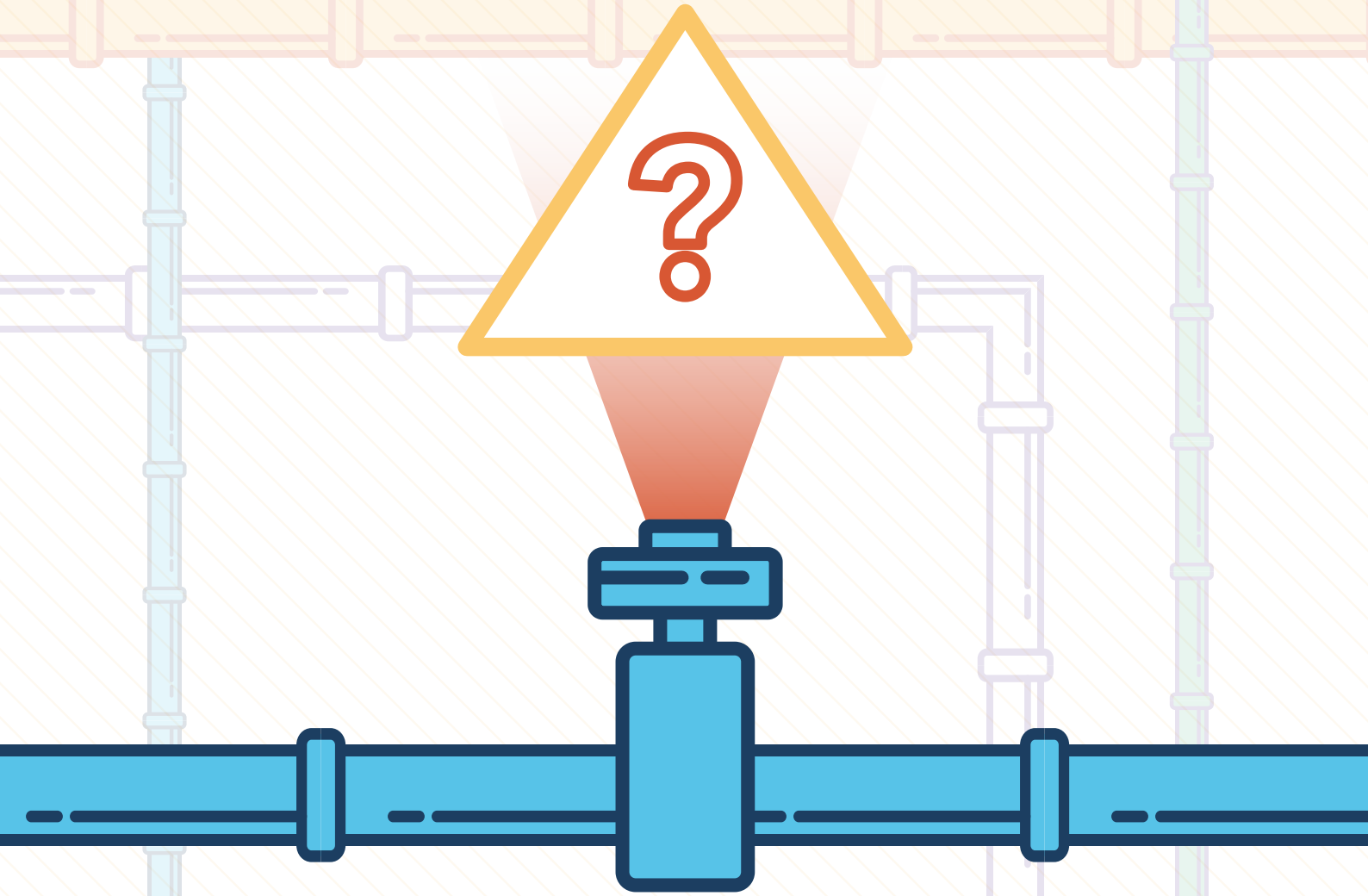


Unidentified Gas

Xoserve Task Force Investigation Guide

February 2019

xoserve

Introduction to Unidentified Gas

Xoserve successfully implemented Project Nexus in June 2017 which introduced a new suite of systems and industry driven process changes. The changes introduced the concept of Unidentified Gas – UIG – but gas shippers have experienced higher UIG base levels and daily volatility than they expected. This is affecting their ability to predict demand and is increasing the cash flow management risk for some of their businesses because UIG is reconciled (corrected) over an extended and unknown future period.

Ofgem approved Uniform Network Code (UNC) Modification 0658 in July 2018, requesting that Xoserve lead a more centralised and focussed approach to the resolution of UIG. Our role as the Central Data Service Provider to the industry means we are uniquely placed to carry out unbiased investigations into the root causes of UIG and to make recommendations to reduce the UIG risk to the whole market.

We rapidly set up a team of subject matter experts from our Operations and Data Office teams and recruited Cambridge Consultants as our external analytics partner, combining their world-class data science and analysis expertise with our in-house knowledge.

We developed a detailed plan for our fortnightly analytics sprints and activities to help us better quantify the impact of potential drivers of UIG. Our existing understanding of UIG root causes guided the plan, and the end of sprint reviews enabled the flexibility to reprioritise as we gained deeper insight into the underlying issues.

