be obtained by adding the count of meter points in the consumption data file to the count of meter points in the "no meter reads" file.

- 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.
- A list of meter installation dates and numbers of meter dials, on an MPRN-by-MPRN basis. This information is used in the processing of meters which appear to have negative consumption to determine if meter index rollover has occurred.

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 Section 6 below. The exact format of the data provided is given in Appendix A.

5.3 Long Term RbD Bias

Both standard and CSEP monthly RbD values split by LDZ have been requested and received.

5.4 Allocation Algorithm Error

The AUGE has requested and received the following information that was required to update the RbD-based method.

1. AQ data broken down by EUC in order to allow calculations to be performed using the deeming algorithm. Separate datasets were provided for loads within CSEPs and loads outside CSEPs.
2. Mod81 data. This provides a more detailed picture of AQ changes between gas years resulting from the AQ review and allows like-for-like tracking of AQs from year to year based only on those sites whose AQs were successfully recalculated. This data also provides all the required information on SSP/LSP threshold crossers that is used in the analysis.
3. New/Lost Meter points. In order to be able to calculate correct aggregate AQs by EUC for each gas day, the AUGE has requested and received a complete set of data for meter points which are new or have been lost. This data includes the date added/lost, AQ and EUC. This data is at aggregate level split by EUC by gas day.
4. Allocation Algorithm Data. This includes ALPS, DAFs, WCFs and SF, and allows the AUGE to replicate results from the allocation algorithm, which is necessary in order for algorithm bias to be calculated (RbD-based method). Data is split by day, LDZ, and EUC.

5.5 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.

5.6 Unregistered/Shipperless Sites

The following information has been requested concerning Unregistered/Shipperless sites. In each case both the number of sites and their aggregate AQ was requested. All data is required to be 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 has been removed and 12 months after removal 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 has been requested:

- 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 Visits.

- Asset meter reads for orphaned sites to determine the proportion which have been flowing gas prior to becoming registered. The current dataset runs to 19/08/11 and hence needs to be brought up to date.

- Asset and shipper details for a sample of confirmed sites. The current dataset contains data up to 20/10/11 and is used to calculate the proportion of UG from Unregistered sites that cannot be backbilled. This needs to be brought up to date.

Updated data for 2012 is yet to be received for these items. Any results contained in this report are therefore based on 2011 values for these.

5.7 Meter Errors

Data for meter error calculations consists of meter capacity, AQ and NDM/DM classification records for all LSP sites. The dataset used in the 2012/13 analysis is accurate as of September 2011 and so a new version applicable to the present time was requested.

This dataset has been supplied.

5.8 Theft

The following data concerning theft has been requested and received:

- A list containing records of each occurrence of alleged and confirmed theft, presented with each occurrence as an individual record. For each record, the following details were provided:
    - Date
    - LDZ
    - Shipper
    - Market sector (LSP band/SSP) based on *current* AQ value
    - Transporter or Shipper responsible
    - Estimated volume (kWh) – where the theft allocation has been pursued

Data from the 2006 to 2012 Theft of Gas summaries has been received.

Note that the meter point AQs provided are the current (latest) AQs and not necessarily the AQs that were in force at the time the theft occurred or was detected. Historical AQs for each site have also been provided from 2000 onwards. The data includes a dummy MPR reference, start date, end date and AQ value.

In addition, meter reads for each theft-affected site have also been received. These are consistent with the meter read records supplied for the consumption analysis, but in this case cover theft-affected sites only. This includes read units, T/P correction, EUC group and meter read frequency. This data allows market sector for each site, at the time of the theft, to be calculated independently without reliance on AQ.
Note that dummy MPR IDs for theft are different to those for meter reads and metered volumes.

6 Methodology

This section describes in detail the methodology for each aspect of UG where the calculation method has changed since the original analysis was carried out in 2011. Where methods have remained the same, details can be found in Section 6 of the 2011 AUGS for 2012/13 [20].

As described in Section 4, the proposed method for calculating the UG total is the meter read/consumption-based calculation initially described in the September 2012 Interim Report [25]. This 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 method presented in the 2011 AUGS for 2012/13 [20], 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 as described in the 2011 AUGS for 2012/13) 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)

6.1.1 Algorithm

Detailed descriptions of the raw data fields, derived data fields and step-by-step method used to calculate consumption figures for each individual meter point are given in Appendices A, B and C. In addition, worked examples of both a standard consumption calculation and a meter index roll-over affected calculation are given in Appendix D.

A flow chart of the steps involved in this procedure is given in Figure 1. Figure 20 below shows the steps involved in calculating the consumption for an individual meter point. The steps refer to the detailed description given in Appendix C. Prior to this procedure, a pre-processing step is performed as described in Appendix B.

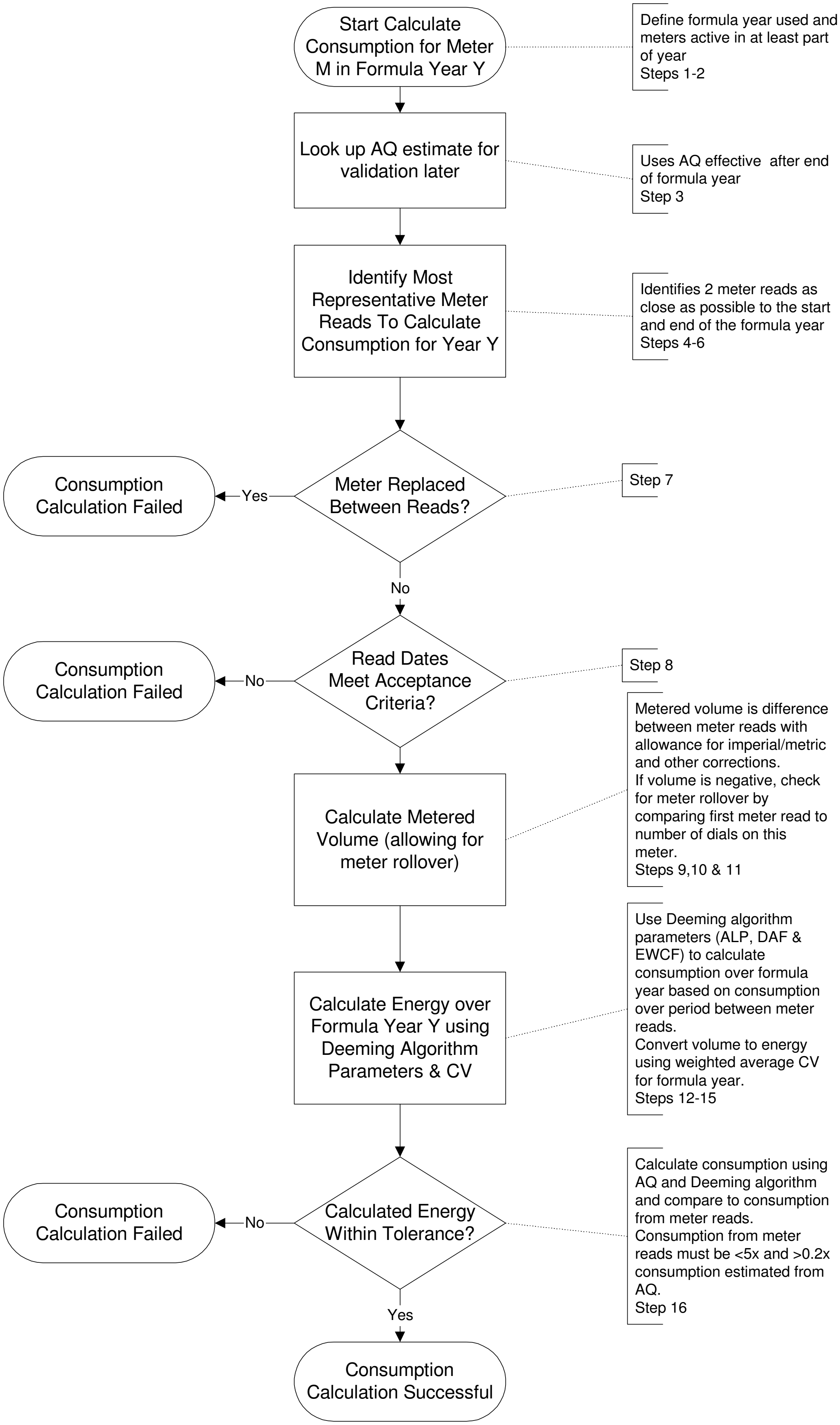


Figure 20: Consumption Calculation Flow chart

When applied to each meter point in any given LDZ, this procedure 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 algorithms:

- Sites for which a consumption could be calculated
- Consumption calculation failures
- Sites in CSEPs (for which meter reads are not available)

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

The sum of these three categories across all EUCs gives the total NDM population of the LDZ. Typical summarised output information for a single LDZ is shown in Table 20 below.

Table 20: Consumption Calculation Output Example (EM LDZ, 2009/10)

EUC	Sample Size	Sample Metered Consumption (GWh)	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption (GWh)
01B	1,855,569	30,016.9	0.009	281,011	132,732	2,269,312	36,709.8
02B	16,418	2,231.4	0.057	4,318	138	20,873	2,837.0
03B	4,008	1,795.0	0.118	828	31	4,867	2,179.8
04B	1,541	1,837.2	0.384	290	94	1,925	2,295.3
05B	419	1,414.0	0.937	78	6	503	1,696.8
06B	138	1,213.8	2.284	23	4	165	1,453.4
07B	45	937.9	4.286	13	0	58	1,211.2
08B	11	404.2	5.003	6	0	17	648.8
09B	1	60.6	0.000	0	0	1	60.6

In this table:

- "Sample Size" refers to sites with non-zero calculated consumption only
- "Failed Sites" are only those sites outside the sample that are assumed to be consuming (decision rule based on AQ and defined below)
- "Population Size" is therefore the best estimate of the consuming population (i.e. non-consuming sites removed)

Therefore in the consumption calculation, for each EUC it is the average consumption of consuming sites only that is calculated, and this is applied only to sites outside the sample that are assumed to be consuming. The criteria for classifying sites into these categories are defined below. This process allows for different proportions of consuming and non-consuming sites within and outside the sample.

The full population (of both consuming and non-consuming sites) for this LDZ is given in Table 21 below for comparison purposes.

