Water Breakthrough Challenge
Transform Stream

Managing Background Leakage

Stage A Report

Final

Date: 07/02/2023

Issue: Final v.1.0

1 Introduction

1.1 What is background leakage?

Background leakage in district meter areas (DMAs) is based on minimum achieved night flows following a sweep of the area to find and fix leaks. By deducting an allowance for night consumption, the minimum achieved leakage (MAL) level is derived. The difference between the MAL level in a DMA and a theoretical minimum achievable level of leakage (MAbL) using background leakage default estimates from the Managing Leakage reports can be considerable. The range of MAL values in DMAs is wide, with a long tail to the distribution. Some DMAs can achieve very low levels of network leakage, whereas other DMAs have a persistently high MAL

1.2 Why is it important to address background leakage?

Water companies in England and Wales have Ofwat targets to reduce leakage by 15% in the current planning period to March 2025, known as AMP7. Water UK has issued a Public Interest Commitment (PIC) to triple the leakage reduction rate by 2030 and has accepted an NIC (National Infrastructure Commission) proposal to halve leakage by 2050. A new roadmap towards achieving that aim refers to the need to address Background Leakage, as clearly the long-term aim cannot be met unless substantial progress is made in reducing it.

Customers and regulators seek a downward trend in leakage which is seen as wasted water; but more importantly wasted power and chemicals for treatment and distribution, adding C02 emissions, impacting the challenge of achieving net-zero carbon, and adding to environmental water abstractions impacting ecology.

The problem faced by the UK water industry is that around 50% of reported leakage is due to **Background Leakage**; the level remaining in a distribution network sector after surveys to find and fix leaks, mainly using acoustic techniques. Without tackling **Background Leakage**, reductions after 2025 will be very challenging and meeting the aim of halving leakage will not be possible.

Background Leakage has since 1994 been defined as the sum of small leaks below a detectable threshold and it has generally been accepted that it can't be reduced without large scale mains and service pipe replacement and / or by reducing pressures in the network. However, there is a hypothesis that some **Background Leakage** comes from old long-running leaks, not detected by current methods, and from other sources not currently understood.

This project aims to redefine the detectable limit of leakage with forensic investigations of 25 district meter areas, deploying flow, pressure and temperature sensors at an intensity never previously done, coupled with digital-twin network models to localise then pinpoint and repair hidden leaks and resolve other factors contributing to background leakage. This benefits customers by creating more sustainable ways of reducing leakage, avoiding increased environmental water abstractions if future leakage targets can't be met by current means.

UKWIR has also identified the need to understand, diagnose, locate and reduce background leakage as high-priority in its Leakage Big Question research roadmap. However, very little progress has been made to date. Substantial knowledge gaps remain on what comprises BL, where it occurs, and whether it is leakage at all. Therefore, UKWIR has confirmed it supports this proposed project; expecting it to make progress on these questions in ways not possible previously.

1.3 What is the estimated level of background leakage?

Leakage comprises three types of leaks and bursts. **Reported bursts** come to the surface or affect customer supplies and are located and repaired quickly. Most bursts are "reported" but have short run times, contributing less to overall annual leakage (estimated as c.10% prior to the start of this project). **Unreported bursts** have to be found by active leakage control (ALC) and have longer run times contributing and estimated 30-40% leakage. The remaining estimated 50-60% of **Background Leakage (BL)** is not a key part of current BAU (business as usual) leakage reduction strategy. It is the residual after all detectable bursts have been found and fixed; defined in the 1994 Managing Leakage reports as being unavoidable; due to the accumulation of small leaks below a detectable limit. Technology has improved since the 1990's when default values for BL were set, but background levels have not reduced greatly.

Water companies AMP7 target is to reduce leakage by minimum 15%. Current focus is on reducing run times of unreported leaks through more detection staff and increased use of sensors. However, there is a limit to how far leakage can be reduced in this way.

1.4 The Ofwat funded Managing Background Leakage Project

The hypothesis being tested in this project is that in certain DMAs some background leakage comes from a small number of old hidden leaks that aren't found by current techniques. For example, acoustic methods work well in metal pipe networks in urban areas where fittings are close together. They are not as effective in plastic networks where leak noise does not travel far, or in rural areas where there can be long distances between fittings and where leak noise has to travel further.