We maintain a UIG investigation tracker detailing our lines of investigation, and at the end of each sprint we published executive summaries communicating our findings, analysis progress and next steps. Over six sprints, we completed 23 lines of investigation and identified 13 issues as contributors to UIG. You can read about the UIG Task Force's finding and recommendations to date on the following pages, and we have published the detailed findings, a range of options and Xoserve's recommended options to resolve each issue on our **website**.

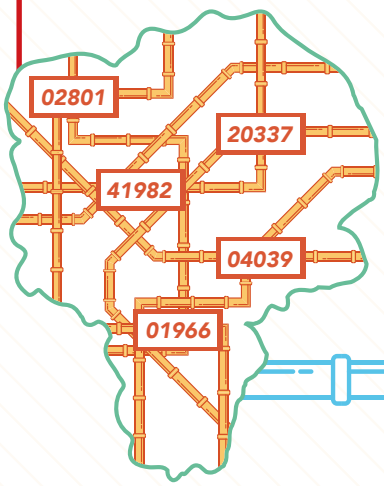
"We have completed 23 lines of investigation and identified 13 issues that contribute to UIG"

Xoserve and the Joint Office hosted a workshop in January 2019 to take the industry through our progress to date and agree the next steps. We made recommendations to the industry that we believe will reduce UIG. Some of our recommendations require the support of the wider industry to deliver. Our customers may not choose to take them forward and can raise their own recommendations.

We have a number of initiatives planned to realise the benefits of the Task Force's findings. We are developing our customer engagement strategy, concentrating on the best way to communicate key areas of focus to our customers and how we can most effectively partner with them to drive forward solution options that we think will reduce UIG. We will support our customers in drafting and raising UNC Modifications and Change Proposals, and we will further develop options to address complex challenges in cooperation with industry forums. Our investigations into the root causes of UIG are ongoing in collaboration with Cambridge Consultants, and we will continue providing regular updates on the progress of our investigations and the impact on UIG as the Industry and Xoserve take action.

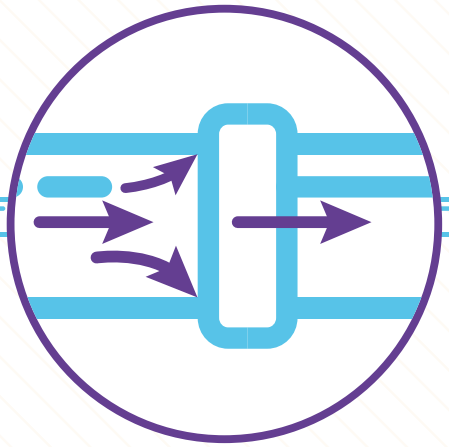
A look beneath the surface at what happens when **gas goes missing**

The majority of gas consumed in Great Britain is ***metered*** and ***registered***



The gas is taken from the Local Distribution Zone (LDZ) System, where it is attributed to an individual Supply Meter Point...

... Or accounted for as acceptable loss through shrinkage



Gas usage by the Transporters for activities such as running the equipment within the LDZ is an example of shrinkage

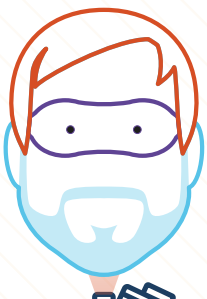
However,

some gas is lost from the system, or not registered...

This is due to:

The gas which cannot be accounted for is referred to as

THEFT



CONSUMPTION

by unregistered supply points

UNKNOWN REASONS

UNIDENTIFIED GAS

Contents *page 1 of 2 (click to navigate)*

Contents *page 2 of 2 (click to navigate)*

Investigation Area 12: Volume Conversion Factors

Gas is metered by volume, but billed in energy. Converting volume to energy uses a standard formula, and the conversion factor is part of that formula. The Standard Conversion Factor is defined in legislation and assumes that gas is metered 12.2°C and at an altitude of 66m above sea level. The standard value is set at 1.02264. If sites do not have an appropriate conversion factor the gas will be under or over metered and will contribute to UIG. Once the reads have been used to calculate an Annual Quantity (AQ), the daily usage estimate for the sites will also be affected.

Item 12.1: Use of standard conversion factors for NDM sites greater than 732,000 kWh AQ

All sites of this size should have a specific conversion factor based on the actual altitude, temperature and pressure at the meter rather than the industry standard value. There are **currently around 5,000 of 26,000 (around 18%) eligible sites without a site-specific conversion factor**, but this is a relatively small section of the market at approximately 1% national of AQ.

Comparing the standard conversion factor to the average of the specific conversion factor in each LDZ suggests that **the energy for the affected sites is 7.4% lower than it should be**, which we think is contributing 0.1% of total throughput to UIG.

We recommended that we monitor for sites that have standard conversion factors but are above the AQ threshold and notify relevant Shippers, highlighting the individual sites, and support shipper action to update the factors. To make this easier for our customers, we think that the conversion factor should be amendable via the Supply Point update process. We also think that introducing incentives to update the conversion factor when required would improve the position, and that automatically applying the last non-standard conversion factor (if available) to the site when the AQ increases above 732,000 kWh would reduce the UIG risk to the market.

