GL Noble Denton

Report Number: 13056	30th August 2012

Allocation of Unidentified Gas: Consumption Analysis and Theft Interim Report

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 1

Customer Reference: 1-6AS9AC

This Report is protected by copyright and may not be reproduced in whole or in part by any means without the approval in writing of GL Noble Denton. No Person, other than the Customer for whom it has been prepared, may place reliance on its contents and no duty of care is assumed by GL Noble Denton toward any Person other than the Customer.
This Report must be read in its entirety and is subject to any assumptions and qualifications expressed therein. Elements of this Report contain detailed technical data which is intended for analysis only by persons possessing requisite expertise in its subject matter.

GL Noble Denton is the trading name of GL Industrial Services UK Ltd
Registered in England and Wales No. 3294136 Registered Office: Holywell Park, Ashby Road, Loughborough, Leicestershire, LE11 3GR UK

Report Issue / Amendment Record

Report Title: Allocation of Unidentified Gas: Consumption Analysis and Theft Interim Report	
Report Number: *13056*	**Project Title:** Allocation of Unidentified Gas Statement

Amendment details

Issue	Description of Amendment	Originator/Author
0.1	First draft	Clive Whitehand / Andy Gordon / Tony Perchard
0.2	Updated following review of calculations	Clive Whitehand
0.3	Final review	Clive Whitehand / Andy Gordon / Tony Perchard
1.0	Finalised for publication and consultation	Clive Whitehand / Andy Gordon / Tony Perchard

Executive Summary

Following the publication of the AUGS for 2012/13 [Ref 1] in 2011 and the first draft of the AUGS for 2013/14 [Ref 2] in 2012, specific issues have been raised by the industry in two key areas of the Unidentified Gas (UG) calculation, as follows:

1. The estimate of total UG:
   It is acknowledged in both of the documents referenced above that the method used to date for estimating the UG total is limited in scope by the availability of data. This method is capable of calculating LSP-assigned UG only and therefore carries an inherent assumption that SSP-assigned UG is negligible. The provision of meter read data for all supply points, both LSP and SSP, allows a different method to be used that calculates the total UG directly and makes no assumptions about the relative size of LSP-assigned and SSP-assigned UG.

2. The split of theft by market sector:
   In both of the documents referenced above, the split of theft between the LSP and SSP market sectors was calculated using, for each site at which theft occurred, the pre-theft AQ as the best estimate of the true AQ of the 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 of each site during the period of theft, which may in turn lead to more accurate assignment of market sector to each site.

The Allocation of Unidentified Gas Expert (AUGE) has therefore developed new methods to allow both the estimate of the UG total and the split of theft by market sector to be calculated using consumption data, alongside relevant existing data where required. This document contains details of these new methods and results from them.

The following conclusions were drawn from the analysis:

Consumption Analysis

- The method of calculating total Unidentified Gas using meter read data for both the SSP and LSP market sectors is more accurate than the approximate method employed in the 2011 AUGS for 2012/13. The AUGE therefore concludes that this method should be used in the 2012 AUGS for 2013/14 and for future years.
- When applied to EA LDZ, the new method results in a total Unidentified Gas estimate of 803GWh. This is of the same order as the figure for this LDZ in the 2011 AUGS for 2012/13, which was 841GWh.
- The Confidence Interval associated with this estimate is from 673GWh to 933GWh. Based on the data used in this analysis, we can therefore be 95% sure that the true Unidentified Gas total for EA LDZ lies between these two figures.
- It is estimated that the Confidence Interval for the new method is of a very similar width to that associated with the old method.

Theft Analysis

- Meter read data can also be used in the theft analysis for allocating sites to market sectors more accurately. This produces more accurate results than those presented in the 2011 AUGS for 2012/13, but a number of issues remain regarding this analysis.
- When using meter read data, the average estimate of the proportion of theft that occurs in the LSP market sector is 21.5%.
- The use of detected theft figures in this analysis means that all Shippers are disincentivised to find theft in their own market sectors, and that mixed Shippers have the opportunity to influence the figures by concentrating detection efforts in one market sector over another. Both of these situations are undesirable.
- The failure rate for consumption calculations for theft affected sites is approximately 50%.
- Meter read data has also identified issues with use of pre-theft and post theft AQ as AQs can be affected in different ways and this is not consistent from case to case.
- If theft is split by market sector using throughput instead, this results in an LSP theft percentage of 23.3%.
- The AUGE concludes that the throughput method should be used in the 2012 AUGS for 2013/14 and for future years because it avoids many of the issues identified with the method based on detected theft and results in a more robust methodology.

I

Contents

1 Introduction 5

2 Estimate of the Unidentified Gas Total using Meter Read Data 5

2.1 Total UG Calculation Method 6

2.2 Total UG Estimate Confidence Interval 12

2.3 Results 13

2.4 Comparison with 2011 Method 15

3 Split of Theft by Market Sector 16

3.1 Summary of Original Method: 2011 AUGS for 2012/13 16

3.2 Inclusions and Exclusions 17

3.3 Theft of Gas in Year Overriding AQ 18

3.4 Alternative Methods 19

3.4.1 Alternative Method 1 – Theft Split using Metered and Unmetered Consumption 19

3.4.2 Alternative Method 2 – Fixed Split using Metered and Unmetered Consumption 21

3.4.3 Alternative Method 3 – Fixed Theft Split Varied by Throughput 21

3.4.4 Alternative Method 4 - Throughput 22

3.5 Theft of Gas Results 22

3.6 Discussion of Results 24

3.7 Industry Participant Questions 25

4 Conclusions 26

5 References 26

Appendix A Theft Statistics 27

A.1 Additional Metrics 27

A.2 Examples of sector movements Original vs. Alternative methods 27

1 Introduction

Following the publication of the AUGS for 2012/13 [Ref 1] in 2011 and the first draft of the AUGS for 2013/14 [Ref 2] in 2012, specific issues have been raised by the industry in two key areas of the Unidentified Gas (UG) calculation, as follows:

1. The estimate of total UG:

   It is acknowledged in both of the documents referenced above that the method used to date for estimating the UG total is limited in scope by the availability of data. This method is capable of calculating LSP-assigned UG only and therefore carries an inherent assumption that SSP-assigned UG is negligible. The provision of meter read data for all supply points, both LSP and SSP, allows a different method to be used that calculates the total UG directly and makes no assumptions about the relative size of LSP-assigned and SSP-assigned UG.

2. The split of theft by market sector:

   In both of the documents referenced above, the split of theft between the LSP and SSP market sectors was calculated using, for each site at which theft occurred, the pre-theft AQ as the best estimate of the true AQ of the 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 of each site during the period of theft, which will in turn lead to more accurate assignment of market sector to each site.

The Allocation of Unidentified Gas Expert (AUGE) has therefore developed new methods to allow both the estimate of the UG total and the split of theft by market sector to be calculated using consumption data, alongside relevant existing data where required. This document contains details of these new methods and preliminary results from them.

In the case of the consumption analysis, in view of the timescales required for Xoserve to extract the very large amount of data required, the methods developed and presented in this document have been tested on a single pilot LDZ. EA LDZ was used for this purpose.