Table 21: Population by EUC Including Non-Consuming Sites (EM LDZ, 2009/10)

EUC	Consumption Calculation Successful	Consumption Calculation Failed	Sites in CSEPs	Total
01B	1,887,961	293,479	132,732	2,314,172
02B	16,434	4,476	138	21,048
03B	4,014	858	31	4,903
04B	1,543	312	94	1,949
05B	421	82	6	509
06B	138	24	4	166
07B	45	13	0	58
08B	11	7	0	18
09B	1	0	0	1

Full details of the calculation method that creates the estimated metered LDZ NDM consumption (i.e. the final column of Table 20) are given in Appendices A-C. The actual data will be supplied separately in an Oracle database.

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, as shown in Table 22.

Table 22: Unidentified Gas Estimate (EM LDZ, 2009/10, GWh)

Total Allocation	50,122	
Metered Consumption	49,093	%
Low CI Bound	921	1.84
UG	977	1.95
High CI Bound	1032	2.06

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.

The low and high Confidence Interval bounds represent a 95% confidence level and are calculated using the Standard Deviation figures given in Table 20. Details of the calculation method are given in Section 6.1.2 below.

Data has not yet been received from Xoserve to allow the AUGE to make the required adjustments for temporary UG and meter error. This is due to the very large volume of data, and the complexity of that data, required for the consumption analysis. At the time of writing, Xoserve are still producing the meter read and consumption data for the remaining three LDZs, and other datasets will not be supplied until this process is complete.

Therefore, final UG estimates cannot be presented at this stage. Interim figures are supplied for each LDZ for which data has been received, however, subject to the following:

- Adjustments for meter errors have been made only for the known large issues in SC, SO and NT LDZs, data for which was retrieved from the Joint Office website.
- An approximate adjustment for temporary UG has been made based on calculated temporary UG from the 2011 analysis.

- The figures have been adjusted for detected theft.

The interim figures for the 10 LDZs for which data has been received are given in Section 4.7. Details of individual elements of the calculation that require further explanation are given in the sections below.

6.1.2 Scaling Up Procedure

As described above, the procedure for scaling up from metered sample demand to estimated population demand is based on consuming sites only, in order to account for potential differences in the proportion of non-consuming sites within and outside the sample. Rules are applied to define both sample sites and failed sites as consuming/non-consuming, and the scaling up procedure takes place based on the reduced population of consuming sites. The criteria applied are as follows:

- In Sample
  Sites are classified as non-consuming if the calculation is successful and returns a consumption of zero. Such sites are not included in the sample (so they do not contribute to the sample size or average consumption for the EUC in question), but they do not qualify as Failed Sites either. In effect they are removed from further calculations.
- Failed Sites
  Sites where the consumption calculation fails are classified as consuming/non-consuming based on AQ, as this is the only reliable data available for such sites. It is recognised that due to changing circumstances for each site, 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. sites within the sample):
  - the proportion of sites with AQ=1 for Year X that are consuming in Year X = A
  - the proportion of sites with AQ>1 for Year X that are consuming in Year X = B

  The consuming population of the non-calculated meter points is then calculated as
  Consuming Popn = A x (sites with AQ=1) + B x (sites with AQ>1)

For each EUC, the average annual demand for a consuming site is calculated, and this is applied to each failed site classified as consuming (according to the rules defined above). This average load is also applied to loads in CSEPs to give the final population total demand.

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 sites which were only active for part of the year are assigned an average demand scaled based on the ALPs for that part of the year.

6.1.3 LDZ Offtake Meter 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 23 below.

Table 23: 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 can therefore be applied to the Unidentified Gas calculation after the consumptions have been calculated and aggregated to EUC level (i.e. after the step whose output is shown in Table 20 above). 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 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

At the time of writing, only the sample data given in Table 23 had been received, and therefore it has not been possible to include the effects of all meter errors in the interim UG figures. This data will be used in the calculation of the final figures, however.

6.1.4 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 when this data becomes available). As described in Section 3.4, temporary UG can exist in the following categories:

- iGT CSEPs (for LSP sites only)
- 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 [20] and has not changed since this time except for an additional adjustment to the unregistered sites to account for the consistent initial AQ overstatement as described in Section 4.2.2. The figures calculated using this method are therefore deducted from the (meter error adjusted) total UG figure to give the final 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.

6.1.5 Total UG Estimate Confidence Interval

When carrying out this estimation of total metered load based on a sample, a 95% confidence interval for the aggregate load can be produced. This is based around the Central Limit Theorem, an extension of standard Confidence Interval calculation procedure, and a Finite Population Correction.

The statistical distribution of individual metered loads, calculated as described in Appendix C, is unlikely to be Normal. Regardless of the statistical distribution of the values themselves, however, the Central Limit Theorem states that the mean of a series of samples of size n will be Normally distributed, with a mean of $\bar{x}$ and a Standard Error of $S/\sqrt{n}$, where $\bar{x}$ is the mean of the sample taken and S is the Standard Deviation.

If a relatively large sample (greater than 5% of the entire population) is taken, it is appropriate to apply a Finite Population Correction to the Standard Error: this reduces the Standard Error in relation to the size of the sample taken, so larger samples lead to greater confidence in the estimate of the population mean, and hence a narrower Confidence Interval. The Finite Population Correction is defined as follows:

$$FPC = \sqrt{\frac{N-n}{N-1}} \qquad (6.1)$$

where *N* is the population size

n is the sample size

This procedure can be used to give a 95% Confidence Interval, which is a range of values between which we can be 95% sure that the true population mean (i.e. the average consumption for a meter point in that market sector) lies.

For large samples such as those used in these calculations, the high and low limits of the Confidence Interval are given by the following formula (which includes the Finite Population Correction):

$$95\%\ CI = \bar{x} \pm 1.96 \times \frac{S}{\sqrt{n}} \times \sqrt{\frac{N-n}{N-1}} \qquad (6.2)$$

where $\bar{x}$ is the sample mean

S is the sample SD

N is the population size

n is the sample size

This equation provides limits for the EUC mean demand, between which we are 95% sure that the true population value lies. The aggregate EUC demand is simply N times the mean, and so from the mean Confidence Interval it is easy to calculate an equivalent for the aggregate, simply by multiplying both the low and the high estimates by N. This calculation can also be carried out at the full population level (as opposed to at EUC level) by combining the estimates of Standard Error across all of the EUCs (i.e. the strata of a stratified population).

This is obtained using the following formula:

$$SE_P = \left(1/N\right) \times \sqrt{\sum_{EUC=1}^{k} \left(N_{EUC}^2 \times S_{EUC}^2\right) / n_{EUC}} \qquad (6.3)$$

Confidence Intervals for the final UG figures will be calculated using this technique and presented with the final 2013/14 figures.

It is important to note that Confidence Intervals calculated in this manner reflect the variability in the UG estimate *caused by the sample size only*. Other factors that will affect the UG figure, in particularly the uncertainty in the meter reads themselves, cannot be quantified. In effect, the CI calculation reflects the variability inherent in the UG estimation technique itself rather than the uncertainly caused by any inaccuracies in the input data.

In any analysis of this type, the input values are regarded as "actuals" – they represent the most accurate data available and are therefore assumed (by the calculation) to be correct. In reality there will be uncertainty associated with the input data, and this could theoretically be of an equivalent or larger order than the variability in the calculation method that is reflected in the CI. Without a dedicated large-scale experiment designed to calculate the uncertainty of meter reads, however, it is not possible to quantify this element. Such an experiment would constitute a separate project in its own right. Therefore, given the data available, for this analysis the meter reads must be regarded as accurate, and the Confidence Interval reflects this assumption.

6.2 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 [20].

6.3 Unregistered and Shipperless Sites

The analysis for this element of the UG calculation remains the same as described in the 2011 AUGS for 2012/13 [20]. The figures will be updated based on the latest data when the final UG estimates for 2013/14 are produced.

6.4 IGT CSEPs

The analysis for this element of the UG calculation remains the same as described in the 2011 AUGS for 2012/13 [20]. The figures will be updated based on the latest data when the final UG estimates for 2013/14 are produced.

6.5 Consumer Metering Errors

The analysis for this element of the UG calculation remains the same as described in the 2011 AUGS for 2012/13 [20]. The figures will be updated based on the latest data when the final UG estimates for 2013/14 are produced.

The AUGE is currently investigating the possibility of extending the Meter Error calculations to cover SSP meters in addition to the LSP NDM and DM market sectors currently covered by the available data.

6.6 Shipper Responsible Theft

As described in Section 4.8, undetected theft (which forms the vast majority of the Balancing Factor) is 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.
- 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 one fundamental assumption of this method that the rates of theft in each market sector and the volumes stolen as a percentage of the market sector total are similar – in other words, the prevalence of theft does not differ by market sector and so throughput can be used as an effective method of splitting total theft. This assumption has been shown to be reasonable, as discussed in Section 4.8.5.

The LSP market sector percentage calculated using this method is 23.3%.

6.7 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 Meter Error for DM sites, and as described in the Worked Example in Section 6.9 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 preparation of the final 2012 AUGS for 2013/14 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.

6.8 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.

This will be based on the total LDZ consumptions for formula years 2009 and 2010 subtracted from their corresponding total NDM allocations with corrections for meter error, detected theft and temporary UG as described in this statement.

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

6.9 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). 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 is calculated over the formula years 2009 to 2010 due to availability of data. The calculation is carried out as described in Section 6.1 above and in Appendices A-C, 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 3.4 and calculated as described in the 2011 AUGS for 2012/13 [20]. 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 at 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 is 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 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.

   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. As stated in Section 6.5 above, the AUGE is currently investigating the possibility of extending this analysis to cover SSP meters.

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 6.6 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.

7 Unidentified Gas Estimates

This section is reserved for a set of tables containing the best estimates of UG calculated using the methods described in Section 6 and in the 2011 AUGS for 2012/13 [20]. These values will be calculated using the appropriate methods and most recent data that is available once the methodology for the latest AUGS has been approved by the UNCC. Estimates will be presented on an LDZ by LDZ basis, with each LDZ's figures split into SSP and LSP market sectors, and also by each category of UG. The Scottish Independents will also be included within the figures for SC LDZ, although their contribution to the overall UG figure has been negligible up to this point. These tables will therefore give a full breakdown of UG by source in each LDZ.

An example (unpopulated) table is shown below. The top section shows the breakdown of UG by category, with different columns for the SSP and LSP market sectors. The individual components of the Shipperless/Unregistered category are shown in grey, with the total for the category in black. The total of the directly measured components is shown, to which the Balancing Factor (i.e. Theft plus Other) is added to give the overall LDZ UG totals for the SSP and LSP sectors, which are shown in bold. All units are GWh.