Item 12.2: Use of a single standard conversion factor for all NDM sites with AQ less than 732,000 kWh regardless of location/geography

We analysed the impact of using actual LDZ temperatures instead of the standard 12.2 degrees to calculate daily, LDZ specific Conversion Factors. We adjusted energy usage over a one year period to see what could happen to energy levels. The results suggest that the annual effect is non-zero: the summer over-recording of actual energy does not fully offset the winter under-recording of actual energy. Between 01/06/2017 and 31/05/2018, this could have contributed up to 0.4% of national LDZ throughput to UIG, would have reduced UIG by up to 3% on peak days, and increased it by up to 4% on the warmest days. This year was colder than average and so the energy differences will vary depending on the actual weather.

"Using a Single Standard Conversion Factor could contribute up to 0.4% of throughput to UIG"

To reduce the impact, we recommended that we use the actual LDZ temperatures instead of the standard temperature to convert the metered consumptions used to develop the daily estimation profiles. We think that this would reduce the UIG seasonal variability without requiring changes to legislation.

Item 12.3: Use of a non-standard conversion factor for all NDM sites with AQ less than 732,000 Kwh

All sites of this size should have the industry standard conversion factor but around 10,000 of these sites have a nonstandard factor. These sites have a total AQ of 2.8bn kWh – around 5% of the total market – and their average AQ is around 270,000 kWh, suggesting that many sites were previously eligible for a site specific conversion factor and have not yet had an update back to the standard factor.

Comparing the standard conversion factor to the specific conversion factor suggests that the energy for the affected sites is 3.8% higher than it should be which **we think is reducing UIG by 0.02% of total throughput**.

We recommended that we monitor for sites that have nonstandard conversion factors and are below the AQ threshold, and notify relevant Shippers, highlighting the individual sites, and support shipper action to update the factors. To make this easier for our customers, we think that the conversion factor should be amendable via the Supply Point update process. This issue could also be monitored by the Performance Assurance Committee to give the wider industry oversight and visibility of progress.

We also think that introducing incentives to update the conversion factor when required would improve the position, and that automatically applying the standard conversion factor if the site AQ decreases below 732,000 kWh would reduce the UIG risk to the market. We also suggested that the standard conversion factor becomes a lookup value in our system rather than being held individually for each site. This would reduce complexity and help facilitate any changes to the standard conversion factor that may arise from item 12.2.

Investigation Area 3.2: Inaccurate or out of date Annual Quantity

We hold an Annual Quantity (AQ) for each live Meter Point. The AQ is the energy a site would use over a year under 'normal' weather conditions. The AQ level determines the usage profile we use to estimate the daily usage for the site, and how much the calculation varies the daily energy estimate as the actual weather differs from 'normal' weather. AQ is a key input in to the formula that calculates the daily energy for market sectors where we do not have daily readings available.

If the AQ is inaccurate then the estimation profile could be different to the actual usage pattern. This creates a variance between the estimated and actual usage at different times of the year and directly contributes to UIG variability. An inaccurate AQ would also result in a different overall level of consumption estimation compared to actual usage, contributing to the base level of UIG.

Item 3.2.1: Non-Daily Metered Sites above the Daily Metered Threshold

Through our thorough investigation into sites where the estimated usage is very different to the actual usage, we have found a small number of large sites which are read periodically but their AQ is above the threshold where they should be daily metered. Some of the sites are read annually so meter reads and AQ updates are infrequent.

"The difference between the estimated and actual usage could be contributing up to 0.4% of throughput to UIG on an average day"

Sites of this size are likely to have a unique usage pattern and the estimation profile is unlikely to be a good fit for each site's usage. **The difference between the estimated and actual usage could be contributing up to 0.4% of throughput to UIG on an average day and causing UIG volatility of up to 0.7% of throughput**.

To reduce the impact of sites like these on the industry, Xoserve recommended that we regularly monitor the market for these sites and engage with relevant Shippers to highlight the individual Meter Points, supporting action to re-confirm.

We also recommended adding new reports to the Performance Assurance Report Register to monitor these sites, using the UIG Weighting Factors to incentivise Shippers to make the relevant sites daily metered, and automatically changing meter read frequency to Monthly when AQ increases above 293,000 kWh which we think will improve read submission performance and therefore reduce the time the AQ is at risk.

Item 3.2.2: Sample sites with different consumption patterns or levels compared with UK Link

Where the consumption we hold on our systems is different from the actual gas usage, the AQ will not be correct and so the daily energy estimate will not be accurate, which will contribute to base level UIG. The UIG will be permanent if the actual energy is not recorded correctly following a meter read.

Xoserve maintains an offline sample where we record the actual daily energy for a small number of sites, and so to understand this issue we have compared the metered energy from our systems against the energy for identical consumption periods for around 6,000 of the sample sites over a 4 year period. The offline sample records more total energy than is recorded on our systems, so the energy recorded on our systems may be understated if the sample data is correct. **This could account for up to 0.25% of LDZ throughput currently attributed to UIG** if the difference in the sample is present in the whole market.

The mismatch could be caused by errors in the daily sample dataset, metering equipment errors, incorrect meter readings supplied to Xoserve, meter reads submitted with a meter read date different to the date the meter was actually read, or incorrect meter setup on our systems. Intelligence from Shippers suggests that **between 1% and 26% of sites could have asset errors requiring corrective updates**.