In the case of the theft analysis, the methods described in this document were applied to all LDZs. Therefore, in this case, results are provided for all LDZs.

For each analysis, the calculation techniques developed by the AUGE are described in full and preliminary results (i.e. using datasets as they exist now) based on these techniques are provided.

2 Estimate of the Unidentified Gas Total using Meter Read Data

The method of estimating total UG for any given LDZ presented as the primary approach in Ref 1 and Ref 2 is based around RbD and an estimate of allocation algorithm bias. This method is described in detail in Ref 1 and is capable of estimating LSP-assigned UG only (i.e. that element of UG placed in the LSP sector by the current allocation process). This restriction in the scope of the calculation is necessary due to the fact that it is not possible to calculate SSP-assigned UG without using full consumption data from this market sector, and in effect creates an assumption in this method that SSP-assigned UG is negligible. LSP-assigned UG is therefore used as the best available estimate of total UG.

This limitation is acknowledged in both Ref 1 and Ref 2, and each document contains a reference to an alternative method of estimating total UG using full meter read (i.e. consumption) data for all MPRNs from both the LSP and SSP market sectors. This data has now been supplied to the AUGE (for EA LDZ only, which is the pilot LDZ used in this study).

An assessment of the effectiveness of the new method can therefore now take place. In Sections 2.1 to 2.4 below, the new method is first described in detail, and then applied to the data supplied by Xoserve for EA LDZ. Results are given and compared to those obtained from the existing method.

2.1 Total UG Calculation Method

Given that SSP and LSP consumption data was not previously available to the AUGE, a default method of estimating total UG using LSP-assigned UG has been developed as described in Ref 1 and Ref 2. This calculation can be stated as follows:

$$\begin{aligned} \textit{LSP Assigned UG} &= \textit{Alloc LSP} - \textit{Metered LSP} - \textit{Model Bias} \quad (2.1)\\ &= \textit{RbD} - \textit{Model Bias} \end{aligned}$$

Now that data for both SSP and LSP (metered) consumptions and SSP and LSP allocations has been made available, then as long as this data is of sufficient quality, a more rigorous approach can be used. In order to use this approach, it is necessary to have enough data to **accurately** estimate total LDZ SSP and LSP metered consumptions and aggregate allocations.

The use of the scaling factor SF in the allocation algorithm ensures that the aggregated LDZ allocations are scaled up to the correct total. Therefore, total UG can be estimated as follows:

$$\textit{Total UG} = (\textit{Alloc SSP} + \textit{Alloc LSP}) - (\textit{Metered SSP} + \textit{Metered LSP}) \quad (2.2)$$

This can be alternatively stated as:

$$\textit{Total UG} = \textit{Aggregate LDZ Load} - \textit{DM Load} - \textit{Shrinkage} - (\textit{Metered SSP} + \textit{Metered LSP}) \quad (2.3)$$

This is the case because the aggregate allocations are scaled to total LDZ load with DM and shrinkage removed.

Using the first version of this equation (2.2), this creates a requirement for the following data for each LDZ for a full UG calculation method 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.

The calculations are 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 meters loads must be used in the UG alternative method calculations.

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

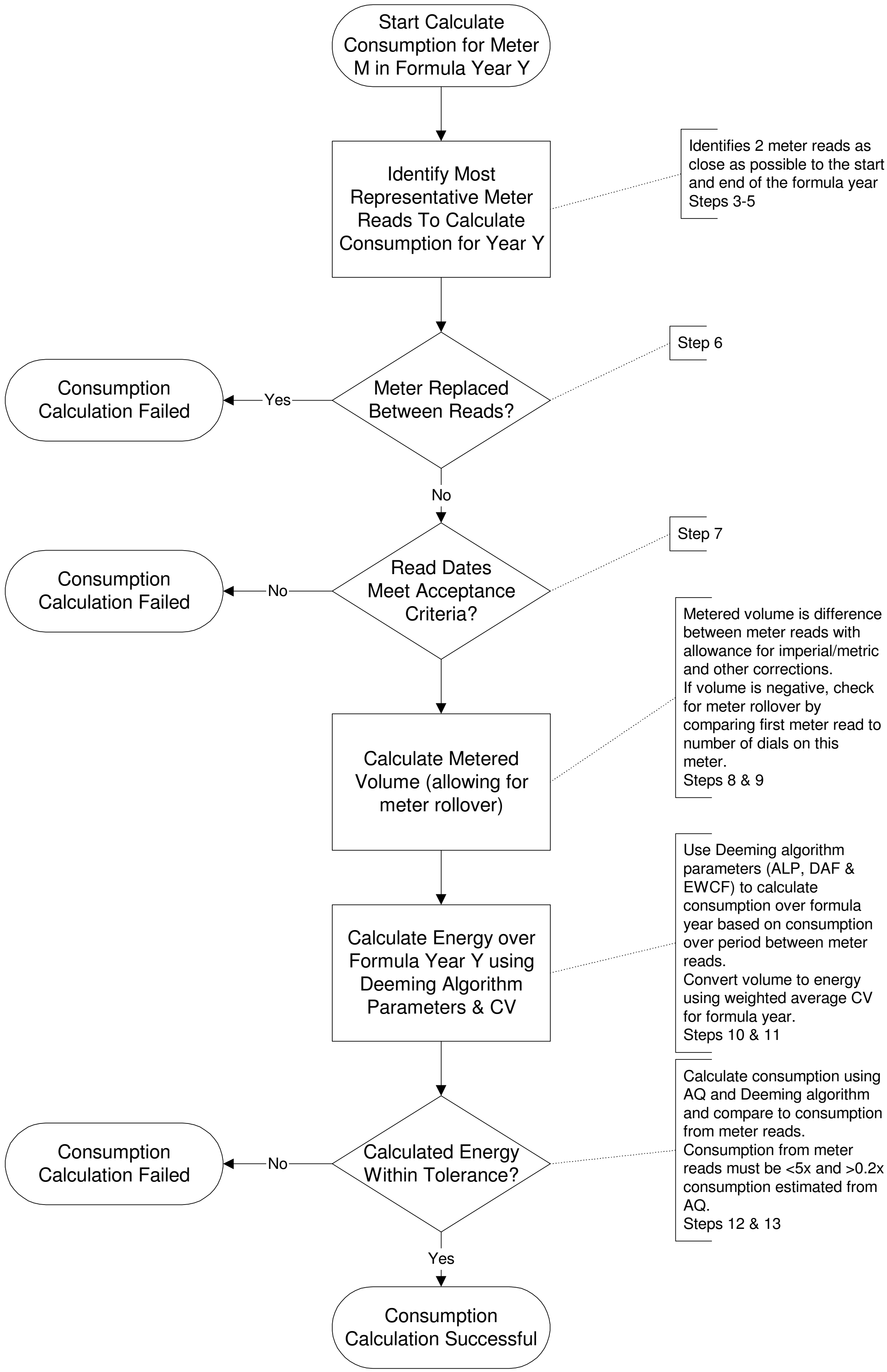
- Allocation data has been provided on a day-by-day basis, split by End User Category (EUC). This data includes CSEP allocations.
- Meter read data has been supplied on an MPRN-by-MPRN basis, with one record for each meter read. Therefore, the volume of data supplied for each MPRN is dependent on the meter read frequency for

that meter. In addition to meter reads, the EUC and the AQ was provided for each MPRN so that metered 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 was also provided. This list also included both EUC and AQ. Therefore, the total number of MPRNs in each EUC could be obtained by adding the count of meters in the consumption data file to the count of meters in the "no meter reads" file.
- Lists of all new sites and lost sites during the analysis period were also supplied including start/end dates. These were used to accurately track the population over time and to ensure that each new or lost site was only included in calculations for the time period for which it was active.
- Aggregate MPRN count and AQ data by EUC for CSEPs was provided in a separate file. 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.
- 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 had negative consumption to determine if meter rollover has occurred.