Table 24: Unidentified Gas Summary (GWh) – Example Table

	XX LDZ		
	SSP	NDM LSP	DM LSP
iGT CSEPs	0.00	0.00	0.00
Shipperless/Unregistered	0.00	0.00	0.00
- Shipper Activity	0.00	0.00	0.00
- Orphaned	0.00	0.00	0.00
- Unregistered <12 Months	0.00	0.00	0.00
- Shipperless PTS	0.00	0.00	0.00
- Shipperless SSrP	0.00	0.00	0.00
- Without Shipper <12 Months	0.00	0.00	0.00
Meter Errors	0.00	0.00	0.00
Total Directly Measured	0.00	0.00	0.00
Theft + Other	0.00	0.00	0.00
Total	**0.00**	**0.00**	**0.00**

7.1 Estimation of SAP price

The estimation of SAP price will be based on the methods used for the AUGS year 2012/13. The SAP price for 2013/14 will be estimated using the latest SAP price data obtained at the time the final figures and rates are calculated.

This 2013/14 SAP price is only used to provide a common basis for estimating the overall cost of UG in the coming gas year. In practice the SAP price actually used will be the daily average SAP price over the reconciliation billing period in question and the shipper's relevant aggregate AQ share. This is described in the TPD [14] section E 10.5.

7.2 Final AUGS Table

To be populated on final draft AUGS approval.

8 Consultation Questions and Answers

This section captures a history of the questions raised by the Industry Bodies during the consultation periods and the AUGE responses. These relate to the 2011 AUGS for 2012/13, the first draft of the 2012 AUGS, and the September 2012 Interim Report. 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" [34], "AUGS Draft2 Query Responses 14_11_2011" [19], "AUGS Query Responses 19_03_2012" [24], "AUGE Responses to 1st Draft 2012 AUGS" [26] and "AUGE Responses to Interim Report Consultation" [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 25 below contains a list of organisations that responded to the first draft of the 2011 AUGS for 2012/13.

Table 25: 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
Centrica	15/06/2011
EDF Energy	16/06/2011
GDF Suez	16/06/2011
Gazprom	17/06/2011
ScottishPower	17/06/2011

Table 26 below contains a list of organisations that responded to the second draft of the 2011 AUGS for 2012/13.

Table 26: 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
Centrica	31/10/2011

Table 27 below contains a list of organisations that responded to the final version of the 2011 AUGS for 2012/13.

Table 27: Responses to the Final Draft of the 2011 AUGS

Organisation Name	Date of Communication
Centrica	20/02/2012
Inexus	08/03/2012
Shell Gas Direct	08/03/2012

Table 28 below contains a list of organisations that responded to the first draft of the 2012 AUGS for 2013/14.

Table 28: 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 29 below contains a list of organisations that responded to the September 2012 Interim Report

Table 29: 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

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 369 Re-establishment of Supply Meter Points – measure to address Shipperless sites

[3] Mod 194 Framework for correct apportionment of NDM error

[4] Mod 194a Framework for correct apportionment of LSP unidentified gas

[5] Mod 228/228A Correct apportionment of NDM Error – Energy

[6] UNC Modification Proposals 228 and 228A Correct Apportionment of NDM Error – Energy. An Assessment by TPA Solutions Ltd, January 2010

[7] Mod 229 Mechanism for Correct Apportionment of Unidentified Gas implemented in UNC Section E 10 v3.54 26th April 2011,

[8] Theft of Electricity and Gas "Next Steps", OFGEM, January 2005

[9] Reducing Supplier Disincentives to Detect and Investigate Gas Theft – Uniform Network Code Proposal UNC231V and other Changes, OFGEM, December 2010

[10] Mod 274 Creation of a National Revenue Protection Service, Version 1.0, 11th November 2009, Eon

[11] Mod 277 Modification Report "Creation of Incentives for the Detection of Theft of Gas (Supplier Energy Theft Scheme)" Version 2.0, 20th January 2011, Joint Office of Gas Transporters

[12] Mod 346 An Alternative to the Supplier Energy Theft Scheme Based on Throughput Version 2.0, 20th January 2011, Centrica

[13] Uniform Network Code (UNC) 194, 194A, 228, 228A and 229: These proposals deal with the identification and apportionment of costs of Unidentified Gas, OFGEM, 26th May 2010

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

[15] Theft of Electricity and Gas, Discussion Document, OFGEM, April 2004

[16] Mod 254 Facilitating the use of forecast data in the UNC section H 1.5

[17] AUGS Draft 1 Query Responses, September 2011

[18] Correct Apportionment of Unallocated Gas Volumes and Mod 228, CEPA LLP

[19] AUGS Draft 2 Query Responses, November 2011

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

[21] Modification 115 and 115A Correct Apportionment of NDM error

[22] Mod 369 UNCC voting record 16th February 2012

[23] Mod 369 OFGEM decision letter 26th March 2012

[24] 2011 AUGS for 2012/13 Query Responses 19_03_2012, March 2012

[25] Consumption Analysis and Theft Interim Report, September 2012, GL Noble Denton

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

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

[28] Modification 0429 Version 2.0, 10th December 2012

[29] Modification 0410 Version 6.0, 07 December 2012

[30] Final Modification Report 0424 Version 2.0, 09 November 2012

[31] Modification 0425 Version 7.0, 09 November 2012

[32] UNC – Transportation Principle Document Section N – Shrinkage

[33] The Gas Act 1986, Schedule 2B "The Gas Code"

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

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
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 parameter. Alternative to WCF based on CWV rather than demand)
IGT	Independent Gas Transporter
LSP	Larger Supply Point
MAM	Meter Asset Manager
MEG	Monoethylene Glycol
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
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 parameter)
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
WCF	Weather Correction Factor (deeming algorithm parameter)
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 analysis.

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	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	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

Issues:

- EUC is by supply point rather than MPRN.

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	Aggregate CSEP AQ at start of formula year
COUNT_OF_SUPPLY_POINTS	Count of supply points at start of formula year

A.4 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
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 - 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

Code	Description
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)

Issues:
- METER_READ_VAL – Incorrectly entered/reported e.g. wrong number of digits recorded.
- METERED_VOL – Roll-over missed, roll-over incorrectly assumed, meter replacement missed.
- ROUND_THE_CLOCK_IND – unreliable