This is a complex issue and more insight would be valuable to fully understand and resolve it. That's why we recommended that we cross-check a sample of the meter asset data we hold against other data sources, to find out the level of asset data quality issues on our systems. This may make the case for a larger data cleanse exercise if we find a large number of mismatches.

Longer term, we recommended that Xoserve obtain smart meter readings and asset data direct from the Data Communications Company (DCC) rather than the shipper. Obtaining this data from the DCC would reduce complexity and improve accuracy of settlement data as sourcing it from a single, central provider reduces the risk of data loss or corruption.

Item 3.2.8: Sample sites consuming energy with a UK Link AQ of 1

Through our analysis of the offline sample data, we have found sites that appear to be using gas but that have (or had) an AQ of 1. Sites with very low AQs will not be allocated any energy when they are estimated and will contribute to base level UIG if they are using gas.

Our systems validate meter reads on receipt, and if the metered energy is outside a tolerance band they will reject the read. Where the AQ for the sample sites is now corrected on our systems, the new AQ is often greater than it could increase to under normal processes. This means there may be process blockers to increasing the AQ in this scenario.

There are around 400,000 live MPRs with an AQ lower than 100 kWh on our systems, and our analysis suggests around 5% of these (20,000 MPRs) have had reads rejected because they were outside the tolerance band and a subsequent read has not been accepted. This line of analysis is ongoing but **we think that that these low AQs could account for up to 0.35% of throughput attributed to UIG**.

We have shared the sites we have identified with the Xoserve Customer Account Managers and we will work with Shippers to get AQs corrected where required.

There are changes underway to resolve Xoserve system issues which can, in some cases, erroneously set the AQ to 1. We recommended that we run the impacted MPRNs through the AQ correction tool once the fix is in place rather than waiting for the AQ to increase through usual processes.

We also recommended enhancements to Performance Assurance Committee (PAC) reporting on the AQ corrections process usage. This is because we have been advised that some shippers may be using the AQ correction process outside of the functionality it was designed for. We have also suggested a number of system and process changes that the industry could make to reduce the risk of understated AQ contributing to UIG which are available in the recommendations pack for this item on our **website.**

Investigation Item 1: Use of Estimates for Daily Metered Sites

Daily Metered (DM) sites are generally very large sites like factories and power stations. Because their gas use is significant and can be unpredictable, these sites should submit meter reads every day. The Industry target is to provide 97.5% of valid daily meter reads to Xoserve the following day.

Where we do not receive actual reads or the reads are rejected, we estimate usage for the day using the same consumption as 7 days ago if available; otherwise we use the Annual Quantity (AQ) divided by 365. This may not be a good representation of the actual consumption and any difference would contribute to UIG.

As at 01/01/2019, **around 5bn kWh of DM AQ – 0.9% of total LDZ AQ** – has not had an actual meter reading accepted for over 3 months, and read submission performance is 86% for Class 1 and 45% for Class 2 against the 97.5% target.

DM sites' consumption can vary by -50% and +100% from the average on any given day, which **could contribute up to 0.9% of LDZ throughput to UIG volatility**. If the site usage has changed by 10% since the last reading it could contribute up to 0.09% of LDZ throughput to base level UIG.

"Daily Metered sites where Xoserve has not accepted a meter reading for over 3 months could be contributing up to 0.9% of daily throughput to UIG volatility"

To reduce this risk to UIG, we recommended that Xoserve monitor Class 1 read rejections engage with Daily Metered Service Providers to understand and resolve the root cause of any rejections. We also recommended that we monitor Class 2 read rejections and engage with Shippers, providing the support enabling them to submit the required readings. Where needed, Xoserve could refer individual sites and associated Shippers to Ofgem to ensure appropriate levels of oversight and reduce the risk to the market.

We also recommended reducing the period of missing readings before the Distribution Network obtains a reading on behalf of the Shipper for Class 2 sites, and extending this service to include Class 1 sites. We think that introducing incentives to increase Class 2 submission rates or extending Class 1 liabilities to Class 2 would also improve read submission performance and reduce the risk to UIG.

Investigation Item 2: Low take-up of Winter: Annual Ratio Band End User Categories

Large Supply Points have a Winter: Annual Ratio Band (the WAR Band) – an assessment of how much gas they consume over the winter period compared with the rest of the year. The WAR Band determines how much effect the weather has on a given site's daily usage estimate for that site. Each End User Category (EUC) in bands 3-8 has 4 WAR bands and a 'Bucket'' band, which the default where there is no WAR information.

The estimation calculation uses a default weather sensitivity profile for sites in the Bucket. This means that sites without a WAR band could be over or under estimated on a given gas day, which will contribute to seasonal UIG variability. If sites were distributed in WAR Bands at the ideal ratio, UIG would be proportionally **up to 2.5% lower on peak winter days**, and **summer would be up to 1.5% higher**. The net annual effect is a slight underestimation of energy.

Around 40% of eligible sites were not in a WAR Band sub-EUC and so are in the Bucket Band as of October 1st 2018.

"Around 40% of eligible sites were not in a WAR Band sub-EUC as of October 1st 2018, which is increasing UIG, especially in winter"

This can be because a suitable read pair was not accepted by Xoserve and so the Winter Consumption Calculation could not complete, and we have also found an an issue in our system which impacted the calculation for around 1,000 MPRNs in the 2018 calculation window.