The provision of this data allows the consumption for each individual meter, for each formula year of interest, to be calculated using the following step-by-step method. Figure 1 shows the consumption calculation element, which covers the first 13 steps of the process. The resulting consumptions are then aggregated and scaled up to cover the full population.

FIGURE 1: CONSUMPTION CALCULATION FLOW CHART



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

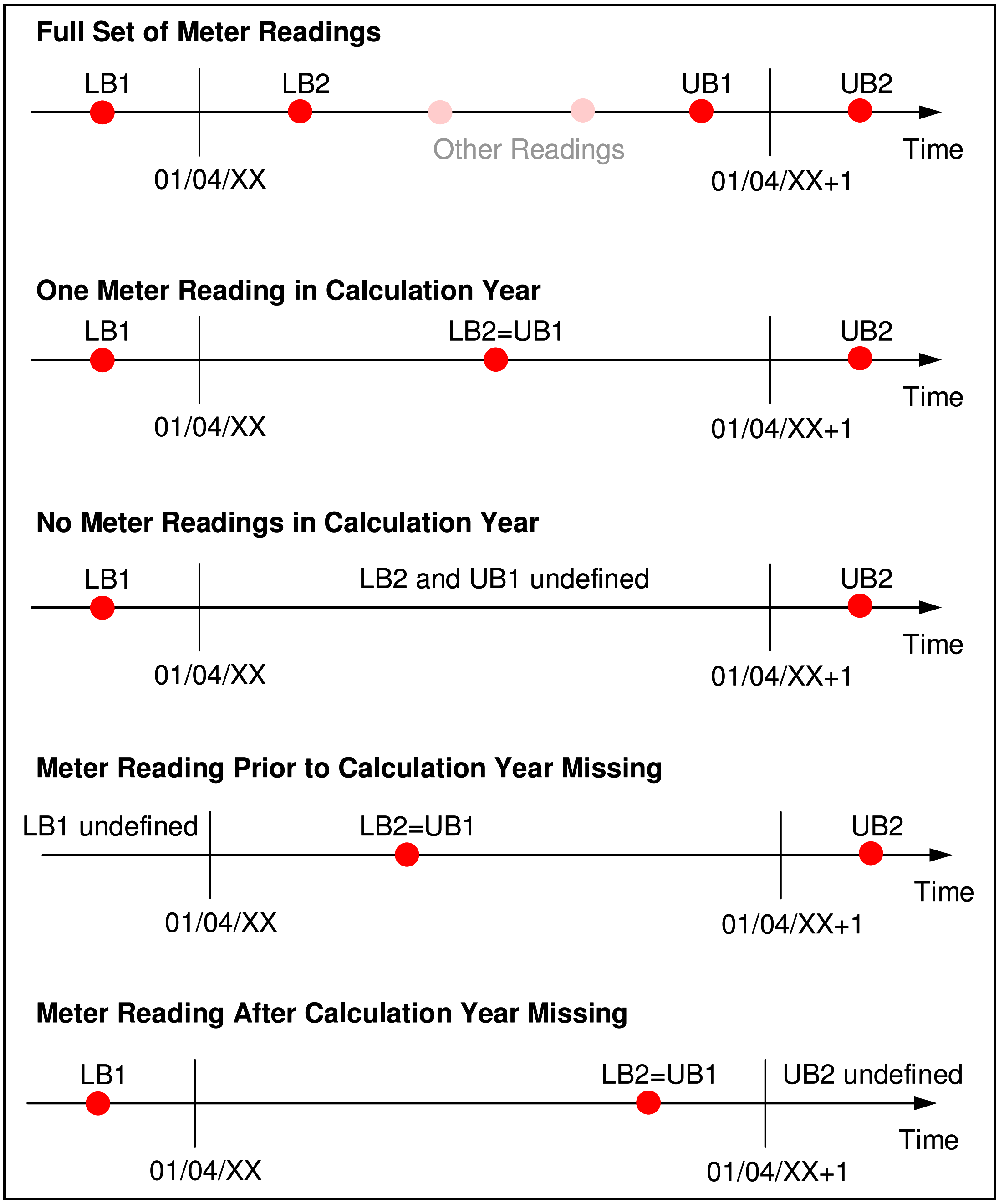
2. Find all meters that were active in a least part of year Y and for which meter reads were taken in and/or around that time period.

3. For each meter 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

   Note that for any given meter, only a subset of this full set of reads may be available. Depending on the nature of this subset, it may still be possible for the consumption algorithm to be carried out for this meter. The full set of meter read availability scenarios that (potentially) allow the formula year consumption for a meter to be calculated are shown in Figure 2 overleaf. If fewer meter reads than shown in any of the examples in this diagram are present, then no consumption can be calculated for the meter in question.

   In order for a meter to have the full set of meter reads, at least two reads within the formula year in question are required (i.e. LB2 and UB1). Depending on meter read frequency there may be more additional reads falling between LB2 and UB1, but these are not used in the analysis.

FIGURE 2: METER READ AVAILABILITY SCENARIOS



4. 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.

5. 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.

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

7. 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.

8. Calculate the volume consumed between the two chosen meter readings. Record whether the meter is metric or imperial and which EUC band it was allocated to when the end meter read was taken.

9. If the volume is negative the meter might have rolled over. Check the number of dials on 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.

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

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

    where ALP^{v} is the ALP divided by the relevant CV value (i.e. a 'volume' ALP rather than the usual energy ALP).
    N.B. if the site is recorded as starting or stopping part way through the year then the demand should not be scaled up to a full year.

11. Then convert this to energy using the appropriate Metric/Imperial conversion factor and a weighted average CV for the formula year, calculated as

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

12. Look up the first AQ estimate effective after the end of the formula year. If none exists use the latest value. Use that value to estimate consumption using the following formula:

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

13. If we have calculated a consumption value from the meter readings that is more than 5 times smaller or more than 5 times larger than the value based on the AQ then reject it.

14. Sum the consumptions for all meters calculated using meter reads, grouped by EUC band.

15. Count the total number of active meters in each EUC band as the number of MPRNs in the consumption file plus the number of MPRNs in the "no meter reads" file plus the number of MPRNs in CSEPs.

16. For each EUC, calculate the percentage of the total number of meters in that EUC for which consumption has been successfully calculated. Scale the calculated aggregate consumption in that band to cover the full population level as follows:

    *Metered EUC Consumption = Sample Consumption * (Total EUC Popn)/(Sample EUC Popn)*

17. Sum the allocations by EUC and formula year to give the total EUC load. Sum the EUC totals to give overall LDZ load.

18. For each EUC, subtract the metered EUC consumption from the total EUC consumption to give UG by EUC. Sum UG over each EUC to give total LDZ UG.