A.5 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.6 Meter Errors

Name	Description
METER_TYPE	Data for all of the following meter point categories is required: DM, NDM, Dom, Unique, CSEP, LDZ Offtake
LDZ	Char[2] - LDZ identifier e.g. EA
START_DATE	Start date of error
END_DATE	End date of error
ADJUSTMENT	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_xx.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
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 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'
```

B.2 FACTORS

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

Name	Description
LDZ	VARCHAR2(2) - LDZ identifier e.g. EA
EUC_BAND	VARCHAR2(5) – EUC band e.g. 08W02
GAS_DAY	Date
ALP	Raw data
DAF	Raw data
EWCF	Raw data
CV	Raw data
ENERGY_PROFILE	ALP * (1 + DAF * EWCF)
VOL_PROFILE	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	VARCHAR2(2) - LDZ identifier e.g. EA
EUC_BAND	VARCHAR2(5) – EUC band e.g. 08W02
F_YEAR	NUMBER
ENERGY_PROFILE	ALP * (1 + DAF * EWCF)
VOL_PROFILE	ENERGY_PROFILE / CV

B.3 METER_INFO_CALC_XX

This table is pre-calculated from the meter reads. There is at least one record for every meter point with reads. There will be two or more if the meter changed from imperial to metric or vice versa.

Name	Description
MPR_ID	Number - Unique ID for meter point used across ALL data
LDZ	Char[2] - LDZ identifier e.g. EA
IMP_IND	Char[1] - Indicator ="Y" for imperial meter read, else "N"
METER_INST_DATE	Date – Effective date for record
UNITS	Number - Multiplier for meter read units (0.1, 1, 10, 100 etc.)

The units are calculated by looking at the ratio of the difference between meter reads and the recorded volume. The units are then taken as the ratio rounded to the nearest power of 10.

B.4 METER_READS_XX

Name	Description
MPR_ID	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
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

The algorithm for flagging bad reads is as follows:

Given subsequent meter reads mr1, mr2, mr3 and mr4 calculate:

con1 = mr2 - mr1
con2 = mr3 - mr2
con3 = mr4 - mr3

As for the consumption algorithm if any of these are negative we check for meter index rollover

- If this gives a negative volume then check if the meter index has rolled over. Based on the number of digits in the first read infer the size of the meter and calculate its maximum possible value. If the start read was >75% of this then calculate the volume on the assumption it rolled over. If this new value is >25% of the max then it was assume a bad reading and reject the meter.

If the meter was replaced we leave the consumption null
Then if the meter was not replaced during the period we check

- If (con3 > 0) and (con2 < 0) and (con1 > 0) then we have a bad reading
    - If con1 > abs(con2) then mr2 is bad
    - Else if con3 > abs(con2) then mr3 is bad

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	VARCHAR2(2) - LDZ identifier e.g. EA

B.6 RESULTS_XX

Name	Description
MPR_ID	NUMBER
LDZ	VARCHAR2(2)
EUC	VARCHAR2(5) – from AQ or meter read
F_YEAR	NUMBER
FY_MR_CON	NUMBER – Consumption (in kWh) for formula year calculated using meter reads / metered volumes
FY_AQ_CON	NUMBER - Consumption (in kWh) for formula year calculated using AQ
METER_READS	VARCHAR2(1) - Indicator ="Y" found two meter reads which satisfy the criteria listed in the algorithm
POSITIVE_VOLUME	VARCHAR2(1) - Indicator ="Y" positive volume calculated after possibly correcting for meter index rollover
AQ_CHECK	VARCHAR2(1) - - Indicator ="N" FY_MR_CON puts the site into the LSP market 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	VARCHAR2(3) – Consumption band calculated based on consumption from meter read data if calculated successful, else on the AQ

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 not null
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 Detailed Description of Consumption Algorithm

The raw data, described in Section 0, was provided by Xoserve in the format shown in Appendix A.

This data is pre-processed to derive various useful values (see Appendix B) and then the following algorithm is used to estimate NDM consumption:

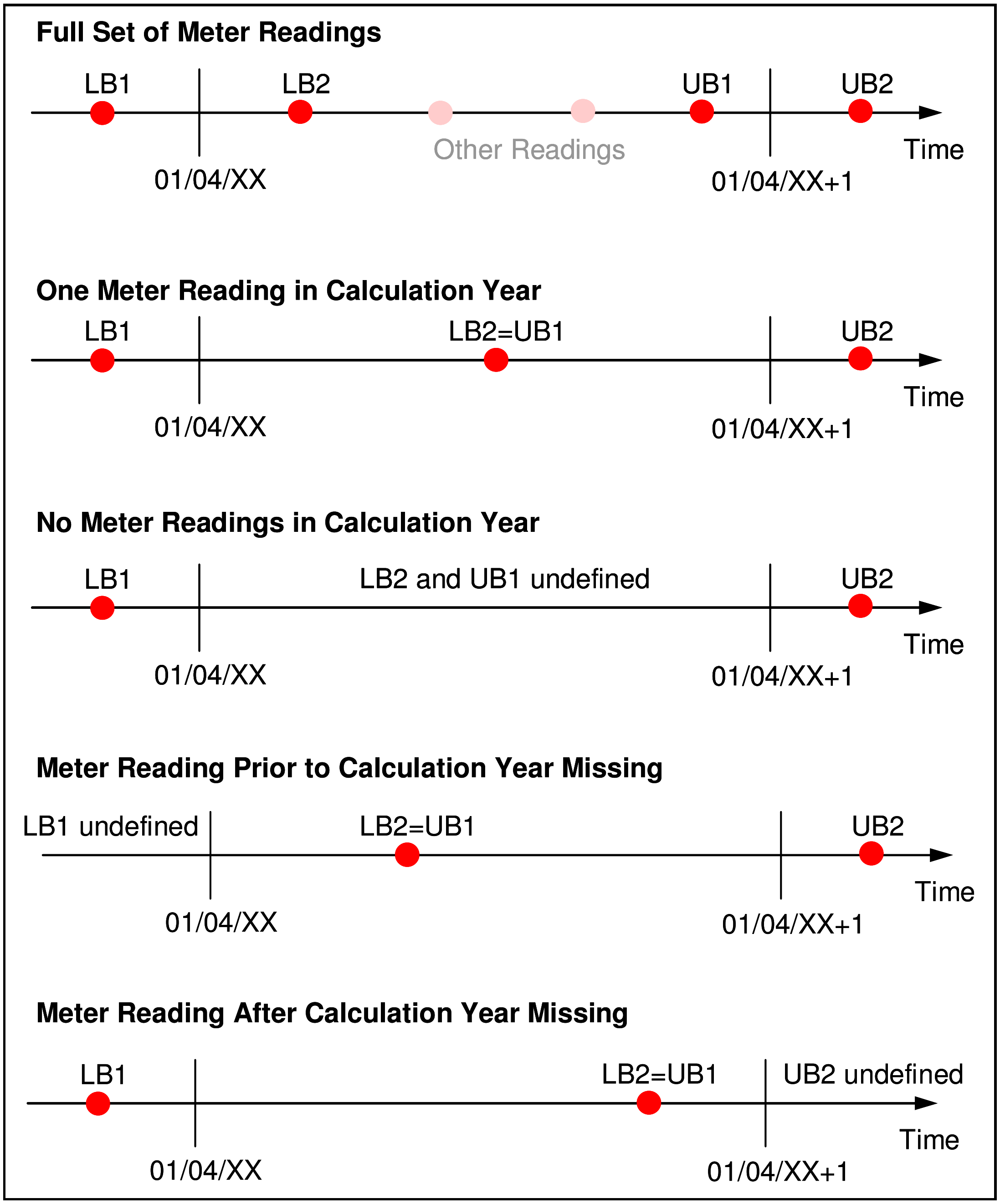
1. Given a formula year Y, define the start and end dates as 01 Apr YY and 31 Mar YY+1

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 use the latest value. Only look at the latest AQ value from each year. (This may fail if a site was only NDM for part of a year, so we relax the condition on using the latest AQ from each year in this case). We also store the EUC associated with the chosen AQ record and calculate our own estimate based on the sites AQ.

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

    excluding 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 the figure below:



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
   - LB2 is recorded as the first reading of a new meter

   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
   - UB1 is recorded as the last reading of the meter

   In which case set it equal to UB1.

7. If the meter was replaced in between LB2 and UB1 then reject the meter

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 EUC of the site:
   A. If the site is 01B then calculate the volume consumed between the two chosen meter readings. If this gives a negative volume then check if the meter index has rolled over. Based on the number of digits in the first read infer the size of the meter and calculate its maximum possible value. If the start read was >75% of this then calculate the volume on the assumption it rolled over. If this new value is >25% of the max then it was assume a bad reading and reject the meter.
   B. Otherwise sum the metered volumes between the two chosen meter readings. If there are any negative volumes in the range, set the sum to -1.

10. If the volume calculated is positive look up whether the meter is/was metric or imperial.

11. Apply either Rule A or Rule B according to the EUC of the site:
    A. If the site is 01B look up the read units. Combine this with the default correction factor (1.022640) and metric/imperial conversion factor to get a final volume.
    B. Otherwise just look up the appropriate metric/imperial factor

12. Calculate the fraction of the year that the meter point was active and NDM weighted by the ALPs.

13. Calculate a consumption estimate based on the AQ using

$$\sum_{\substack{formula\ year \\ or\ part\ thereof}} \frac{AQ * ALP}{365} * (1 + DAF * EWCF)$$

14. Calculate the volume taken over the formula year by multiplying the metered volume by

$$\frac{\sum_{\substack{formula\ year \\ or\ part\ thereof}} ALP^{v} * (1 + DAF * EWCF)}{\sum_{metered\ period} ALP^{v} * (1 + DAF * EWCF)}$$

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

15. Then convert this to energy using the a weighted average CV for the formula year, calculated as

$$\frac{\sum_{\substack{formula\ year \\ or\ part\ thereof}} ALP * (1 + DAF * EWCF)}{\sum_{\substack{formula\ year \\ or\ part\ thereof}} ALP^{v} * (1 + DAF * EWCF)}$$

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

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

Appendix D Worked Example of Consumption Algorithm

This appendix shows the consumption algorithm, described in Appendix C, applied to example data.

D.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 in 2009: Yes

2. Pick an AQ

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

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**
913600	28/04/2009	N	6400	128	0	5523	A	SSP	01B	EA	U	N
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
913600	**12/04/2010**	**N**	**7438**	**485**	**0**	**9457**	**A**	**SSP**	**01B**	**EA**	**U**	**N**
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

Vol = 7438 – 6275 = 1163 m^3

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

Therefore the final volume is 1163 * 1.022640 = 1189.33 m^3 (Compared to 1190 m^3 if we had used the metered volumes in this case)

7. Calculate consumption for formula year 2009 based on AQ

   The meter was active for the whole year.

   The AQ consumption estimate is AQ * sum over 2009 energy profile / 365 = 10477 * 370.46/365 = 10633.82 kWh

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

   Volume taken over the year is = 1189.33 * sum volume profile over 2009 / sum volume profile over metered period

   =1189.33 * 9.40 / 10.19

   =1097.12 m^3

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

   Therefore consumption = 1097.12 * 39.40 / 3.6 = 12007.51 kWh

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

D.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 mr_vol = 299 – 9257 = -8958 ft^3