Xoserve has a number of initiatives underway to improve this issue. The system issue will be fixed for the next Winter Consumption calculation window in May 2019, and we will support shippers to use the WAR Band update process to assign the impacted sites into the correct WAR Band. This support extends to all meter points which do not have a WAR Band assigned and we have shared full portfolio extracts with the Shippers so they can update them as appropriate. We will also highlight the reporting available following the annual Winter Consumption calculations and proactively manage the next steps with our customers.

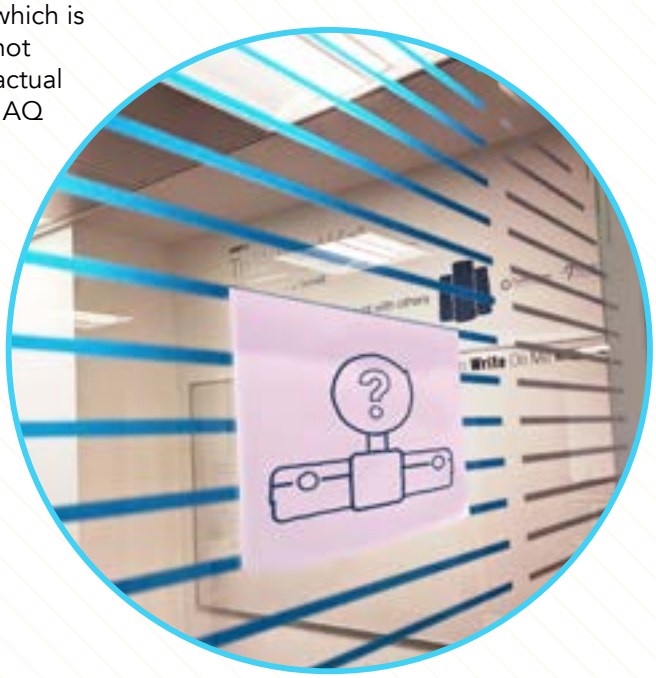
Performance Assurance reports are under development to support UNC Modification 0652 (Introduction of winter read/consumption reports and associated obligations), and we recommend that the report incorporates the current meter read frequency indicator as the WAR Band calculation can fail where the Meter Read Frequency is not Monthly for a site in a WAR Band eligible EUC. We also think there would be value in Xoserve monitoring sites which cross in to a WAR Band eligible EUC following Rolling AQ calculation, and proactively engaging with Shippers to use the WAR Update process when needed.

Investigation Item 3.1: AQ Calculation errors – Reads rejected because Uncorrected Read Value is lower than previous Uncorrected Read

Some sites have a corrector fitted to their meter, which adjusts the metered volume using the actual gas temperature and pressure. Sites with corrector provide both uncorrected and corrected readings to Xoserve.

We have found an issue where our system rejects reads when the new uncorrected read is lower than the previously loaded uncorrected read. The Through-the-Zeroes (TTZ) count in the read submission file applies to the corrected read as this is the read used for billing. There is no way to identify when the uncorrected read has gone through the zeroes and so the read will be rejected when the uncorrected register goes through the zeroes.

This issue impacts around 1,200 MPRNs and has resulted in 12,000 read rejections to date. Around 350 of the MPRNs have no subsequent read accepted. These sites have a total AQ of 2.8bn kWh, which is around 0.5% of total LDZ AQ. The AQ will not recalculate without meter reads, and if the actual consumption is significantly different to the AQ the site will not be estimated appropriately and will contribute to UIG.



We will flag this issue to shippers and in the short term, Xoserve can manually enter reads by exception. We recommended that we monitor uncorrected read rejections and manually load the reads to UK Link if the read has passed all other validations. This will require a level of Xoserve resource to maintain on an ongoing basis, and so longer term we recommended a change to our system to remove this validation check. This would allow valid reads to load.

Investigation Area 13.2: Accuracy of Non-Daily Metered (NDM) Algorithm – Use of Weather Data

Weather is a key input in to the NDM Allocation calculation, and so we asked our analytics partner to examine the way the NDM algorithm uses weather to calculate the daily energy estimate.

Item 13.2.2: Weather Sensitivity Analysis

If the NDM algorithm is particularly sensitive to certain weather factors but does not use those factors as effectively as possible, then the daily energy estimate could be higher or lower than it should be which would contribute to UIG volatility.

Our analytics partner created a simulation where they could vary the behaviour of the different weather factors to see what the impact on UIG would be. The table below shows which weather factors are very sensitive to this variation and so could contribute to UIG volatility.

Parameter or Input	UIG Baseline Sensitivity	UIG volatility Sensitivity
Weather Correction Factor	High	High
Seasonal Normal Composite Weather Variable	High	High
Cold weather coefficient	Low	Low
Wind chill coefficient	High	Medium
Effective temperature coefficient	Medium	Medium
Cut off coefficient	Medium	Low
Cold weather start temperature	Low	Low
Transition start temperature	High	Medium
Transition end temperature	Medium	Low
Temperature	High	Medium
Wind Speed	Low	Low
Seasonal Normal Effective Temperature	Medium	Medium
Wind chill factor limit temperature	Low	Low
Wind limit speed	Low	Low
Effective temperature weight	High	High

This does not mean that there is an issue with the way the algorithm uses these factors, but shows us where it is especially important that the algorithm represents the relationship between the factor and energy use as small changes to the value of these factors can cause large changes in UIG.