The key assumptions associated with this approach to calculating metered LDZ consumption by EUC are as follows:

- For each MPRN, the consumption for the time period between the meter reads used in the calculation is representative of the consumption over the full formula year, taking seasonal factors into account.
- The sample of meters that pass validation and are used in the analysis is representative of the population as a whole for each EUC.

The validation rules described above are designed to ensure that both of these assumptions are reasonable by including only meters with reliable data whilst at the same time creating samples that are large enough to be representative for each EUC.

2.2 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 Section 2.1 above, 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:

$$FPC = \sqrt{\frac{N-n}{N-1}} \qquad (2.4)$$

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 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):

$$95\%\ CI = \bar{x} \pm 1.96 \times S/\sqrt{n} \times \sqrt{N-n/N-1} \qquad (2.5)$$

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 the EUCs, (i.e. the strata of a stratified population) is obtained using the following formula:

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

Confidence Intervals calculated using this technique are given alongside the best estimate of Unidentified Gas in Section 2.3 below.

2.3 Results

Results from the analyses described in Sections 2.1 and 2.2 above are given here. The data supplied by Xoserve allowed estimates of UG to be made for the formula years 2009/10, 2010/11 and 2011/12. The number of meters in EA LDZ and the number for which consumptions could be estimated using meter read data are given in Table 1 below. In all cases, any meter that existed and was (potentially) passing gas at any point in the formula year in question is included in the population.

TABLE 1: POPULATION AND SAMPLE SIZES FOR CONSUMPTION ANALYSIS

Formula Year	Population Size	Sample Size	Sampling %
2009/10	1,874,737	1,580,055	84.3%
2010/11	1,893,209	1,597,240	84.4%
2011/12	1,907,835	1,302,069	68.3%

The sampling percentage for 2011/12 is lower that for the previous two years due to the fact that the current data only runs up to the end of March 2012, meaning that for this formula year only, the UB2 meter read is unavailable for all meters. Therefore, in this case a successful estimate of consumption can only be made where meter read UB1 is present and valid. As more data becomes available and the UB2 meter reads are populated, the sampling percentage for this year will increase. Data for all LDZs has been requested from Xoserve and this will include updated information for EA. The final calculations will be based on this new dataset and therefore in this version the sampling percentage for 2011/12 will be higher.

The estimates of UG given by these samples, including both the best estimate and the upper and lower confidence bounds around it, are given in Table 2 below. All figures are in GWh.

TABLE 2: UNIDENTIFIED GAS ESTIMATES USING CONSUMPTION ANALYSIS

	Allocation	Metered Consumption	UG (Year)	Best Estimate %	Low %	High %
2009/10	39,805	38,541	1,264	3.18%	2.80%	3.56%
2010/11	41,095	39,985	1,110	2.70%	2.40%	3.00%
2011/12	34,080	33,673	407	1.19%	0.73%	1.66%
Average				*2.36%*	*1.98%*	*2.74%*
				Best	Low	High
UG Estimate				**803**	673	933

Note that the Best Estimate figure in this table is obtained by taking an average of the UG percentage (i.e. UG expressed as a percentage of throughput) for the three years in question and applying this to the throughput figure for 2011/12. The full data and calculations used can be found in the spreadsheet "Unidentified Gas Estimate.xls" that has been made available on UKLink to accompany this document.

As described in Section 2.1 above, this estimate is obtained using a technique based on calculating the aggregate consumption for the sample of meters for which valid data was available, and then multiplying this up on and EUC-by-EUC basis to cover the full population.

An alternative method of substituting AQs for consumptions from meters without valid read data was considered, but this was rejected due to the inaccurate results it returns. The problem with this approach is that AQ is designed to reflect a typical consumption for the meter in question under seasonal normal conditions rather than the actual consumption in a specific year. Whilst a factor to convert between the two has been included in the code used by the AUGE for these calculations, this is necessarily a relatively crude approach when compared to using actual meter reads for the year in question. In addition, bias in the AQs also has an impact on the total consumption figure and can result in negative estimates of UG (as shown below). This issue results in the addition of a considerable amount of variation to the total UG estimate when (adjusted) AQs are used in place of consumptions for meters with missing or unusable meter read data.

Results obtained from this method (best estimate of UG only) are given in Table 3 below.

TABLE 3: UNIDENTIFIED GAS ESTIMATES USING AQ SUBSTITUTION

	Allocation	Metered Consumption	UG (Year)	Best Estimate %
2009/10	39,805	38,535	1,270	3.19%
2010/11	41,095	39,910	1,184	2.88%
2011/12	34,080	34,183	-103	-0.30%
Average				*1.92%*

Whilst the 3-year UG average that arises from this analysis is of the same order as the previous result (1.92% compared to 2.36%), it is the year-on-year variation that causes concern here. In particular, this technique gives negative UG for 2011/12 which is a physical impossibility given the values of the directly calculated components.

Therefore, based on this analysis, the AUGE recommends that the technique of scaling up sample demands to cover the full population, as detailed in Section 2.1 above, is used in preference to the substitution of AQs for meters outside the sample.

The chosen technique results in a best estimate of UG of 803GWh for EA LDZ, which is of a similar order to the value of 841GWh estimated for this LDZ in the 2011 AUGS for 2012/13.

The Confidence Interval (CI) for this estimate runs from 673GWh to 933GWh, meaning that based on the samples used for this particular analysis, we can be 95% sure that the true UG total for this LDZ lies between these two figures. The best estimate of total UG (803GWh) lies at the centre of this interval and would be the figure carried through to the remained of the UG calculations.

2.4 Comparison with 2011 Method

As stated in Section 2.3 above, the new method of estimating total UG based on meter reads gives a result for EA LDZ which is of the same order as that from the previous method: 803GWh as opposed to 841GWh.

In addition to the best estimate, the other key issue regarding this estimate is the relative size of the Confidence Intervals around each one. The method that produces the narrowest Confidence Interval can be regarded as providing a more reliable estimate of total UG.

The Confidence Interval for the consumption method is calculated in Section 2.3 and runs from 673GWh to 933GWh.

It is not simple to calculate Confidence Interval for the old method directly because this would require algorithm bias calculations to be carried out on a meter-by-meter 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 approach relies on the fact that the old UG (RbD-based) method and the new (meter read based) method both use consumptions based on meter reads in one form or another.

Old Method
Meters Read → AQs → Algorithm Bias (LSP and SSP)

New Method
Meters Read → Calculated consumptions

As stated in the 2011 AUGS for 2012/13 [Ref 1], the old method uses approximately 76% of the LSP population and 85% of the SSP population in the algorithm bias calculations.

Given that we cannot easily calculate algorithm bias on the meter-by-meter 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 old and the new 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 4: COMPARISON OF SAMPLE SIZE AND FINITE POPULATION CORRECTION (FPC)

Formula Year	Consumption Method		Old Method	
	Sampling %	FPC	Sampling %	FPC
2009/10	84.3%	0.40	84.9%	0.39
2010/11	84.4%	0.40	86.4%	0.37
2011/12	68.3%	0.56	88.3%	0.34

This table shows that for 2009/10 and 2010/11, the Finite Population Corrections from both methods are similar and will lead to CIs of comparable width. For 2011/12 the sampling proportion for the consumption method drops due to current data availability, leading to the FPC being larger in this case. When data becomes available the sample proportion will rise, however, which is likely to lead to very similar FPCs again in this year.