As this is negative we test for roll-over

num_dials = round(log(10, start_mr)) = round(log(10, 9257)) = round(3.97) = 4

max_read = $10^{num\_dials} = 10^4$ = 10000

difference in meter reads gives mr_diff = (max_read - start_mr + end_mr) = 10000 – 9257 + 299 = 1042 ft^3

The check is:

if (start_mr > 0.75 * max_read) and (mr_diff < 0.25 * max_read) then mr_value = mr_diff

In this case we have (9257 > 0.75 * 10000 = 7500) and (1042 < 0.25 * 10000 = 2500) so we set mr_value = 1042 ft^3

This can then be used to calculate a correct meter volume.

Appendix E Example Data Issues

This meter point fails automatically due to the negative volume

MPR ID	METER READ DATE	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	EUC
14770125	31/12/2010	22416282	139936	0	02B
14770125	**31/01/2011**	**22534617**	**-102142986**	**-1**	**02B**
14770125	28/02/2011	22631123	98691	0	02B

This meter point fails due to consumption > 5 * AQ

MPR ID	METER READ DATE	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	EUC
13975325	26/07/2010	85437	409	0	05B
13975325	26/08/2010	85449	1227	0	05B
13975325	31/08/2010	85449	0	0	05B
13975325	**28/09/2010**	**85447**	**30678995**	**3**	**05B**
13975325	26/10/2010	85473	2659	0	04B
13975325	31/10/2010	85479	614	0	04B
13975325	24/11/2010	85502	2352	0	04B

This meter point fails automatically due to the negative volume

MPR ID	METER READ DATE	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	EUC
13979065	21/12/2010	8648510	1787700	0	06W01
13979065	**21/02/2011**	**102832**	**-231410400**	**1**	**06W01**
13979065	**28/02/2011**	**133194**	**0**	**0**	**06W01**
13979065	**01/03/2011**	**9487904**	**982285900**	**0**	**06W01**
13979065	**17/03/2011**	**224221**	**0**	**0**	**06W01**
13979065	28/03/2011	255891	33743	0	06W01
13979065	28/04/2011	322608	69522	0	06W01

This meter point fails automatically due to the negative volumes.

MPR ID	METER READ DATE	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	EUC
13988551	11/09/2008	610290	284231000	0	07B
13988551	06/01/2009	701907	465219300	0	07B
13988551	**27/07/2009**	**107828**	**50013600**	**1**	**07B**
13988551	**24/08/2009**	**730925**	**-985154300**	**0**	**07B**
13988551	23/09/2009	801244	9390200	0	07B
13988551	27/10/2009	875425	9986700	0	07B
13988551	28/10/2009	875426	100	0	07B
13988551	24/11/2009	933954	7879500	0	07B
13988551	15/12/2009	979056	6152300	0	07B

MPR ID	METER READ DATE	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	EUC
13988551	26/01/2010	985826	935400	0	07B
13988551	23/02/2010	985826	0	0	07B
13988551	30/03/2010	985826	0	0	07B
13988551	29/04/2010	985916	12300	0	07B
13988551	27/05/2010	3331	2371000	0	07B
13988551	30/06/2010	82985	10720300	0	07B
13988551	29/07/2010	152618	63970200	0	07B
13988551	**06/08/2010**	**172001**	**-52145600**	**0**	**07B**
13988551	27/08/2010	219720	6342600	0	07B
13988551	29/09/2010	304111	11260600	0	07B

This meter point fails due to consumption > 5 * AQ

MPR ID	METER READ DATE	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	EUC
13975686	18/12/2009	46892	305600	0	03B
13975686	27/01/2010	55551	884900	0	03B
13975686	23/02/2010	61000	552300	0	03B
13975686	01/04/2010	61449	46300	0	03B
13975686	**11/06/2010**	**69287**	**912842800**	**0**	**03B**
13975686	26/11/2010	76978	1221000	0	03B
13975686	**14/12/2010**	**81414**	**0**	**0**	**03B**
13975686	25/01/2011	91140	985780	0	03B

This meter point fails due to consumption > 5 * AQ

MPR ID	METER READ DATE	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	EUC
1553522	26/05/2009	1214	511	0	04B
1553522	29/06/2009	1220	614	0	04B
1553522	**27/07/2009**	**999991**	**102138318**	**0**	**04B**
1553522	24/08/2009	999991	0	0	04B
1553522	29/09/2009	999995	409	0	04B
1553522	03/10/2009	1222	125478	1	04W04
1553522	04/10/2009	0	0	0	04W04
1553522	**26/10/2009**	**999999**	**102263898**	**0**	**04W04**
1553522	27/10/2009	1222	125069	1	04W04
1553522	26/03/2010	11	11	0	04W04
1553522	13/04/2010	13	2	0	04W04

MPR_ID	METER READ DATE	IMP_IND	METER READ VAL	METERED VOL	ROUND THE CLOCK IND	AQ	METER READ FREQ	SSP_LSP	EUC	READ_TYPE_CODE
16423383	04-Mar-10	Y	3250	550283	1	40905	A	SSP	WN:E0901B	UR
16423383	04-Mar-10	Y	3250	-472357	0	40905	A	SSP	WN:E0901B	UR
16423593	26-Aug-08	Y	54468	88356	0	465947	M	LSP	WN:E0703B	UR
16423593	26-Aug-08	Y	53604	0	0	465947	M	LSP	WN:E0703B	UR
16492640	11-May-12	N	3951	353	0	8415	6	SSP	WN:E1101B	UR
16492640	11-May-12	N	10394	6942	0	8415	6	SSP	WN:E1101B	UR
16625454	03-Mar-10	Y	849	178042	1	74119	6	LSP	WN:E0902B	UR
16625454	03-Mar-10	Y	849	-844598	0	74119	6	LSP	WN:E0902B	UR

Example duplicate meter reads. These are manually processed for large sites.

MPR_ID	AQ_EFFECTIVE_DATE	EUC	AQ	SITE_TYPE_FLAG	LDZ	LATEST	EUC_CALC
1773289	01/10/2007	04B	67015	N	EA	Y	01B
1773289	01/10/2008	03W03	64922	N	EA	Y	01B
1773289	01/10/2009	03W01	64922	N	EA	Y	01B
1773289	01/10/2010	03B	64338	N	EA	Y	01B
1773289	01/10/2011	03B	1	N	EA	Y	01B

Example of inconsistent EUCs. There were 9848 sites in EA where the EUC was inconsistent with the AQ.

Appendix F Average Consumption by EUC Band and LDZ

2009	EA	EM	NE	NO	NT	NW	SC	SE	SO	SW	WM	WN	WS	Average
01B	16,085	16,175		16,401		16,505	17,517		15,889	14,548	16,284	15,212	16,344	16,096
02B	132,867	135,744		135,921		134,259	133,046		135,806	134,386	136,877	137,417	134,580	135,090
03B	451,155	446,984		446,293		453,530	463,569		447,868	446,128	450,688	441,087	449,458	449,676
04B	1,200,057	1,190,358		1,214,550		1,188,242	1,210,405		1,193,914	1,168,392	1,204,857	1,185,729	1,214,505	1,197,101
05B	3,389,271	3,353,033		3,332,055		3,396,346	3,303,928		3,271,230	3,306,426	3,414,839	3,500,815	3,501,154	3,376,909
06B	8,933,413	8,772,846		8,762,077		8,981,293	8,688,459		9,544,058	8,867,677	8,663,026	8,868,863	8,753,373	8,883,509
07B	19,274,714	21,150,206		19,532,804		19,451,144	18,493,830		20,761,850	19,193,096	19,606,662	18,344,380	18,407,408	19,421,610
08B	35,436,931	37,846,272		41,584,214		36,240,311	39,790,081		40,387,341	36,823,274	37,836,534		37,868,543	38,201,500
09B		60,366,036				67,220,768	222,109,824				84,473,844		131,326,752	113,099,445

2010	EA	EM	NE	NO	NT	NW	SC	SE	SO	SW	WM	WN	WS	Average
01B	16,313	16,341		16,358		16,398	17,217		16,068	14,475	16,353	15,149	16,096	16,077
02B	133,894	136,159		136,538		134,658	133,743		135,998	135,995	136,446	136,812	134,906	135,515
03B	452,077	447,907		449,435		453,877	460,375		448,462	450,971	449,811	439,756	444,805	449,748
04B	1,173,416	1,192,628		1,229,461		1,194,371	1,217,038		1,194,649	1,154,545	1,215,642	1,215,446	1,207,217	1,199,441
05B	3,386,681	3,412,600		3,397,446		3,432,399	3,258,180		3,277,335	3,382,643	3,408,457	3,200,693	3,523,225	3,367,966
06B	9,397,010	9,446,624		8,386,279		8,841,677	8,709,096		9,075,589	9,338,154	8,687,807	8,332,907	9,055,709	8,927,085
07B	19,907,768	20,676,637		18,904,493		19,130,560	20,118,848		18,816,770	19,719,584	19,666,438	16,546,622	18,790,928	19,227,865
08B	38,661,325	39,574,121		36,809,568		34,171,190	36,646,345		36,113,507	38,296,302	40,516,792		39,252,698	37,782,428
09B				60,904,960		73,783,696	102,945,886		66,424,256		67,775,856		135,237,200	84,511,976

2011	EA	EM	NE	NO	NT	NW	SC	SE	SO	SW	WM	WN	WS	Average
01B	13,827	13,831		13,989		14,112	15,202		13,290	12,012	13,708	12,882	13,646	13,650
02B	135,366	137,529		134,443		137,137	136,690		137,969	136,104	138,303	137,349	136,314	136,720
03B	449,077	445,638		443,816		452,743	457,224		445,212	450,882	445,358	441,616	443,345	447,491
04B	1,161,219	1,188,383		1,234,669		1,182,974	1,214,317		1,184,056	1,174,932	1,197,583	1,178,659	1,166,083	1,188,287
05B	3,354,554	3,448,365		3,506,881		3,403,468	3,362,024		3,362,831	3,446,399	3,381,753	3,234,733	3,592,488	3,409,350
06B	9,141,769	9,437,365		8,670,746		9,526,664	9,440,801		9,178,380	9,206,930	9,027,132	9,492,308	9,109,686	9,223,178
07B	19,330,090	20,424,943		18,798,787		20,157,472	22,118,872		19,755,461	19,534,867	19,997,959	18,371,392	18,031,042	19,652,088
08B	38,742,099	41,202,652		39,722,311		37,305,334	44,353,772		43,592,406	36,065,977	40,660,385		38,564,968	40,023,323
09B		62,165,988				75,169,724					61,798,412		132,316,880	82,862,751

Appendix G Consumption Calculation Output

All consumptions shown in these tables are in kWh.

G.1 EA LDZ

G.1.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,492,470	24,005,811,150	9,464	201,361	108,381	1,802,212	28,987,897,897
02B	14,277	1,896,978,990	56,272	3,598	221	18,097	2,404,435,941
03B	3,350	1,511,209,655	120,094	745	58	4,153	1,873,575,979
04B	1,331	1,597,574,868	383,211	264	41	1,637	1,963,968,071
05B	325	1,100,955,129	958,781	43	30	398	1,348,780,981
06B	87	775,227,550	2,456,176	27	17	130	1,164,862,448
07B	23	443,318,423	2,770,854	5	4	32	618,638,442