To test this forward-looking hypothesis that some background leakage is due to hidden leaks not located by current mainly acoustic techniques, despite intensive surveys, the project partners have been working with the supply chain on emerging methods using accurate flow, pressure and temperature measurements. Smarter, more integrated use of current technology at greater density in the right places combined with deployment of emerging technology can redefine background leakage.

The project aims to further industry understanding of background leakage by developing an emerging technique and aiming to halve background leakage in a sample of DMAs, going beyond current BAU by localising, pinpointing and repairing long-running hidden leaks on water company assets and customer supply pipes.

The project involves surveying 25 DMAs to assess background leakage as estimated by the Minimum Achieved Level (MAL) using BAU techniques, then innovating new intensive forensic approaches to determine whether some leakage comes from long-running hidden leaks/bursts, not located by current techniques, so re-defining the MAL level.

Emerging techniques for pinpointing leaks are impracticable for widespread deployment. A key project aim is to localise leaks to smaller sectors within which such techniques become practicable, sustainable, cost-effective.

By challenging the accepted wisdom on background leakage being due to leaks below a detectable threshold; the aim is to redefine achievable minima, assessing the lowest level of leakage possible.

1.5 The Phases of the project

The overall objective is to show whether background leakage can be reduced using more intensive novel techniques. The aim is to make cost-effectiveness of the technique comparable to other ways of meeting leakage targets and bridging supply-demand gaps (<£1m/Ml/d).

The project has several phases:

- Phase 1 collates information on current understanding/scale of background leakage, selecting 25 pilot DMAs within 5-months of start.
- Phase 2 records DMA starting leakage levels, reviews leakage following acoustic surveys, then re-assesses true night-flow components split into customer night-use, customer-side leakage, and network leakage using logged data.
- Phase 3 utilises innovative digital-twins comparing actual and modelled data, identifying anomalies to localise hidden leaks to small sectors; aiming to locate 5 leaks per DMA.
- Phase 4 employs emerging technologies to pinpoint leaks, and records leakage levels after repairs. We aim to halve background leakage in each DMA.
- Phases 2,3,4 in 5 to 10 Group 1 DMAs provide the ultimate answer with over-instrumentation in a 16-month period. Remaining Group 2 DMAs maximise operation efficiency over 12-months.
- Phase 5 over 12 months uses advanced data-analytics to capture components of error, uncertainty, statistical and hydraulic effects to explore the effect on background leakage estimates
- Phase 6 develops a Dissemination Plan for publicity, Exploitation Plan for roll-out, and Project Report on work done, results, outcomes and implications for the industry, within 1 month of completion.

Our approach involves steps not yet tested at scale:

- Current background leakage estimates rely on customer night-use assumptions. The project will utilise customer-meter data (smart-meters plus logged customer-meters), fitting novel Stop.Watch temperature loggers to unmetered connections to determine true consumption profiles; hence true levels of leakage, considering flows across the whole logging period, not just the night-time window when leakage levels are normally assessed.
- In the first 5-10 DMAs (Group 1) over-deploying instrumentation (BAU and emerging technology), applying data analytics showing how different data can be analysed/utilised in different combinations, spanning from existing data-driven analytics (anomaly detection) to latest-generation digital-twins (focussing on hybrid data-driven deterministic modelling). Key to defining ultimate background leakage is quantifying legitimate demands, customer-side, network leakage, errors and uncertainties.
- Combining individual-property consumption profiles, using state-estimation techniques, flow, temperature, high-frequency (100Hz) pressure data and acoustics (surface mounted and in-pipe hydrophones) into digital-twin DMA models, comparing logged and modelled data; to localise hidden leaks to small sectors, ideally to single streets.
- Then using a variety of techniques in suspect sectors, currently not practicable for whole DMA adoption, to pinpoint the hidden leak, including in-pipe and surface-mounted technologies.
- Selectively down-sampling the Group 1 data-set, showing how the background leakage changes as a function of combinations/intensity of data collection; defining functions of achievable leakage levels for different levels/strategies of deployment; using stochastic demand-estimators to better inform real usage, so that every house doesn't need a smart-meter for the method to be successful.
- After the down-sampling findings, then sequentially progressing through 15-20 more DMAs to refine, then fully confirming and demonstrating what is possible.