Therefore, when full data is available for 2011/12, the two methods will produce CIs of similar width around their best estimate of total UG.

It is also important to bear in mind the difference between variability in the UG estimate and sources of error in the UG estimate. The above analysis considers only the level of variability. With regard to sources of error, the new method minimises these by working with directly-calculated consumption values for all market sectors wherever possible. By contrast, the previous method makes many more assumptions and hence contains more potential error sources. These error sources include:

1. Inconsistencies in RbD due to retrospective corrections and close-out period end.
2. Error in the “Best AQ” value used to calculate algorithm bias.
3. Inherent assumption that SSP-assigned UG is negligible.

Each of these elements can and will introduce systematic errors into the old method without necessarily increasing the variability of the estimate.

Based on the above analysis, the AUGE believes that the consumption method is more accurate by design. Given the fact that when full data is available the variability in its estimate is no greater than that of the old method, the AUGE recommends that the consumption method is used to calculate total UG for 2013/14 and for future years.

3 Split of Theft by Market Sector

In all UG calculations carried out by the AUGE to date, 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 of each site during the period of theft where valid meter reads are present, which could lead to a more accurate assignment of each site to market sector.

The sections below set out alternative methods for assigning market sector and compare these to the original method. Note that whilst the total UG estimate analysis described in Section 2 above has only been carried out for EA LDZ due to the vast amount data required, full theft data including consumptions for all LDZs that have theft affected sites is available. Therefore this analysis has been carried out for all LDZs and results across all LDZs are presented.

3.1 Summary of Original Method: 2011 AUGS for 2012/13

Key Features:

- Apportions detected theft to calendar year of occurrence.
- Applies flat profile to assign theft that crosses year-to-year boundaries into different years (i.e. no seasonal factor is included).
- Uses pre-theft AQ to set market sector.
- Market sector is fixed for the full period of multi-year thefts.

Issues:

- This method creates a disincentive for all shippers to detect theft (in both LSP and SSP market sectors). This is due to the fact that any detected theft increases the percentage split for that market sector and hence will contribute to the amount of UG assigned to that sector in the AUGE's calculations for subsequent years.
- Examination of meter readings shows that pre-theft AQ can be affected by theft if the estimate of theft start date is incorrect.
- Pre-theft AQ may not represent the demand level at the time of theft if there is a very long period of theft or if the occupancy/usage has changed.
- The method carries an inherent assumption that the split of unknown theft between market sectors is the same as the detected theft split.
- The calculation relies on the accuracy of the theft estimate and the estimate of the period of theft.
- The potential exists for external influence on the theft split if mixed shippers focus on detecting theft in one sector over another.
- The method does not account for thefts in the year of occurrence exceeding the LSP threshold (i.e. the AQ is used even if the theft in the year is greater than 73,200kWh).

3.2 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. This is a separate issue from temporary and permanent UG where temporary UG can be picked up via retrospective corrections to RbD over time. The aim of the theft split is to estimate what proportion of total theft occurs in each market sector. Any credits will need to be dealt with separately as part of the direct UG calculations.

Network code 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 part 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 3] and the Gas Act 1986 [Ref 4].

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.

This rule has been applied to the dataset used by the AUGE to calculate the split of theft by market sectors. Approximately 20 sites were removed from the dataset for this reason, of which 5 were LSPs. These included one very large site whose theft was discovered in 2011 and stole over 17GWh of gas over a 10-year period, but was Unregistered for the entire duration of the theft.

3.3 Theft of Gas in Year Overriding AQ

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 the estimate of theft that occurred in the year exceeds the 73,200 kWh threshold. The number of these and the associated amount of theft is shown in Table 5 below.

Closer inspection of these instances confirms that the sector classification should be LSP rather than SSP.

TABLE 5: NUMBER OF INSTANCES WHERE THEFT IN FORMULA YEAR EXCEEDS 73,200KWh AND OVERRIDES AQ (TABLE REF 16082012_1)

Formula Year	Number of Instances	Theft (GWh)
2007	23	2800.1
2008	28	3409.7
2009	23	2858.4
2010	13	1480.7

When applied to the original method, 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.

The AUGE has therefore adjusted the original method to incorporate 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.

The alternative methods described in the following sections also incorporate this approach so that a like for like comparison can be made in each case.

3.4 Alternative Methods

This section considers alternative methods to the original method summarised above. Following clarifications of the theft split at the May 2012 UNCC meeting, the first new method considered uses meter read data and the theft estimate to calculate both metered and unmetered consumption at the site in question, which in turn is used to derive an AQ for market sector classification. A variation of this method is also considered where the theft split is fixed at a point in time in order to avoid potential external influence, and a further version of this is also considered where this fixed theft split is tied to throughput and can therefore vary over time.

A further method considered is the use of throughput to split theft.

3.4.1 Alternative Method 1 – Theft Split using Metered and Unmetered Consumption

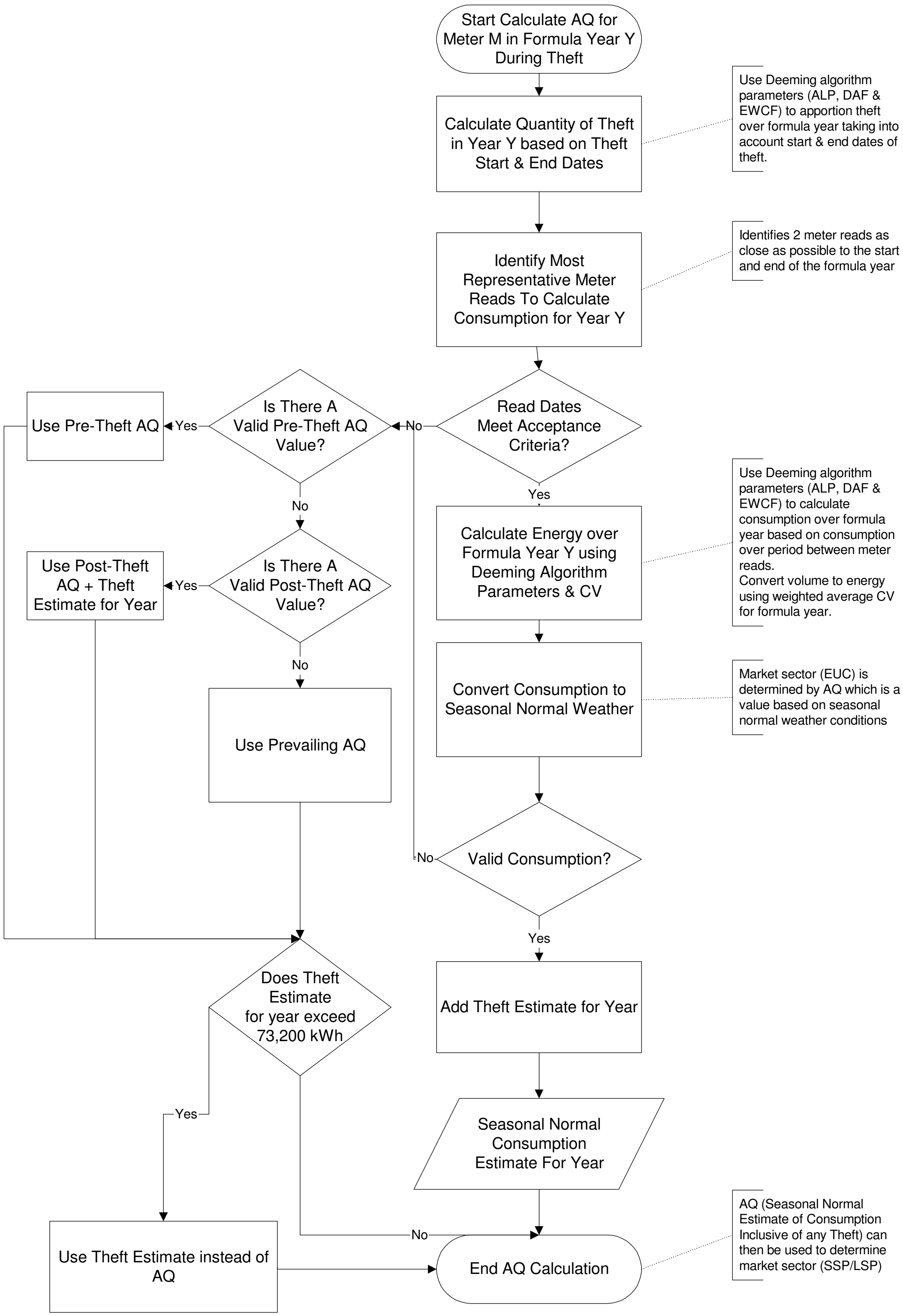
Key Features:

- Apportions theft to formula year of occurrence.
- Splits theft that crosses year-to-year boundaries 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 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 analysis 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.
- 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.

If theft apportioned to a given year is greater than 73,200 kWh this overrides the AQ in cases where the consumption calculation fails.

Figure 3 shows how the theft split method works using metered and unmetered consumption where possible, with AQ used as the backup option when the primary method fails.

FIGURE 3: THEFT SPLIT METHOD FLOW CHART



Issues:

- This method still creates a disincentive for all shippers to detect theft in both the LSP and SSP market sectors.
- 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, as shown in Table 6 below. This can be due to constant meter reads or missing meter reads during the period of theft. Such problems are to be expected given that the calculations are all for theft-affected sites. In all, approximately 50% of all consumption calculations fail.
- When using AQs (in the case of consumption failures) from periods before and after theft 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, but not to the point where conclusions can be drawn about any potential impact on the market sector split for theft.

TABLE 6: METERED CONSUMPTION CALCULATION FAILURES FOR THEFT AFFECTED SITES

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

3.4.2 Alternative Method 2 – Fixed Split using Metered and Unmetered Consumption

This method is based on Alternative Method 1, but in order to protect the theft split from potential external influences it is frozen at a point in time and hence only uses theft detections up to the end of December 2011. This prevents any effect on the figures from Shippers concentrating theft detection activities on one particular market sector over another.

The remaining issues are the same as Alternative Method 1. In addition, over time the theft split may be challenged as being out of date and require re-analysis.

3.4.3 Alternative Method 3 – Fixed Theft Split Varied by Throughput

This method is based on Alternative Method 2, but the theft split initially obtained is then expressed as a proportion of throughput for each market sector. In subsequent years the theft split is calculated based on this proportion rather than remaining fixed and hence will vary in line with throughput.

The remaining issues are the same as Alternative Method 2, and there would still be some potential for the theft split being out of date over time if initiatives to improve theft detection and prevent theft change the base levels of theft within the population.

3.4.4 Alternative Method 4 - Throughput

Key Features:

- 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 which can 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, 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. The validity of this assumption can be checked by comparing results (in terms of the market sector split of theft) when calculated using Alternative Methods 1-3 and the throughput method. If results are similar, it can be concluded that this assumption is valid.

3.5 Theft of Gas Results

In this section, the market sector splits of theft as calculated using the alternative techniques described in Section 3.4 are given. Note that only two alternatives are shown due to the fact that at this point in the calculation cycle, Alternative Methods 1, 2 and 3 are equivalent because the time periods used by methods 1 and 2 are currently the same, and the 3rd method will only vary in future years as throughput changes.

The percentage split figures given here are calculated across all LDZs and would be applied to the Balancing Factor for each LDZ, which is overwhelmingly composed of unknown theft. It is only possible to present the market sector split at this stage rather than any estimates of UG from unknown theft itself because full UG calculations (and hence calculation of the Balancing Factor) cannot be carried out at this point.

As noted above, the consumption calculation failure rate for theft affected sites is high with approximately a 50% failure rate. Therefore, for the remaining records an alternative based on AQ was used, as described in Section 3.4.1 above.

In order to make a like for like comparison of the methods, two different subsets of the full population of theft records were created. The first contained those records where the consumption calculation was successful, and results from this subset provide a direct comparison between the old and new methods on a set of records where the new method was always applied. The second subset contains those records where the consumption calculation failed. In this case the comparison is between the backup method (which is a development of the original method and hence much closer to it) and the original method, again on the same set of records.

A comparison of the original method and Alternative Method 1 (which is also equivalent to Alternative Methods 2 and 3 at this stage) is shown in Table 7 below.

TABLE 7: COMPARISON OF ORIGINAL METHOD AND ALTERNATIVE METHOD 1 THEFT LSP % (TABLE REF 16082012_3)

	2007		2008		2009		2010	
	Original Method	Alternative Method 1	Original Method	Alternative Method 1	Original Method	Alternative Method 1	Original Method	Alternative Method 1
Consumption Calculation Successful	11.5%	13.1%	15.7%	13.0%	19.4%	15.5%	15.8%	8.9%
Consumption Calculation Failed	20.3%	26.7%	29.6%	31.0%	25.6%	24.1%	28.6%	26.4%
Combined	16.4%	19.3%	22.5%	21.4%	23.0%	22.6%	24.8%	20.6%

Where the consumption calculation was successful the two methods provide results with a similar level of variation, with the annual LSP percentage running from 11.5% to 19.4% for the original method (a range of 7.9%) and from 8.9% to 15.5% for the new method (a range of 6.6%). The new method gives a lower average LSP percentage across all years at 12.6% as opposed to 15.6%.

For records where the consumption calculation failed the LSP percentage from both methods is significantly higher: an average of 26.0% for the old method and 27.1% for the new method. It is not unexpected for the two sub-populations to have a different SSP/LSP split because they have been divided on the basis of their theft characteristics: where the consumption calculation is successful, the level of theft is such that meter reads are still made and are considered valid. Where the calculation is unsuccessful this indicates that either meter readings could not be obtained or were constant or invalid. In this latter case it is likely that the level of theft is much higher and could be as much as the entire consumption on the site. It is interesting to note, however, that for these sites that are potentially stealing a higher proportion of their gas, it is the LSP percentage rather than the SSP percentage that is higher.