08B	4	141,747,722	2,567,323	3	7	14	514,536,565
09B	0	0	0		1	1	113,099,445

G.1.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,575,230	25,697,058,819	9,598	121,675	120,327	1,817,232	29,644,886,593
02B	15,180	2,032,443,754	56,739	2,713	256	18,149	2,430,010,365
03B	3,423	1,547,451,255	119,763	563	65	4,051	1,831,197,034
04B	1,371	1,608,711,077	372,897	272	47	1,690	1,983,435,544
05B	322	1,091,192,607	943,553	68	25	415	1,404,585,102
06B	87	814,743,384	2,423,740	20	17	124	1,162,031,755
07B	24	477,786,428	4,077,890	6	2	32	635,002,667
08B	3	115,983,976	3,692,985	4	3	10	380,460,970
09B	0	0	0	0	0	0	0

G.1.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,594,596	22,048,114,716	8,570	109,048	133,029	1,836,672	25,395,253,378
02B	12,783	1,730,400,789	56,892	2,588	263	15,634	2,116,348,353
03B	2,958	1,328,258,166	120,160	531	74	3,563	1,600,027,287
04B	1,131	1,313,662,191	364,552	249	50	1,430	1,660,746,803
05B	255	856,167,560	993,149	62	30	348	1,165,899,017
06B	82	752,353,028	2,336,926	18	14	115	1,047,072,948
07B	23	444,592,061	3,705,174	9	1	33	641,112,945
08B	4	154,968,396	6,338,672	3	0	7	258,943,763
09B	0	0	0	0	0	0	0

G.2 EM LDZ

G.2.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,855,569	30,016,856,205	9,102	281,011	132,732	2,269,312	36,709,816,822
02B	16,418	2,231,407,873	57,426	4,318	138	20,873	2,836,994,738
03B	4,008	1,795,039,609	117,569	828	31	4,867	2,179,752,855
04B	1,541	1,837,202,983	383,908	290	94	1,925	2,295,306,504
05B	419	1,414,039,126	937,363	78	6	503	1,696,818,223
06B	138	1,213,845,168	2,284,032	23	4	165	1,453,370,214
07B	45	937,909,961	4,285,816	13	0	58	1,211,229,264
08B	11	404,182,914	5,003,465	6	0	17	648,848,843
09B	1	60,584,952	0	0	0	1	60,584,952

G.2.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,950,885	31,882,515,396	9,204	188,226	144,940	2,284,051	37,327,314,201
02B	17,090	2,327,950,451	57,306	3,442	204	20,736	2,824,575,142
03B	4,066	1,821,689,427	118,453	797	43	4,905	2,197,993,410
04B	1,572	1,880,727,683	390,447	310	104	1,986	2,375,608,176
05B	434	1,487,990,520	972,959	58	7	499	1,711,212,127
06B	158	1,487,278,705	2,450,443	18	5	181	1,699,662,898
07B	40	819,706,463	4,122,718	11	1	52	1,058,334,728
08B	21	827,561,380	7,613,292	4	0	25	990,767,361
09B	0	0	0	0	0	0	0

G.2.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,979,597	27,379,293,833	8,210	165,918	158,805	2,304,320	31,870,450,770
02B	14,885	2,047,166,580	57,621	2,933	243	18,062	2,484,005,107
03B	3,468	1,545,267,566	117,957	678	48	4,193	1,868,617,491
04B	1,300	1,544,916,034	385,135	295	114	1,709	2,031,058,594
05B	371	1,279,031,503	974,167	64	10	445	1,534,033,451
06B	138	1,305,836,463	2,433,107	28	2	169	1,590,936,550
07B	46	930,274,171	4,105,838	12	2	60	1,215,794,519
08B	22	888,182,244	7,819,854	1	0	23	928,205,567
09B	2	124,331,976	3,964,855	00	0	2	124,331,976

G.3 NO LDZ

G.3.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	974,069	15,975,784,735	9,046	156,801	56,616	1,187,486	19,476,048,625
02B	8,504	1,155,880,739	56,870	1,977	43	10,524	1,430,403,202
03B	2,332	1,040,558,452	113,493	456	18	2,806	1,252,250,522
04B	890	1,081,367,386	407,552	166	9	1,065	1,293,978,272
05B	245	815,580,701	965,593	47	3	295	983,355,722
06B	74	648,393,706	2,432,574	15	0	89	779,824,862
07B	17	332,057,676	3,917,600	2	0	19	370,900,895
08B	6	249,505,284	10,885,585	1	0	7	287,706,784
09B	0	0	0	0	0	0	0

G.3.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,030,361	16,854,161,330	9,108	100,843	60,133	1,191,337	19,487,325,219
02B	8,810	1,202,898,184	56,735	1,498	59	10,367	1,415,528,221
03B	2,335	1,049,294,603	115,513	397	19	2,751	1,236,227,620
04B	867	1,065,593,376	400,704	161	8	1,035	1,272,842,569
05B	244	827,846,505	992,028	44	2	290	985,609,512
06B	73	609,948,547	2,033,218	9	0	82	685,425,058
07B	18	340,280,882	2,936,659	5	0	23	430,262,220
08B	6	220,857,410	7,921,439	2	0	8	296,422,265
09B	1	60,904,960	0	0	0	1	60,904,960

G.3.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,044,778	14,615,488,589	8,156	89,145	68,719	1,202,642	16,823,857,676
02B	6,740	906,162,854	56,577	2,395	101	9,236	1,241,779,234
03B	1,676	743,710,171	114,450	757	20	2,452	1,088,449,263
04B	641	791,305,590	386,667	289	17	947	1,169,618,707
05B	166	581,066,355	1,017,229	74	6	246	861,818,470
06B	64	552,927,321	2,048,241	11	0	75	651,709,973
07B	17	319,579,374	2,395,949	6	0	23	429,372,520
08B	8	317,778,484	7,602,648	1	0	9	357,801,807
09B	0	0	0	0	1	1	82,862,751

G.4 NW LDZ

G.4.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	2,223,768	36,703,970,463	9,693	362,971	102,698	2,689,436	44,389,977,835
02B	21,161	2,841,098,889	56,852	5,738	312	27,211	3,653,318,810
03B	4,845	2,197,488,409	119,376	1,108	68	6,022	2,731,019,681
04B	2,075	2,465,505,451	383,622	373	40	2,488	2,956,779,221
05B	541	1,836,549,719	956,522	119	9	668	2,270,294,915
06B	120	1,080,901,537	2,328,145	32	2	154	1,382,455,301
07B	35	686,777,311	3,759,759	9	4	49	944,725,879
08B	7	253,682,180	7,854,346	3	7	17	647,849,031
09B	1	67,220,768	0	2	0	3	293,419,658

G.4.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	2,107,676	34,560,831,193	9,701	481,509	111,321	2,700,506	44,281,820,608
02B	19,271	2,595,021,711	57,135	6,830	353	26,454	3,562,253,100
03B	4,313	1,957,563,379	119,811	1,543	76	5,932	2,692,359,960
04B	1,796	2,145,404,973	382,113	633	38	2,467	2,946,297,144
05B	471	1,615,349,150	948,506	157	12	640	2,196,784,022
06B	124	1,100,122,985	2,380,009	37	4	166	1,464,764,038
07B	34	650,439,036	3,813,950	10	2	46	870,968,088
08B	11	372,584,684	3,576,286	4	1	16	561,496,823
09B	1	73,783,696	0	3	6	10	819,144,760

G.4.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	2,367,773	33,412,995,491	8,756	225,625	121,654	2,715,052	38,313,650,324
02B	18,561	2,545,378,215	57,489	4,350	385	23,296	3,194,669,747
03B	4,442	2,011,305,205	119,927	847	76	5,366	2,429,288,826
04B	1,704	2,015,400,197	376,339	350	44	2,097	2,481,096,007
05B	478	1,627,141,504	910,046	94	16	588	2,001,388,972
06B	123	1,176,138,177	2,515,466	28	8	159	1,515,851,289
07B	39	789,792,399	3,821,575	9	0	48	967,332,927
08B	9	325,677,236	6,235,812	2	0	11	408,876,441
09B	2	150,339,448	6,596,192	1	7	10	813,241,456

G.5 SC LDZ

G.5.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,477,562	25,882,227,022	10,479	227,263	130,937	1,835,762	32,156,760,981
02B	15,202	2,022,611,246	57,235	3,920	176	19,298	2,567,575,987
03B	4,114	1,907,050,422	121,787	1,068	90	5,272	2,443,728,413
04B	1,938	2,346,216,382	392,630	377	33	2,349	2,842,651,454
05B	491	1,621,466,814	1,014,333	105	19	615	2,032,157,928
06B	108	938,353,573	2,164,023	22	2	132	1,143,020,596
07B	25	462,345,762	3,902,779	9	1	35	650,584,008
08B	3	119,370,242	13,105,954	2	6	11	424,982,242
09B	1	222,109,824	0	0	0	1	222,109,824

G.5.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,555,242	26,776,771,108	10,419	156,388	148,279	1,859,910	32,022,257,492
02B	15,944	2,132,350,663	56,902	2,882	245	19,071	2,550,588,279
03B	4,292	1,975,984,639	120,469	742	98	5,132	2,362,473,070
04B	2,018	2,455,974,615	403,963	354	43	2,415	2,938,595,244
05B	493	1,607,377,300	973,675	98	21	612	1,995,110,379
06B	132	1,149,600,729	2,420,385	16	5	153	1,336,386,202
07B	26	514,541,968	4,311,960	8	0	33	666,245,592
08B	3	109,939,036	6,444,718	3	1	7	243,140,917
09B	2	255,124,368	24,443,428	0	0	2	255,124,368

G.5.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,575,668	23,953,918,672	9,599	145,887	162,920	1,884,475	28,648,526,844
02B	14,186	1,939,022,947	58,161	2,666	324	17,175	2,347,674,017
03B	3,990	1,824,424,735	120,435	677	127	4,794	2,191,973,978
04B	1,755	2,131,206,120	390,530	334	45	2,134	2,591,255,276
05B	411	1,382,550,004	1,020,636	93	28	532	1,788,201,773
06B	117	1,104,575,138	2,345,305	26	8	151	1,428,227,730
07B	26	584,873,233	4,534,536	7	7	40	852,407,193
08B	5	203,459,636	6,777,777	2	1	8	323,529,604
09B	0	0	0	1	0	1	82,862,751

G.6 SO LDZ

G.6.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,331,419	21,155,070,134	9,451	189,909	91,059	1,612,386	25,619,401,958
02B	13,816	1,876,242,434	57,258	3,210	181	17,207	2,336,820,121
03B	3,273	1,466,054,135	119,621	636	52	3,962	1,774,341,520
04B	1,373	1,638,691,948	378,060	210	18	1,601	1,911,381,220
05B	282	922,115,018	916,060	63	9	354	1,158,200,197
06B	86	816,180,678	2,592,409	13	5	103	985,986,862
07B	25	510,036,531	3,891,558	4	1	30	607,144,579
08B	6	227,899,370	8,828,145	1	0	7	266,100,870
09B	0	0	0	0	1	1	113,099,445

G.6.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,389,322	22,323,227,430	9,525	135,392	100,605	1,625,319	26,115,157,804
02B	14,397	1,957,939,944	56,923	2,445	220	17,062	2,320,424,544
03B	3,371	1,511,669,898	120,333	555	44	3,969	1,780,097,549
04B	1,379	1,647,160,664	382,915	226	28	1,633	1,950,276,726
05B	288	945,240,484	909,950	53	5	346	1,135,124,927