1.6 Phase 1

Phase 1 collates the information on current understanding and scale of background leakage in the five participating water companies and selects 25 pilot DMAs within 5-months of project start. This report sets out the results of the Phase 1 work. The scope of Phase 1 included:

Start Up Meetings: A start-up meeting with participants to review the project plan/methodology.

Several on-line meetings have been held with the representatives of the 5 participating water companies.

A full day workshop was held at University of Sheffield attended by 40 representatives from 9 water companies, the contractors Invenio and UoS, the 3 technical experts, and people for the water industry supply chain.

At these meetings the project team, with the participants, reviewed the methodologies and the programme taking account of partner comments, other information sources, relevant changes between award and project start, sources of background leakage information and work with partner water companies

The information obtained was used in the selection of the 25 DMAs for fieldwork.

Knowledge review: In Phase 1, the project team reviewed the status of current club contracts, ongoing work by water companies, and other research, to ensure the project takes account of current knowledge. A summary of the current definition and estimated levels of background leakage is included as Appendix 1

Data analysis: The Invenio team analysed DMA leakage data using fixed night-use values and varied night-use values according to new consumption algorithms. The results are presented in this report.

DMA selection: From the analysis of data a process was established to select the DMAs to be surveyed in the field work.

2 Data Analysis Method

This section sets out the method of data analysis undertaken in Phase 1 of the project.

2.1 Stage 1 summary data

Each of the five water companies provided water-company level data to draw conclusions on what background leakage was being reported, its consistency with other companies and then subsequently compared to results from stage 2 work.

The following data was requested:

- The reported KPI Ml/d leakage.
- The bottom-up Ml/d value in the MLE calculation.
- The level of trunk main and service reservoir leakage included in bottom-up.
- The top-down Ml/d value in the MLE calculation.
- The background level of leakage included in the most recent SELL refresh for AMP7. This should be for DMA leakage only, excluding leakage due to reported bursts, based on MAL

levels in DMAs, a theoretical MAbL value, or a combination of the two, along with a brief description of how it was derived.

- An estimate of the level of leakage due to reported bursts based on the number in 2020/21, their run time, and the average flow rate.

2.2 Stage 2: Detailed Data

Data was collected covering individual DMAs in the participating companies.

2.2.1 Data request

The following data were requested:

Individual DMA data for the three years 2020/21, 2019/20 and 2018/19 to include:

- The weekly MNF (Minimum night flow) values.
- The MAL (Minimum achieved leakage) level achieved in that period.
- The weekly average daily flow values.
- The length of mains.
- The number of properties.
- The non-household night use allowance.
- The household night use allowance.

Also:

- Data on the methods used to estimate bottom-up leakage in DMAs.
- The methods used to determine the components of night consumption.
- The estimates of trunk main and service reservoir leakage and how they are. derived
- Burst numbers and run times for reported and unreported bursts.

The quality and comprehensiveness of data varied between the companies, and there were some differences of interpretation and methods.

2.2.2 Data selection

The initial analysis aimed to identify the lowest achieved leakage over a three-year period in each valid DMA. In order to obtain a reliable result, a number of filters were applied to the raw data. These included:

- Reject DMAs with data quality flags provided by the water company, if applicable.
- Reject any DMAs where the commercial allowance makes up more than 25% of the total night use allowance.
- Reject any DMAs with fewer than 50 properties.
- Reject any DMAs with fewer than 5 unique values of MNF or NFL (Night flow leakage) over the entire 3-year period, or where more than 50% of the MNF or NFL values are identical (this indicates that missing data may have been infilled with a single value).
- Reject DMAs where the gradient of the MNF data is constant for more than 25 consecutive weeks (this indicates that missing data may have been linearly interpolated).
- Weekly Data Filters:
    - Remove any weeks where the MNF or DD (Daily demand) are negative.
    - Remove any weeks where the DD is greater than 1500 l/s.
    - Remove any negative peaks lasting only a single week.
    - Remove negative weeks where there are fewer than 10 negative values in total.
- Reject any DMAs where fewer than 52 weeks (1 year) of data remain after the previous weekly filters.

It is worth noting that data was not rejected simply because the calculated leakage was negative. The prevalence and distribution of apparent negative leakage DMAs provides a way to estimate errors in measurement.

2.2.3 Leakage estimation

Night flow leakage was estimated in these DMAs using two alternative methods:

- Subtracting the night consumption allowance in the established manner.
- Subtracting an allowance for night consumption that was proportional to the difference between minimum night flow and average daily demand (the ratio method).