When these two subsets are combined to form the full theft-affected population, the overall LSP theft percentage is extremely consistent from year to year for Alternative Method 1. The LSP percentage in this case runs from 19.3% to 22.6%, a range of 3.3%. By contrast the overall LSP percentage from the original method runs from 16.4% to 24.8%, a range of 8.4%. If, however, we omit the oldest year (2007), the ranges produced by each method are similar at 2.0% and 2.3% respectively.

The full set of year-on-year results from Alternative Method 1 (for the full set of theft-affected sites) is shown in Table 8 below, where the consistency of the yearly values can be seen.

TABLE 8: ALTERNATIVE METHOD 1 LSP PERCENTAGE (TABLE REF 16082012_2)

Year	LSP	SSP	Total	LSP%
2007	4112289	17171071	21283360	19.3%
2008	5913913	21768639	27682553	21.4%
2009	5069834	17373857	22443691	22.6%
2010	3250190	12552272	15802462	20.6%
2008-2010				**21.5%**

Note that at this point in time it has not been possible to incorporate data concerning customer changes into the analysis. This is due to the fact that information requested from the industry is still awaited from some of the Shippers.

The final market sector split method examined in this document is Alternative Method 4, which is a simple split of theft by market sector throughput. The results of splitting theft by throughput are given in Table 9 below.

TABLE 9: MARKET SECTOR SPLIT BASED ON THROUGHPUT IN TWh, ALL LDZs (TABLE REF 16082012_4)

Sector	2007	2008	2009	2010	2011	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%**

Note that due to the fact that the LSP percentage is consistently falling rather than varying randomly, in this case the projected value for 2013 based on the observed trend is taken as the final figure rather than the average across years.

This is a relatively simple method that produces a consistent set of LSP percentage values year on year and the strengths and weaknesses of this approach are discussed in Section 3.6 below.

3.6 Discussion of Results

All of the alternative methods described in the above sections provide a more detailed analysis of theft than the original method. Each has its own strengths and weaknesses, however, and these are discussed here.

The key advantage of Alternative Method 1 is that it uses consumption data to calculate the AQ of the theft-affected sites more accurately than the current method, and also deals with multi-year thefts more accurately via the use of ALPs to assign theft to the correct time period. It still retains some key drawbacks of the original technique, however. Whenever calculations are based on detected theft figures, whatever method is used, this will always create a disincentive for Shippers to detect theft going forward. In addition, the use of detected theft data provides mixed Shippers with the opportunity to influence the market sector split by concentrating detection efforts in one market sector or another, a situation which needs to be avoided.

In order to mitigate this drawback, Alternative Method 2 limits the period of theft detections used to the end of 2011. This fixes the theft split figure at a point in time, preventing Shippers from exerting any influence over the market sector split via theft detection policy. It also reduces the disincentive to prevent and detect theft for all Shippers. With this method, both prevention and detection of theft reduces the (permanent) Unidentified Gas total and hence results in savings to the Shippers, without there being any associated effect on market sector split (which could cause costs to the Shipper).

The drawback in this case is that if the relative size of the market sectors or the prevalence of theft within them were to change, the balance of theft across market sectors could also change. This could cause the split to become out of date and require revision.

Alternative Method 3 attempts to address this issue by linking theft to throughput. Theft is calculated in the same way as in Alternative Methods 1 and 2 initially, but is then expressed as a proportion of throughput. This proportion is fixed and applied to the throughput from future years to estimate theft for these. This amendment to the technique guards against theft changes due to relative market sector size, but will not account for any fundamental changes in the pattern of theft within market sectors should they occur.

In addition to these general points, Alternative Methods 1-3 share a number of additional issues as follows:

- 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
- 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 Alternative Method 4, 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 validity of this assumption can be investigated using data from Tables 8 and 9 above. Table 8 shows the LSP theft percentage as calculated using Alternative Methods 1-3 (which all produce the same estimate at the present point in time). This method returns an LSP percentage of 21.5%, a figure based on a full analysis of detected thefts with the market sector of each record based on metered consumptions and theft estimates where available.

Table 9 gives the breakdown based on the simple throughput method, which returns an LSP percentage of 24.8%. These two figures are similar in magnitude, and a statistical 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 sizes, i.e. the 3-year period used for the calculation of the results using each method). The statistical test used was a 1-sample t-test based on Alternative Method 1 results tested against a hypothesised mean of 23.7%, the value from Alternative Method 4. This approach is necessary due to the fact that data from Alternative Method 4 does not vary randomly.

These results show that the two methods return market sector splits with no statistically significant difference between them, which in turn indicates that the underlying assumption of Alternative Method 4 is reasonable.

As stated in the list above, Alternative Methods 1-3 are highly sensitive to the quality of their 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 (Alternative Method 4), 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.

3.7 Industry Participant Questions

As part of the consultation of this analysis the AUGE poses the following questions to the industry participants for their views on the results presented in this report. These will be incorporated into the next draft of the full AUGS, which will set out the proposed methods for calculating total UG and splitting theft between market sectors, and also incorporate updates on the other elements of UG.

1) To what extent do you believe that the estimate of UG using consumption data provides a better estimate of UG than the previous AQ bias method?
2) Regarding theft split, to what extent do you think the original method and Alternative Methods 1,2 and 3 act as a disincentive to you in terms of theft detection/prevention?
3) Of the theft split methods presented, which one do you think provides the most robust and reliable method of splitting theft into the market sectors given the issues described above? What are your reasons for drawing this conclusion?

4) What are your remaining key concerns regarding the methods presented?

4 Conclusions

The following conclusions are drawn:

Consumption Analysis

- The method of calculating total Unidentified Gas using meter read data for both the SSP and LSP market sectors is more accurate than the approximate method employed in the 2011 AUGS for 2012/13. The AUGE therefore concludes that this method should be used in the 2012 AUGS for 2013/14 and for future years.
- When applied to EA LDZ, the new method results in a total Unidentified Gas estimate of 803GWh. This is of the same order as the figure for this LDZ in the 2011 AUGS for 2012/13, which was 841GWh.
- The Confidence Interval associated with this estimate is from 673GWh to 933GWh. Based on the data used in this analysis, we can therefore be 95% sure that the true Unidentified Gas total for EA LDZ lies between these two figures.
- It is estimated that the Confidence Interval for the new method is of a very similar width to that associated with the old method.

Theft Analysis

- Meter read data can also be used in the theft analysis for allocating sites to market sectors more accurately. This produces more accurate results than those presented in the 2011 AUGS for 2012/13, but a number of issues remain regarding this analysis.
- When using meter read data, the average estimate of the proportion of theft that occurs in the LSP market sector is 21.5%.
- The use of detected theft figures in this analysis means that all Shippers are disincentivised to find theft in their own market sectors, and that mixed Shippers have the opportunity to influence the figures by concentrating detection efforts in one market sector over another. Both of these situations are undesirable.
- The failure rate for consumption calculations for theft affected sites is approximately 50%.
- Meter read data has also identified issues with use of pre-theft and post theft AQ as AQs can be affected in different ways and this is not consistent from case to case.
- If theft is split by market sector using throughput instead, this results in an LSP theft percentage of 23.3%.
- The AUGE concludes that the throughput method should be used in the 2012 AUGS for 2013/14 and for future years because it avoids many of the issues identified with the method based on detected theft and results in a more robust methodology.