06B	85	769,242,878	2,399,751	16	2	103	930,561,928
07B	19	357,518,623	3,313,407	5	2	26	482,967,635
08B	8	288,908,058	9,648,330	0	0	8	303,088,200
09B	1	66,424,256	0	0	0	1	66,424,256

G.6.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,420,582	18,879,016,206	8,389	111,504	107,993	1,640,079	21,796,049,678
02B	12,129	1,673,487,287	57,105	2,474	260	14,863	2,050,640,070
03B	2,795	1,244,588,487	120,291	540	55	3,390	1,509,426,865
04B	1,120	1,325,820,201	381,758	228	34	1,381	1,635,712,229
05B	237	796,914,983	978,837	56	6	299	1,006,565,650
06B	73	668,061,678	2,280,627	14	4	90	829,133,603
07B	17	326,271,260	4,217,178	6	0	22	437,056,480
08B	6	261,554,438	11,235,021	3	1	10	428,101,727
09B	0	0	0	0	0	0	0

G.7 SW LDZ

G.7.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,420,582	18,879,016,206	8,389	111,504	107,993	1,640,079	21,796,049,678
02B	12,129	1,673,487,287	57,105	2,474	260	14,863	2,050,640,070
03B	2,795	1,244,588,487	120,291	540	55	3,390	1,509,426,865
04B	1,120	1,325,820,201	381,758	228	34	1,381	1,635,712,229
05B	237	796,914,983	978,837	56	6	299	1,006,565,650
06B	73	668,061,678	2,280,627	14	4	90	829,133,603
07B	17	326,271,260	4,217,178	6	0	22	437,056,480
08B	6	261,554,438	11,235,021	3	1	10	428,101,727
09B	0	0	0	0	0	0	0

G.7.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,259,293	18,228,519,908	9,020	106,952	104,090	1,470,336	21,283,400,889
02B	11,328	1,540,542,782	56,430	2,255	168	13,751	1,870,092,512
03B	2,529	1,140,496,594	119,847	512	58	3,099	1,397,551,086
04B	970	1,119,550,162	373,225	214	17	1,201	1,386,451,086
05B	208	703,360,883	1,012,022	43	7	258	872,650,268
06B	86	806,382,798	2,607,449	20	2	109	1,014,879,232
07B	22	433,830,841	3,505,355	6	1	29	568,425,895
08B	4	153,185,208	5,436,638	0	0	4	153,878,389
09B	0	0	0	0	0	0	0

G.7.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,259,293	18,228,519,908	9,020	106,952	104,090	1,470,336	21,283,400,889
02B	11,328	1,540,542,782	56,430	2,255	168	13,751	1,870,092,512
03B	2,529	1,140,496,594	119,847	512	58	3,099	1,397,551,086
04B	970	1,119,550,162	373,225	214	17	1,201	1,386,451,086
05B	208	703,360,883	1,012,022	43	7	258	872,650,268
06B	86	806,382,798	2,607,449	20	2	109	1,014,879,232
07B	22	433,830,841	3,505,355	6	1	29	568,425,895
08B	4	153,185,208	5,436,638	0	0	4	153,878,389
09B	0	0	0	0	0	0	0

G.8 WM LDZ

G.8.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,634,482	26,616,423,148	9,339	256,529	92,241	1,983,252	32,295,902,288
02B	15,409	2,109,167,717	57,312	3,777	109	19,295	2,641,107,177
03B	3,910	1,762,157,834	119,935	750	34	4,694	2,115,515,316
04B	1,692	2,039,036,961	388,991	294	14	2,000	2,410,019,737
05B	435	1,486,797,114	998,197	99	7	541	1,847,648,943
06B	100	870,606,128	2,136,115	29	4	134	1,160,802,959
07B	33	647,019,854	3,406,737	10	0	43	843,086,476
08B	19	718,894,146	8,026,192	5	1	25	930,961,409
09B	2	168,947,688	27,713,093	0	0	2	168,947,688

G.8.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,634,482	26,616,423,148	9,339	256,529	92,241	1,983,252	32,295,902,288
02B	15,409	2,109,167,717	57,312	3,777	109	19,295	2,641,107,177
03B	3,910	1,762,157,834	119,935	750	34	4,694	2,115,515,316
04B	1,692	2,039,036,961	388,991	294	14	2,000	2,410,019,737
05B	435	1,486,797,114	998,197	99	7	541	1,847,648,943
06B	100	870,606,128	2,136,115	29	4	134	1,160,802,959
07B	33	647,019,854	3,406,737	10	0	43	843,086,476
08B	19	718,894,146	8,026,192	5	1	25	930,961,409
09B	2	168,947,688	27,713,093	0	0	2	168,947,688

G.8.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	1,746,541	23,941,139,336	8,289	151,906	108,773	2,007,220	27,514,462,233
02B	13,401	1,853,376,152	57,476	3,019	198	16,617	2,298,235,695
03B	3,285	1,462,978,602	119,270	765	49	4,099	1,825,596,518
04B	1,402	1,678,674,661	382,359	336	20	1,758	2,105,222,819
05B	369	1,246,667,845	984,960	74	9	452	1,527,476,843
06B	109	981,437,458	2,369,120	22	4	134	1,212,793,147
07B	40	809,770,430	4,349,301	4	1	45	900,864,962
08B	17	691,226,544	8,506,318	1	1	19	754,133,597
09B	1	61,798,412	0	0	0	1	61,798,412

G.9 WN LDZ

G.9.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	200,611	3,051,633,702	9,093	30,190	10,473	241,274	3,670,179,836
02B	1,753	240,840,922	56,979	574	26	2,353	323,358,965
03B	427	188,253,368	116,152	116	6	549	242,225,521
04B	207	245,445,900	383,959	40	1	248	293,546,238
05B	63	220,532,994	1,122,960	7	1	71	249,308,356
06B	20	177,377,263	2,087,152	2	0	22	196,728,470
07B	4	73,377,520	2,818,447	3	1	8	151,063,958
08B	0	0	0	0	0	0	0
09B	0	0	0	0	0	0	0

G.9.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	211,480	3,203,672,241	9,062	20,268	11,786	243,534	3,689,253,237
02B	1,966	268,906,738	56,400	392	24	2,382	325,855,787
03B	478	210,287,791	117,512	81	6	565	248,504,436
04B	201	244,017,827	389,716	35	3	239	290,339,808
05B	56	179,238,793	928,871	10	0	66	211,245,720
06B	21	174,991,044	1,968,879	3	0	24	203,619,124
07B	5	82,733,111	1,550,137	3	2	10	177,538,740
08B	0	0	0	0	0	0	0
09B	0	0	0	0	0	0	0

G.9.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	216,830	2,793,258,239	8,154	16,425	12,897	246,152	3,170,997,519
02B	1,797	246,849,203	57,882	361	22	2,181	299,520,219
03B	387	170,801,133	116,291	75	5	466	205,945,152
04B	195	230,041,108	368,603	28	4	227	268,133,018
05B	41	132,457,002	885,837	11	1	53	171,273,797
06B	21	196,173,734	2,825,315	5	1	27	256,122,830
07B	6	110,228,350	2,725,634	2	0	8	149,532,527
08B	0	0	0	0	0	0	0
09B	0	0	0	0	0	0	0

G.10 WS LDZ

G.10.1 2009

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	653,687	10,683,650,740	9,190	132,989	42,538	829,214	13,552,412,742
02B	5,176	696,631,939	56,567	1,578	87	6,841	920,729,753
03B	1,247	560,373,632	119,659	321	39	1,606	722,047,970
04B	494	600,199,862	392,699	117	8	619	751,560,279
05B	134	468,746,082	1,000,450	26	2	162	566,737,433
06B	46	402,655,177	2,315,619	15	5	66	575,953,113
07B	7	128,851,857	4,216,388	3	1	11	198,842,971
08B	4	151,474,172	1,708,815	1	0	5	189,675,672
09B	1	131,326,752	0	0	0	1	131,326,752

G.10.2 2010

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	705,447	11,354,749,837	9,113	83,069	45,053	833,569	13,416,985,861
02B	5,598	755,166,183	56,768	999	101	6,698	903,563,532
03B	1,291	574,198,032	117,021	238	38	1,567	696,816,940
04B	507	612,630,920	374,108	95	12	614	741,354,350
05B	141	496,774,794	1,058,653	23	3	167	588,295,919
06B	47	428,929,740	2,321,645	7	3	58	521,445,844
07B	10	187,909,282	3,410,501	3	1	14	265,486,057
08B	5	196,263,490	7,595,356	0	0	5	196,263,490
09B	1	135,237,200	0	0	0	1	135,237,200

G.10.3 2011

EUC	Sample Size	Sample Metered Consumption	SD	Failed Sites	Sites in CSEPs	Population Size	Population Metered Consumption
01B	724,311	9,883,808,302	8,186	67,378	49,099	840,789	11,473,236,991
02B	4,653	634,229,767	56,843	1,127	114	5,894	803,389,671
03B	1,075	476,650,217	117,728	224	37	1,336	592,527,051
04B	445	519,415,354	358,631	93	11	549	640,677,045
05B	114	411,336,705	1,078,215	33	3	150	540,429,636
06B	40	367,840,058	2,089,241	12	3	56	508,627,078
07B	13	227,089,370	2,287,643	2	1	16	293,546,358
08B	5	192,824,838	5,938,054	1	0	6	246,758,114
09B	1	132,316,880	0	0	0	1	132,316,880

Appendix H UG Summary

LDZ	Formula Year	Allocation (GWh)	Consumption (GWh)	Temporary UG and Identified Theft (GWh)	UG (GWh)	Low (%)	Best Estimate (%)	High (%)
EA	2009	39,805	38,990	47	768	1.82%	1.93%	2.04%
EA	2010	41,095	39,472	47	1,576	3.74%	3.84%	3.94%
EA	2011	34,080	33,885	46	149	0.32%	0.44%	0.55%
EM	2009	50,122	49,093	53	977	1.84%	1.95%	2.06%
EM	2010	51,601	50,185	53	1,363	2.53%	2.64%	2.75%
EM	2011	43,028	43,647	51	-670	-1.68%	-1.56%	-1.44%
NE	2009		Data received					
NE	2010		Data received					
NE	2011		Data received					
NO	2009	26,386	25,874	36	476	1.64%	1.80%	1.97%
NO	2010	26,711	25,871	35	805	2.89%	3.01%	3.14%
NO	2011	22,677	22,707	34	-64	-0.42%	-0.28%	-0.15%
NT	2009		Data received					
NT	2010		Data received					
NT	2011		Data received					
NW	2009	60,438	59,270	91	1,077	1.67%	1.78%	1.89%
NW	2010	60,507	59,283	89	1,134	1.78%	1.87%	1.97%
NW	2011	51,158	52,058	85	-985	-2.03%	-1.93%	-1.82%
SC	2009	45,827	44,484	74	1,270	2.58%	2.77%	2.96%
SC	2010	45,973	44,370	74	1,529	3.20%	3.33%	3.45%
SC	2011	40,062	40,255	73	-266	-0.79%	-0.66%	-0.53%
SE	2009		Data received					
SE	2010		Data received					
SE	2011		Data received					
SO	2009	35,408	34,772	39	596	1.56%	1.68%	1.81%
SO	2010	36,078	35,084	39	955	2.54%	2.65%	2.76%
SO	2011	29,917	29,693	39	186	0.47%	0.62%	0.78%
SW	2009	28,915	28,381	44	490	1.56%	1.70%	1.83%
SW	2010	29,162	28,547	44	571	1.83%	1.96%	2.08%
SW	2011	24,137	24,199	43	-105	-0.56%	-0.43%	-0.30%
WM	2009	44,697	44,414	40	243	0.40%	0.54%	0.69%
WM	2010	45,639	44,789	39	811	1.67%	1.78%	1.88%
WM	2011	37,866	38,201	38	-372	-1.11%	-0.98%	-0.86%
WN	2009	5,340	5,126	2	211	3.64%	3.95%	4.26%
WN	2010	5,373	5,146	2	225	3.95%	4.18%	4.42%