2.2.4 MAL estimation

The minimum achieved leakage in each DMA was then calculated. For each of the two methods of estimating night flow leakage, the minimum value of the resulting three years of leakage data was determined. Because the absolute minima are sensitive to outlying low leakage values, the fifth percentile value was also calculated for each of the two different leakage estimates. This meant that, in total, four different values of the lowest achieved leakage were calculated for each DMA.

2.2.5 Selection of best estimate of MAL

Out of the four values obtained for each DMA, the most suitable was selected based on the reliability of the results, particularly the amount of leakage that was apparently negative.

2.2.6 Interpretation of Results

The results of the analysis were examined to understand the distribution of lowest-achieved leakage between DMAs. This included:

- Assessing how minimum achieved leakage was related to property count, mains length and other readily available parameters
- Assessing the cumulative minimum-achieved leakage as a function of minimum-leakage cut-off. This is intended to identify how much minimum-achieved leakage might be reduced by targeting a relatively small number of long-term high leakage DMAs
- Identifying how minimum-achieved leakage is distributed by property and mains length for each company and as a whole.
- Comparing minimum achieved leakage to the estimate of background leakage reported in stage 1.

At the stage of surveying each DMA the logged pressure data will be used to derive the average zone pressure (AZP) which will allow an estimate to be made of the theoretical MAbL value (see Appendix 1) for comparison with the actual MAL level.

2.3 Results of DMA analysis

The results for each of the participating water companies are shown in graphical form in Appendix 2. In this section the results are summarised for all the companies.

At this stage, company level data is included for Affinity Water but the DMA level analysis is not because it is currently based on MAL values supplied by the water company and not analysed by Invenio from raw flow profiles. Therefore, it is inconsistent with the other 4 companies. However, it will be included in a future update of this report and will be reviewed prior to the field work.

The reported background leakage at company level is shown in the tables below.

Company Reference	Background leakage in DMAs Ml/d	Adjusted Unreported DMA Leakage (Ml/d)	Adjusted reported DMA leakage (Ml/d)	Total DMA leakage (Ml/d)	Trunk main and service reservoir leakage (Ml/d)	Total MLE KPI leakage (Ml/d)
Anglian	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DCWW	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Portsmouth	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Severn Trent	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Affinity*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Total	456.1	289.8	47.9	793.8	141.6	935.4

*Table 1: Summary results of stage 1 reporting. *Background in Affinity Water has been estimated from minimum achieved in individual DMAs whereas other results are based on background estimates from previous SELL updates.*

Table 1 shows the component values of the reported KPI level of leakage for each company in Ml/d. The total of 935.4 Ml/d is about 30% of the total for England and Wales; so it is a good size sample.

Company ref	Background leakage in DMAs (%)	Revised Unreported DMA Leakage (%)	Revised reported DMA leakage (%)	Trunk main and service reservoir leakage (%)	Background in l/prop/day
Anglian	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
DCWW	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Portsmouth	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Severn Trent	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Affinity*	[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
Average	49%	31%	5%	15%	51.1

*Table 2: Proportions of leakage components in individual companies. *Background in Affinity Water has been estimated from minimum achieved in individual DMAs whereas other results are based on background estimates from previous SELL updates*

Table 2 shows the same data as Table 1 as a percentage of the reported KPI leakage level and shows background leakage in litres/property/day.

Background leakage varies from 37% to 71% of the total with the property weighted average across the companies being 49%. Adding the 5% for reported DMA leakage which is difficult to reduce further makes 54%. This supports the original premise for the project that if background leakage is half of the total then the industry must address it if it is to meet the challenge of halving leakage by 2050.

Reported DMA leakage varies from 1% to 10% with the average being 5%. There is some uncertainty in these values due to the method of estimation. It is possible that the actual average figure is higher due to the 1% reported for [redacted] being an under-estimate.

Trunk mains and service reservoir leakage upstream of DMAs varies from 0 (as [redacted] include all mains in DMAs) to 21% with the average being 15%. Again, there is uncertainty over these values due to the different methods of estimating upstream leakage.

The unreported DMA leakage is the balance in the sum i.e. it is 100% less the other three components as there is not a way of estimating it directly. It varies from 19% to 36% with the average being 31%.