The algorithm has a temperature 'memory', where it uses the previous day's weather to smooth out the effects of rapidly changing temperature. We set out to understand if reducing the influence of the temperature memory in the algorithm, or removing it entirely, would improve the daily energy estimate and so reduce UIG volatility. We found that removing the weather memory can improve the estimate and reduce UIG volatility over winter, but it can make the estimate worse and increase volatility in the summer, and so using the temperature memory selectively throughout the year may improve UIG levels.

The algorithm uses 12 temperature measurements and 6 wind speed measurements taken each day to create a single factor to represent that day's weather. The measurements are weighted because

the weather at different times of day can have varying influence on energy use. The sensitivity analysis revealed that the weightings have more influence on the estimate in spring and autumn, and highlights that energy usage is sensitive to the temperature in the early morning and afternoon. This time-of-day and time-of-year sensitivity **might be contributing up to 0.2% of throughput to UIG**, and so adjusting the weightings, potentially for each season, could improve the accuracy of the daily energy estimate and reduce UIG.

We have shared these results with Xoserve's Demand Estimation Team and in partnership with the Demand Estimation Subcommittee they will decide whether we will change the way the algorithm uses the weather memory and weather weightings in the future.

Item 13.2.5: Basic Machine Learning and the Use of Additional Weather Data

The NDM algorithm is based on the proven, strong relationship between weather and energy usage. This is a 'linear' relationship, which in simple terms means that for every degree the actual temperature differs from the 'normal' temperature, the amount of energy used will move away from 'normal' use by a fixed amount. For instance, for each 1° change in temperature, the energy use will change by 8%. A sophisticated version of this basic principle forms the foundation of the NDM algorithm.

There have been significant developments in predictive algorithms in recent years. These algorithms 'learn' from what actually happened in the past to estimate future behaviour. In particular, modern computer power has given rise to the use of Artificial Intelligence (AI) and Machine Learning (ML), and we wanted to understand whether using these cutting edge techniques could estimate energy use better than the existing NDM algorithm.

Our analytics partner is world-class in this space, and has deployed a number of sophisticated algorithms to try and create a better estimate of daily energy usage. We found that accurately estimating daily energy usage is a very complex task, and that the complexity of the current NDM algorithm is necessary to achieve the level of accuracy we see today.

The analysis demonstrated that AI and ML have the potential to create a better energy estimate than the NDM algorithm. Using the same inputs as the current formula, the ML algorithms reduced the simulated level of UIG by 0.8% of national throughput. Including additional weather data – the warming effect of sunlight, rainfall, humidity, and atmospheric pressure – as inputs to the algorithm improved the accuracy of the estimate further and could potentially **reduce UIG by 1.2% of throughput**, which is a 30% improvement over the current algorithm. **The ML algorithm also reduced the largest spikes in UIG volatility by around 40%**.

"Machine Learning shows potential to reduce the largest spikes in volatility by around 40%"

We have shared these findings with the Demand Estimation Team and in partnership with the Demand Estimation Subcommittee they will assess whether including this additional weather data in the current NDM algorithm could improve the accuracy of the energy estimate and reduce UIG.

These positive results have generated new lines of analysis into how we can improve the NDM algorithm.

You can read more about this work on page 20.

Investigation Area 13.1: Accuracy of Non-Daily Metered Algorithm - Including End User Category Definitions

Non-Daily Metered (NDM) Sites account for over 60% of total energy usage on a given day. We don't know how much energy a site has actually used until the meter is read, which can take over a year. The industry relies on the NDM algorithm to create a daily estimate of how much energy each type of site will use, and to allocate that energy to each market participant. Xoserve develop the inputs to the algorithm in partnership with the industry, and the work is coordinated and reviewed at the Demand Estimation Subcommittee (DESC).

The NDM algorithm is very complex, and there are numerous input variables which can influence the accuracy of the daily estimation. We tasked our analytics partner to examine the components of the NDM algorithm and find out if it is working as well as it can be. In this investigation area, we examined the way the algorithm uses fixed factors to influence the daily estimate.

Item 13.1.1: 2018-19 EUC01 Uplift Factors

The NDM algorithm is reviewed every year and for the 2018-19 Gas year, DESC voted to use a set of 'Uplift Factors' developed by one market participant. These factors increase the amount of energy allocated to End User Category (EUC) band 01. EUC01 is mostly domestic properties with some smaller commercial users; sites like small shops and offices. The factors only apply to the daily estimate, and so their impact will be removed when a meter read is accepted on to our system.

Our analysis suggests that these factors would have reduced estimated UIG by around 2.5% of national throughput on average in the year following the introduction of UIG, but they mask the true UIG position. This appears to be the case in the first months of the 2018-19 gas year, where UIG is running at around -0.5% of throughput but **without the factors the true UIG position would be around 2% of throughput**. DESC will monitor the performance of the factors and decide whether to continue using them going forward.