5 References

[1] AUGS 2012/13 Final (Version 4), December 2011, GL Noble Denton

[2] AUGS 2013/14 First Draft. May 2012, GL Noble Denton

[3] Uniform Network Code – Transportation Principal Document, Section N - Shrinkage

[4] The Gas Act 1986, Schedule 2B “The Gas Code”

Appendix A Theft Statistics

A.1 Additional Metrics

From the history of theft data provided by Xoserve, the number of theft records where the theft start date is greater or equal to 01/04/2007 and less than or equal to 31/03/2011 for registered sites is 3147.

Having assigned each theft to the year in which it occurred, the number of thefts by sector by formula year is given in table 10 below.

TABLE 10: NUMBER OF THEFTS OCCURRING IN FORMULA YEAR BY SECTOR (TABLE REF 16082012_7)

	Number of thefts occurring in formula year by sector	
Year	**LSP**	**SSP**
2007	94	1906
2008	112	2483
2009	116	2658
2010	84	1498

As noted in the report, consumption could not be calculated for approximately 50% of the theft affected sites. An AQ substitution method was used and the prevalence of each type is shown in Table 11.

TABLE 11: NUMBER AND TYPE OF AQ SUBSITUTION USED FOR CONSUMPTION FAILED RECORDS (TABLE REF 16082012_9)

AQ used	Count
Pre Theft	5127
Post Theft	678
Current AQ	68

A.2 Examples of sector movements Original vs. Alternative methods

During the analysis, a subset of sites were classified differently by each method. As part of the validation and checking of the reasons for this, a closer examination of the data was carried out on a site by site basis. Examples are given in this section to highlight some of the differences identified. This included examination of the meter reads, AQs, customer changes where data was available, theft estimates and periods of theft.

In the main the differences identified between methods can be justified, however, there are some examples where the real sector cannot clearly be determined without further information (e.g. specific customer and site details which are not available to us).

MPR dummy reference 7344

Start Date	End Date	Theft amount
26/09/2010	05/04/2011	125,401

AQ Data

Start Date	AQ
01/10/2005	17302
01/10/2007	25925
01/10/2010	14370
01/10/2011	14370

For this site, the consumption calculation failed so pre-theft AQ used – the prevailing AQ prior to the theft start period was 25,925 kWh.

In the original method the theft was split by calendar year so approximately 60,000 kWh was assigned to 2010 and a similar amount to 2011. In the revised method using formula year, the bulk of the theft sits in 2010 and exceeds 73,200 kWh hence this site is set to LSP by virtue of the amount of theft.

MPR dummy reference 7749

Start Date	End Date	Theft amount
20/08/2009	28/09/2011	76,816

AQ Data

Start Date	AQ
01/10/2005	10,990
01/10/2006	20,909
01/10/2008	94,997
01/10/2009	82,774
01/10/2010	53,063

The Original method used pre-theft AQ of 94,997 kWh, hence was categorised as LSP. The new method calculated consumption from meter reads in 2010 as 13,034 kWh plus 39,065 kWh of theft, hence SSP.

MPR dummy reference 7293

Start Date	End Date	Theft amount
10/09/2007	04/04/2011	154,559

AQ Data

Start Date	AQ
01/10/2005	122,960
24/11/2005	89,938
01/10/2009	77,020
01/10/2010	21,554

The Original method used pre-theft AQ of 89,938 kWh, hence was categorised as LSP. The new method calculated consumption from meter reads in 2010 as 13,755 kWh plus 43,111 kWh of theft, hence SSP.

MPR dummy reference 6440

Start Date	End Date	Theft amount
18/04/2007	03/05/2011	31,291

AQ Data

Start Date	AQ
01/10/2004	1
01/10/2005	99,112
01/10/2007	742
01/10/2008	1
01/10/2009	200
01/10/2010	1
01/10/2011	1

The Original method used pre-theft AQ of 99,112 kWh, hence was categorised as LSP. The new method calculated consumption from meter reads in 2010 as 0 plus 8213 kWh of theft, hence SSP.

MPR dummy reference 920

Start Date	End Date	Theft amount
15/09/2008	26/10/2010	294,059

AQ Data

Start Date	AQ
01/10/2004	22,286
01/10/2005	48,204
01/10/2006	50,628
01/10/2007	34,353
01/10/2008	685
01/10/2009	126
01/10/2010	338
01/10/2011	4,026

The Original method split the theft by calendar year as 40,810 kWh, 139,211 kWh and 113,657 kWh (for 2008, 2009 and 2010 respectively). The new method split the theft by formula year as 106,845 kWh, 140,630 kWh and 46,582 kWh. Hence the site is assigned to SSP in 2010.

MPR dummy reference 3557

Start Date	End Date	Theft amount
26/05/2009	22/07/2010	102,887

AQ Data

Start Date	AQ
01/10/2004	21,993
01/10/2005	26,890
01/10/2006	29,709
01/10/2007	39,680
01/10/2008	6,110
01/10/2009	1,121
01/10/2010	1,003

The Original method allocated 53,394 kWh of theft to calendar year 2009. The new method allocated 85,102 kWh to formula year 2009 so the site is assigned to LSP.

MPR dummy reference 7347

Start Date	End Date	Theft amount
01/05/2009	07/05/2010	84,800

AQ Data

Start Date	AQ
01/10/2005	1
01/10/2008	1
01/10/2009	1
01/10/2010	1

The Original method allocated 55,771 kWh of theft to calendar year 2009. The new method allocated 76,583 kWh to formula year 2009 so the site is assigned to LSP.

MPR dummy reference 3625

Start Date	End Date	Theft amount
07/04/2004	17/04/2008	153,747

AQ Data

Start Date	AQ
01/10/2004	13,512
01/10/2005	37,880
01/10/2008	66,235
01/10/2009	23,218
01/10/2010	30,863
01/10/2011	21,580

Meter reads

Read Date	Read Value
22/06/2006	1,701
28/06/2007	6,925
17/04/2008	0
17/12/2008	2,798
04/10/2009	0
17/11/2009	2,568
16/07/2010	4,198
21/10/2010	4,624
29/01/2011	5,396
28/01/2012	7,317

The consumption calculation fails for 2007. There is no pre-theft AQ available so we use the first post-theft AQ of 66,235 kWh plus 35,541 kWh making the site LSP. The Original method did not take account of the theft bringing consumption over the 73,200 kWh threshold and so had categorised the site as SSP.

MPR dummy reference 1200

Start Date	End Date	Theft amount
31/08/2004	21/07/2009	206,715

AQ Data

Start Date	AQ
01/10/2004	191,454
01/10/2005	111,603
01/10/2007	22,090
01/10/2010	21,030

The consumption calculation fails and there is no pre-theft AQ so we use the first available post-theft AQ which is 21,030 kWh. The theft is split across 2006-2009 as 39,295 kWh, 41,393 kWh, 43,826 kWh and 7,948 kWh so the site is always categorised as SSP. This is suspicious given the previous AQs > 100,000.