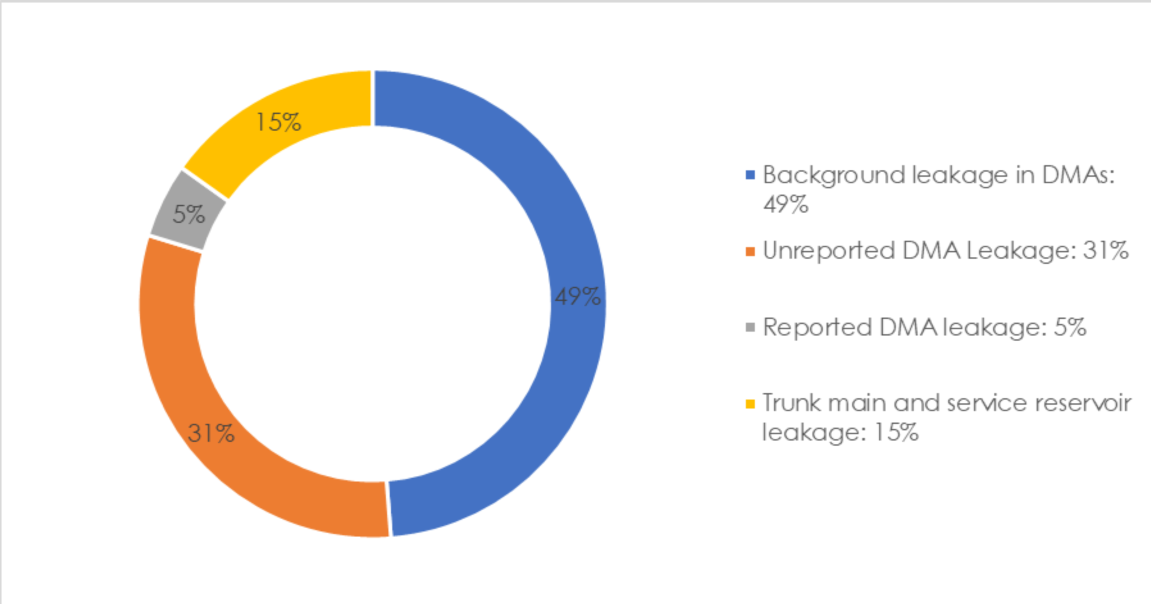


Figure 1: Proportions of leakage components in the companies as a whole

Figure 1 shows the property weighted average values in doughnut graph form. Current focus in the industry is on reducing run times for unreported DMA leaks by greater monitoring and having ALC staff for find and fix. However, that only impacts 31% of the total.