Item 13.1.2: Weekend vs. Weekday correlation

Consumers use different levels of energy on different days of the week. Generally domestic customers will use more energy at weekends and factories and offices will use more energy in the week. We use over 500 different energy estimation profiles to estimate energy more accurately for users with different levels of energy consumption. These profiles use the day of the week to inform the estimation, and so we asked our analytics partner to see if they could find a relationship between different days of the week and UIG.

"We use over 500 different energy estimation profiles to estimate energy more accurately"

Their analysis did not find a relationship between UIG and the day of the week, which tells us that the NDM algorithm is correctly estimating the change in use we see on different week days.

Item 13.1.3: Holiday Factor Analysis

Consumers also use different levels of energy on bank holidays, and holiday demand can influence surrounding days. We call these events holiday factors. Generally domestic consumers will use more energy on holidays but factories and offices will use less. The estimation profiles use these holiday factors to inform the estimation, and so we asked our analytics partner to see if they could find a relationship between holidays and UIG.

Their analysis did not find a relationship between UIG and holiday periods, which tells us that the NDM algorithm is correctly using the holiday factors to estimate the holiday period usage changes.

Item 13.1.4: Influence of Geographical Factors on Demand Estimation

Xoserve maintains an offline sample of sites where we record the daily energy usage for each site. We have thousands of sample sites across the country representing different types of consumer. If the sites in the sample are spread out across the country in a different pattern to the full population of gas users, then the energy patterns we use to create the estimation profiles might not be a good fit for actual gas usage. Any difference between the estimated and actual energy usage will directly contribute to UIG.

"We have thousands of sample sites across the country representing different types of consumer"

We asked our analytics partner to compare the location of the sample sites to the full population to see if there was a relationship with levels of UIG. Our goal was to understand if there were any geographical factors (for example, proximity to coastline, relative population density, and the distance from the weather stations where we measure the temperature and wind speed used in the daily estimation formula) which could influence the estimation calculation and contribute to UIG.

The analysis did not find any relationship between the location of the sample sites and UIG. While there are pockets of sites where the estimate could be improved, there was no pattern or trend we could find. This tells us that there could be other factors that influence gas usage which are not related to geography. While we have no recommendations at this stage, we have used this analysis to shape the work we are doing to improve the NDM algorithm.

Investigation Area 13.3: Accuracy of Non-Daily Metered (NDM) Algorithm – NDM Sample Data

Xoserve maintains an offline sample of sites where we record the daily energy usage for each site. We have thousands of sample sites across the country representing different types of consumer. If the sites in the sample behave differently to the full population of gas users, then the energy patterns we use to create the estimation profiles might not be a good fit for actual gas usage. Any difference between the estimated and actual energy usage will directly contribute to UIG, we asked so our analytics partner to examine the sample to make sure it represents the whole market.

Item 13.3.1: NDM Sample Representation Across End User Categories

The daily energy estimation uses over 500 custom usage profiles to represent sites with different energy usage levels or patterns. We call these groups End User Categories (EUCs). We review the algorithm inputs every year using a subset of the sample sites for each EUC. If the sample site energy usage in an EUC is different from the actual energy use of the all the sites in the EUC then the daily energy estimation might not reflect actual usage which would contribute to UIG.

"Breaking EUC01 into sub-bands has the potential to reduce the amount of energy reconciled when meters are read"

We asked our analytics partner to compare the spread of users in each EUC between the sample and the full market. We found that while the larger EUCs in the sample have a similar spread to the full market, the sample for the lowest EUCs contains a greater proportion of sites with higher energy use than the wider market. If sites with different levels of energy consumption also use energy in a different pattern over the year then this could mean that the daily estimate is not a good fit for actual daily usage, which would contribute to seasonal UIG variation.

From the analysis in Item 13.3.2 below, we know that EUC 01 is the largest contributor to UIG at allocation, and so the majority of any impact to UIG caused by the different spread of sites in the sample compared to the whole market would be in EUC01. To understand if this was having an effect on the daily estimate, our analytics partner created a number of simulations of the Non-Daily Metered algorithm input analysis for EUC01. Each simulation was adjusted in different ways to see what could happen if the makeup of the sample was more like the wider market.

To begin, they increased the influence that sample sites with lower energy usage had on the development of the simulated estimation profiles. This sample weighting reduced UIG in the winter and increased it in the summer, but the effect was very small and made virtually no difference to UIG over the year.

They then subdivided EUC01 into five smaller sub-bands, and created a unique energy estimation profile for each sub-band. These new profiles did reduce the level of UIG for the sites in each sub-band, but the overall difference to UIG when all the estimates from the sub bands were added together was again very small.

This means that the makeup of sites in the EUC01 sample does not appear to contribute to UIG, and so Xoserve have no recommendation from these findings for reducing market level UIG. We did learn that breaking EUC01 into sub-bands has the potential to reduce the amount of energy reconciled for EUC01 when the meters are actually read. Reducing the amount of reconciliation energy is a key objective for the Demand Estimation Subcommittee, and these findings show potential to help achieve that objective so we have shared them with Xoserve's Demand Estimation Team for further analysis.