WN	2011	4,495	4,522	2	-29	-0.98%	-0.64%	-0.31%
WS	2009	18,287	17,609	23	655	3.43%	3.58%	3.78%
WS	2010	18,099	17,465	23	611	3.23%	3.37%	3.54%
WS	2011	15,062	15,232	22	-192	-1.39%	-1.27%	-1.06%

Appendix I 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).

I.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.

I.2 FACTORS

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

Name	Description
LDZ	VARCHAR2(2) - LDZ identifier e.g. EA
EUC_BAND	VARCHAR2(5) – EUC band e.g. 08W02
GAS_DAY	Date
ALP	Raw data
DAF	Raw data
EWCF	Raw data
CV	Raw data
ENERGY_PROFILE	ALP * (1 + DAF * EWCF)
VOL_PROFILE	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	VARCHAR2(2) - LDZ identifier e.g. EA
EUC_BAND	VARCHAR2(5) – EUC band e.g. 08W02
F_YEAR	NUMBER
ENERGY_PROFILE	ALP * (1 + DAF * EWCF)
VOL_PROFILE	ENERGY_PROFILE / CV

I.3 METER_DIALS

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

I.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.

I.5 T_RESULTS

Name	Description
MPR_ID	NUMBER – dummy MPR ID
SSP_LSP	Varchar2 – 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	Varchar2 – LDZ that the theft occurred in

I.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

Appendix J Theft of Gas – Unregistered Meters

This table contains a list of sites provided by Xoserve that were unregistered at the time of the occurrence of theft. Note that the dummy MPR IDs do not match those for the meter read/consumption population.

MPR ID	TOG Start Date	TOG End Date	Min Conf Date	Comments
233	28/05/1993	15/05/2006	01/10/1995	TOG Prior to First Conf
1342	01/01/2008	22/04/2009	01/03/1996	No Shipper between 2002 and 2008
1936	02/02/2002	15/04/2008	01/03/1996	No Shipper between 2002 and 2002
2485	14/01/1998	25/04/2007	01/03/1996	Period of Non Ownership During TOG
3189	23/05/1996	23/05/2008	01/03/1996	No Shipper between 2001 and 2008
3280	01/05/2009	28/07/2010	01/03/1996	No Shipper between 2004 and 2010
3584	07/10/1994	24/02/2006	01/03/1996	TOG Prior to First Conf
7130	08/02/2006	03/03/2011	01/03/1996	No Shipper between 2005 and 2006
3869	01/05/2006	15/02/2007	01/03/1996	No Shipper between 2002 and 2006
4008	05/04/2003	26/04/2006	25/07/1999	No Shipper between 2007 and 2008
4061	05/01/2005	21/07/2007	12/04/1999	No Shipper between 2001 and 2006
4107	17/08/2009	07/10/2009	25/09/2009	TOG Prior to Min Conf Start
4108	13/10/2005	18/01/2006	28/09/2000	No Shipper between 2006 and 2007
4126	15/01/2008	18/03/2008	14/08/2000	No Shipper between 2004 and 2010
4182	29/05/2001	06/12/2006	13/08/2001	TOG Prior to Min Conf Start
4183	29/05/2001	12/09/2001	27/05/2002	TOG Prior to Min Conf Start
7338	17/03/2004	12/02/2010	05/04/2004	TOG Prior to Min Conf Start
4257	04/12/1991	23/11/2007	09/01/2004	TOG Prior to Min Conf Start
4270	09/02/2002	01/02/2008	21/02/2003	TOG Prior to Min Conf Start
4285	10/07/2002	04/05/2007	02/12/2004	TOG Prior to Min Conf Start
4303	12/01/2006	12/01/2007	28/04/2006	TOG Prior to Min Conf Start
4312	30/12/2005	09/05/2006	12/01/2006	TOG Prior to Min Conf Start
4313	28/05/2004	15/07/2009	23/06/2004	TOG Prior to Min Conf Start
4329	24/01/2007	08/04/2008	20/06/2007	TOG Prior to Min Conf Start
4338	19/02/2006	29/07/2008	25/05/2006	TOG Prior to Min Conf Start
4340	01/04/2005	01/04/2006	25/08/2005	TOG Prior to Min Conf Start
4343	03/06/2004	10/02/2006	16/05/2005	TOG Prior to Min Conf Start
7379	17/06/2010	19/04/2011	24/09/2010	TOG Prior to Min Conf Start
4359	06/07/2003	28/04/2006	23/04/2008	TOG Prior to Min Conf Start
4367	17/10/2005	17/10/2006	16/03/2006	TOG Prior to Min Conf Start
4368	24/05/2005	05/10/2010	30/03/2006	TOG Prior to Min Conf Start
4372	28/05/2004	25/01/2011	04/10/2006	TOG Prior to Min Conf Start
4375	15/07/2005	21/04/2006	13/02/2006	TOG Prior to Min Conf Start
4382	08/06/2005	08/06/2006	17/10/2006	TOG Prior to Min Conf Start
4386	23/11/2006	30/03/2007	30/11/2006	TOG Prior to Min Conf Start
4397	05/04/2006	05/11/2009	08/03/2008	TOG Prior to Min Conf Start
4400	28/11/2005	07/02/2006	26/02/2008	TOG Prior to Min Conf Start
4401	13/08/2007	21/09/2009	11/06/2009	TOG Prior to Min Conf Start
4402	18/06/2008	20/01/2010	21/11/2009	TOG Prior to Min Conf Start
4404	17/09/2007	17/09/2008	02/12/2008	TOG Prior to Min Conf Start
4406	15/12/2008	29/09/2009	14/05/2009	TOG Prior to Min Conf Start

MPR ID	TOG Start Date	TOG End Date	Min Conf Date	Comments
4407	15/03/2008	05/10/2009	22/07/2009	TOG Prior to Min Conf Start
4408	11/09/2005	11/09/2006	30/09/2008	TOG Prior to Min Conf Start
4409	17/11/2009	18/05/2010	19/05/2009	TOG Prior to Min Conf Start
4410	20/01/2009	19/01/2010	01/06/2009	TOG Prior to Min Conf Start
4412	01/01/2001	15/09/2009	31/10/2009	TOG Prior to Min Conf Start
4416	08/01/2009	08/09/2009	06/09/2009	TOG Prior to Min Conf Start
4417	17/10/2008	16/10/2009	16/10/2009	TOG Prior to Min Conf Start
4418	29/07/2006	04/05/2011	11/05/2010	TOG Prior to Min Conf Start
7412	01/07/2009	28/01/2011	21/02/2011	TOG Prior to Min Conf Start
7414	31/12/2006	14/12/2010	14/12/2010	TOG Prior to Min Conf Start
7915	29/05/2001	07/03/2011	26/05/2011	TOG Prior to Min Conf Start

Appendix K Meter Points where Estimated Theft in Formula Year Over-Rides Market Sector Classification

This table shows the meter points where the AQ (where consumption blank) or consumption if calculated successfully over the formula year is overridden by the estimated theft for that formula year. Note that the dummy MPR IDs do not match those for the overall meter read/consumption population although meter reads/metered volumes are provided separately for the theft affected meter points.

The theft in this table is the raw theft apportionment to the formula year, the seasonal normal form is used when deciding on market sector although at a similar level to the figures in this table. This is for information only and does not form part of the proposed methodology.

There are some meters in this table that occur in each formula year.

MPR ID	Formula Year	Theft (kWh)	Consumption (kWh)	AQ (kWh)
25	2007	88994	0	93258
49	2007	75879	22883	103401
53	2007	158779		8789
252	2007	114268		28266
720	2007	372116		12558
832	2007	116643	18439	142419
1583	2007	106919	0	112839
1722	2007	73927		1939
2765	2007	74873	9501	88606
2850	2007	73988	15909	94367
2896	2007	83737	46098	136131
3861	2007	121385	0	126761
4051	2007	189019	73	191509
4128	2007	81189	11889	94927
4205	2007	82552		1795
4319	2007	156263		20600
4322	2007	158566	5843	166677
4391	2007	76686	11721	91464
7187	2007	108192		7284
21	2008	83933	53473	136809
25	2008	95451	0	92667
34	2008	115329	11059	125059
49	2008	80407	24248	102851
53	2008	169375		8789
252	2008	144723		28266
720	2008	387392		12558
832	2008	123500	0	122202
920	2008	106846	153	104543
1165	2008	75493	0	74699
1583	2008	102376		9533
1585	2008	76537	6752	82413
1995	2008	119459		22129
2121	2008	79435	5748	82992
2695	2008	82084		31722
2850	2008	78925	17456	94420

MPR ID	Formula Year	Theft (kWh)	Consumption (kWh)	AQ (kWh)
2894	2008	136167	3021	136418
2896	2008	89325	30657	117530
3098	2008	102082	10001	108083
3648	2008	80442	0	76180
3861	2008	130628	0	126019
4140	2008	77292		14166
4301	2008	144178	85	139174
4327	2008	74233		20600
6380	2008	146934		32081
7187	2008	117702		7284
720	2009	385157	131	389566
832	2009	123330		24411
920	2009	140630	355	140042
1182	2009	74554	7121	85739
1846	2009	115260	4610	119058
2156	2009	156408	4032	167222
2894	2009	135553	0	132822
3557	2009	85102	494	82525
3639	2009	148315		25294
4380	2009	105760		9560
5919	2009	86212	0	86879
5921	2009	129155	0	125218
5968	2009	83207		24467
6320	2009	99523		47773
6380	2009	146732	4595	150305
7079	2009	166702		65263
7187	2009	116855		7284
7347	2009	76584	0	74969
7396	2009	81646		20017
3639	2010	107877		25294
5943	2010	182373		53450
6064	2010	76386		8933
6380	2010	155294	4863	152120
6398	2010	73453	0	68124
6815	2010	115005		7585
7079	2010	174857	85	165023
7187	2010	102285		7284
7292	2010	123155		63480
7303	2010	75053	14803	86233
7344	2010	123506		25925
7396	2010	86410		20017
7527	2010	73659		50144

Appendix L Summary of Theft Split Results: 2011 Method vs Metered + Unmetered Consumption Method

This table shows the summary of theft split results for the 2011 method compared to the metered plus unmetered consumption method described in the Interim Report.

	2007		2008		2009		2010	
	2011 Method LSP%	Alternative Method 1 LSP %	2011 Method LSP %	Alternative Method 1 LSP %	2011 Method LSP %	Alternative Method 1 LSP%	2011 Method LSP%	Alternative Method 1 LSP %
Consumption Calculation Successful	12.7	13.1	15.7	13.0	19.4	15.5	15.8	9.0
Consumption Calculation Failed	20.3	26.6	29.0	29.8	25.6	29.0	28.6	26.5
Combined Consumption and AQ estimate	16.4	19.2	22.5	21.3	23.0	22.6	24.9	20.6

END OF DOCUMENT