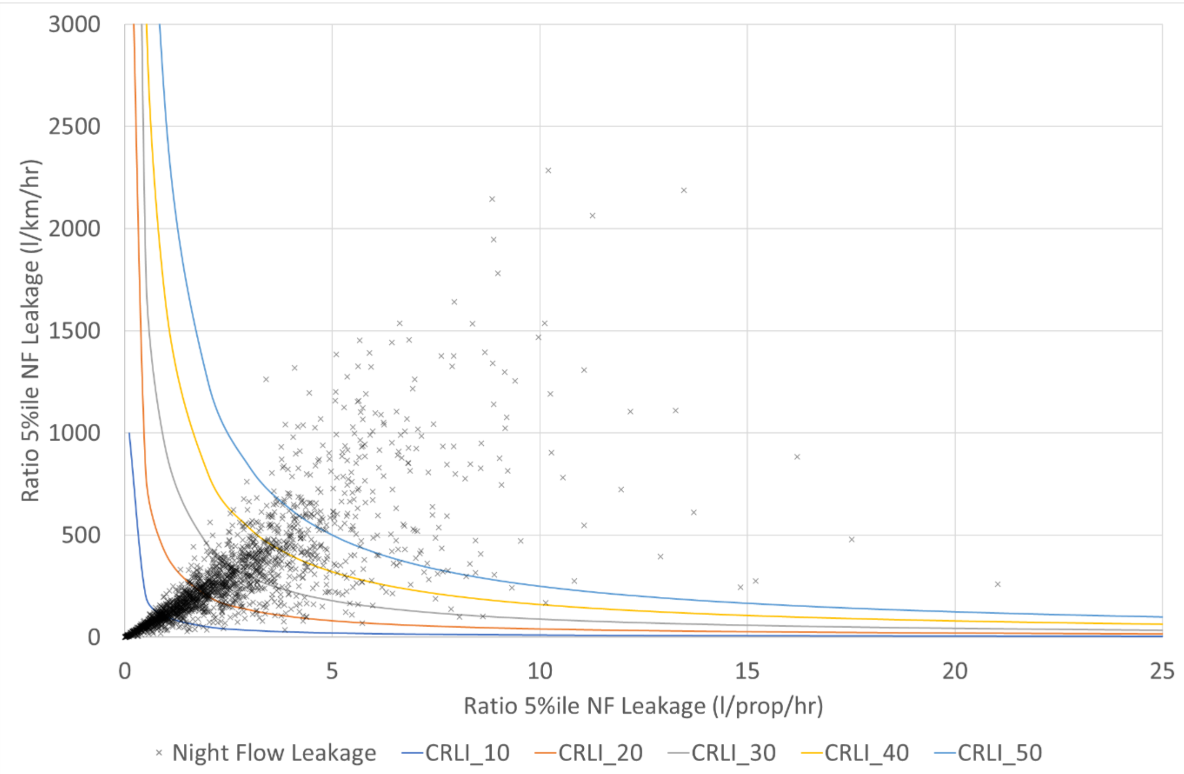


Figure 2: Range of DMA-level lowest achieved night flow leakage using three different measures

Figure 2 shows the range of 5%ile MAL values across the DMAs in the participating companies on two axis graph scaled by the length of mains and the number of properties. The CRLI contours are there to show the lines of equal ‘leakiness’ for urban and rural DMAs with different lengths of mains per property supplied.

It can be seen that there is wide range of CRLI values with a grouping at the low end tending towards zero. At the other extreme there are DMAs with very high CRLI values in both the urban sector (top left) and rural sector (bottom right). This challenges the current definition that background leakage is due to a large number of small leaks. Were that the case a far narrower range of MAL values would be expected.