Item 13.3.2: NDM Sample Population Analysis

Our analytics partner compared the recorded daily usage from the sample against the Annual Quantity (AQ) held on Xoserve's systems. This revealed that only around half of the sites had an AQ within 10% of the actual usage, and there were a small number of sites which had actual usage over five times greater or smaller than their AQ. The general trend identified that energy appeared to be understated on our systems. This analysis continued under items 3.2.2 and 3.2.8, and you can read our findings and recommendations on Page 9.

We also asked our analytics partner to attempt to identify which EUCs were the biggest contributors to UIG. They completed a detailed investigation into the difference between the actual energy recorded and the daily energy estimate for the sample sites: the sample's 'UIG'. They then created a method to scale the sample UIG to the full market, which tells us which EUCs are most likely contributing to the national UIG level.

We discovered that **sites in EUC01 and non-daily metered sites in EUC 09 are the biggest contributors to both base level UIG and daily UIG volatility.** EUC02 is also appears to contribute to UIG volatility.

"Improving the daily energy estimate for these EUCs has the most potential to reduce UIG base levels and volatility"

We have analysed NDM sites in EUC09 in more detail, and our findings and recommendations can be found in Item 3.2.1 on page 8. The results from this analysis have focussed the Machine Learning tasks on EUCs 01 and 02 because improving the daily energy estimate for these EUCs has the most potential to reduce UIG base levels and volatility.

Other Items

The Task force examined a number of existing industry processes and known issues to understand whether they were impacting UIG. We did not find a material impact to UIG from these items and so have no recommendations at this time.

Investigation Item 6: Shipperless and Unregistered Sites

Shipperless and Unregistered sites are Meter Points which are live on our systems, but are not associated with a Gas Shipper. If these sites are actually using energy then they will contribute to UIG.

The Allocation of Unidentified Gas Expert (AUGE) is an organisation appointed by Xoserve to assess known and suspected causes of UIG, producing sharing factors to allocate UIG across the industry. The AUGE assessment is that Shipperless and Unregistered sites account for between 0.01% and 0.015% of throughput attributed to UIG. There are currently around 5,000 Shipperless and Unregistered sites, and as the UIG impact is very small and Xoserve has a dedicated team to monitor and resolve these sites, we have no recommendations for action to make at this time.

Investigation Item 8: Daily Metered Nomination Accuracy

Daily Metered Sites are very large, and because unexpected changes in usage can impact other sites and the wider gas network, Shippers must tell us how much energy they think their sites will use before the Gas Day. This is called Nomination. If Nominations do not match with the actual energy use then this could cause uncertainty which might contribute to the estimated pre-Gas Day UIG. Some Shippers will buy and sell gas based on these predictions, so getting them right is very important. There are already nomination incentive charges in place; the more accurate the final Nomination, the lower the incentive charge, but this doesn't apply to earlier Nominations.

We analysed the full Nomination history for a two year period, both before and after Project nexus implementation, to understand if Nomination accuracy could pose a risk to UIG levels. We found that Nominations are usually very accurate by the second nomination run, and there is not a significant level of variability in nomination accuracy over time.

We compared nomination accuracy to UIG levels and could find no relationship, so DM Nominations do not appear to contribute to UIG.

Investigation Item 18: Isolated Sites with Incoming Reads

When a site is Isolated, the Shipper warrants that it is not capable of using gas. We found around 4,000 isolated sites where shippers are submitting meter reads to Xoserve that tell us the sites are using energy.

Isolated sites are not included in the daily energy estimate and so these sites are directly contributing to the UIG baseline. The sites have a total AQ of 33m KWh, or 0.006% of annual throughput attributed to UIG.

This issue will be progressed for resolution through the appropriate industry processes, and the low materiality means that we have no further recommendations to make.

Next Steps

Xoserve are planning a number of activities for the coming weeks to realise the benefits of the Task Force's findings and to explore more potential root causes of UIG.

The UIG Task Force will work with Xoserve's customer advocates and other colleagues to share relevant, key messages with our Customers, and effectively partner with the industry to drive forward solution options that we think will reduce UIG. Xoserve will support our Customers in drafting UNC Modifications and Change Proposals. We will work with industry forums to further develop options to address complex challenges like the Use of a Standard Conversion factor as highlighted in item 12.2.

Our investigations into UIG root causes are ongoing. We are continuing to explore the use of Machine Learning to improve the NDM algorithm, and our Analytics Partner are currently developing a Neural Network model with the goal of increasing the accuracy of the daily energy estimate and reducing UIG. They are also investigating historic differences between the energy estimate and actual metered energy. We want to see if we can find any new trends which could help us make the estimation profiles better, and to understand how much UIG could be accounted for as meters are read in the future.

Xoserve are investigating long term AQ trends and the impact rolling AQ could be having on UIG levels. There are 17 more lines of investigation to pursue, and Xoserve will assess these for their potential UIG impact and start our analysis where we feel there could be most benefit to the industry.

We will continue providing regular updates on the progress of our investigations and the impact on UIG as the Industry and Xoserve take action on our website at www.xoserve.com/uig. We will continue reporting the progress of our analysis on the Investigation Tracker, and in partnership with the Joint office, we will record the actions we are taking to resolve UIG on the UIG Task Force Recommendation Tracker.

If you would like to discuss UIG with Xoserve, you can reach the UIG Task Force via email at **uigtaskforce@xoserve.com**; the team will be happy to help.