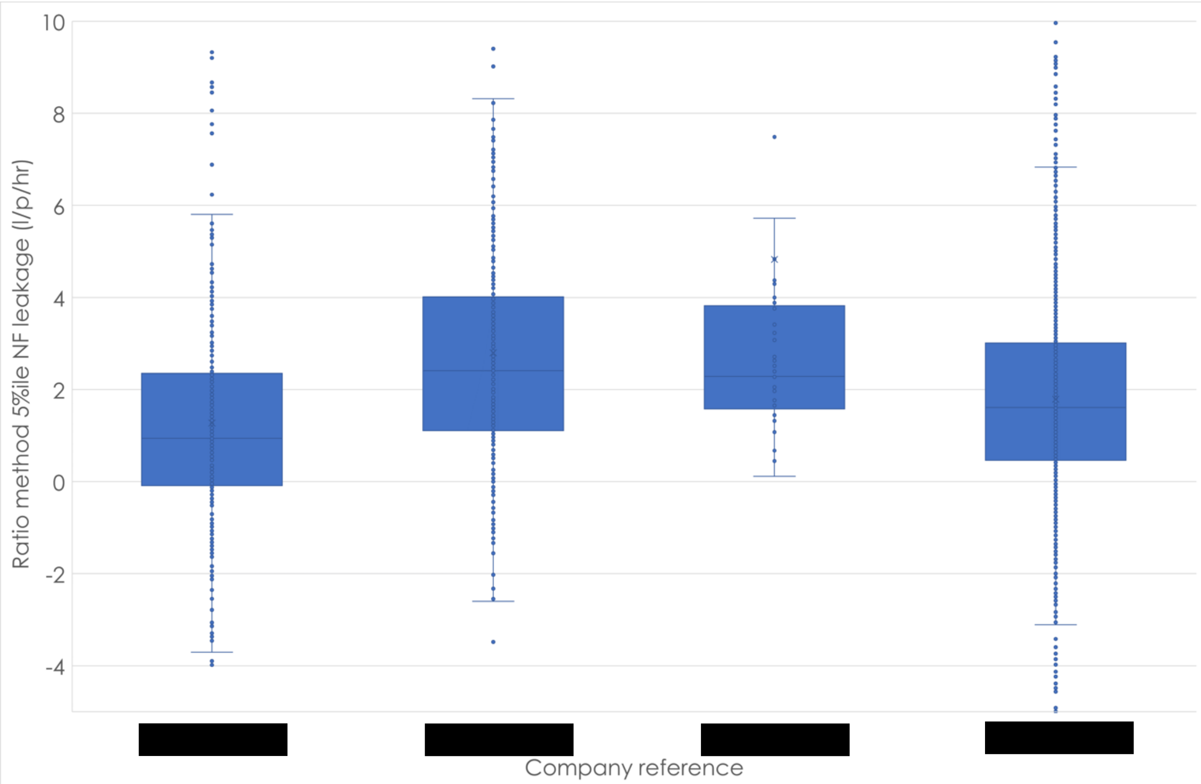


Figure 3: Box and whisker chart of ratio-method 5%ile leakage by water company

Figure 3 shows the range of lowest achieved DMA leakage (using the 5%ile ratio method) for the individual companies. The box shows the interquartile range and the whiskers the 95%ile range, with values for individual DMAs shown by dots.

Figure 4 shows the distribution of MAL levels across the DMAs in the participating companies. There are a significant number of DMAs with negative MAL levels due to over-estimation of the actual night consumption. Most of these DMAs are small rural areas which contribute a lower proportion to the total leakage level. The peak of the distribution is between 0.5 and 2.0 litres/property/hour. There is then a long tail with some DMAs having MAL levels in excess of 12.5 l/prop/hr. The field work stage of the project will focus on 25 DMAs in this tail with the hypothesis being that some contain long running hidden leaks which if found and fixed would reduce the MAL to what appears to be a reasonable range of say 0 to 4.5 l/prop/hr.

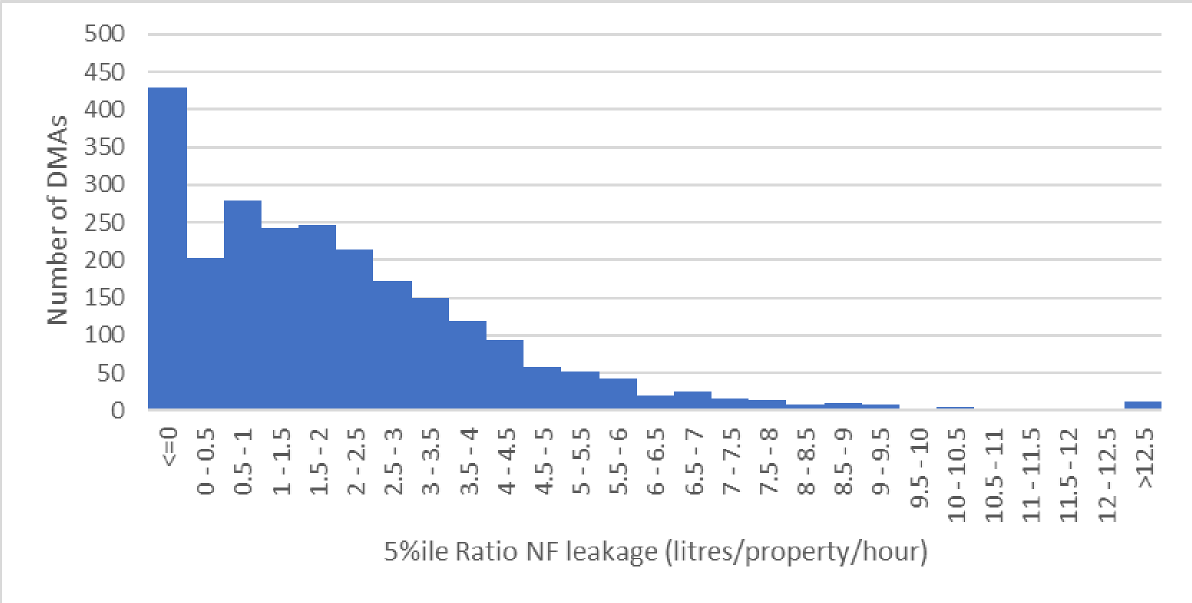


Figure 4: Histogram of number of DMAs by 5%ile ratio-method night flow leakage (l/prop/hr)

A MAL of 4.5 l/prop/hr at a 22 hour day is 99 l/prop/day which is about twice the average background leakage level from table 2.

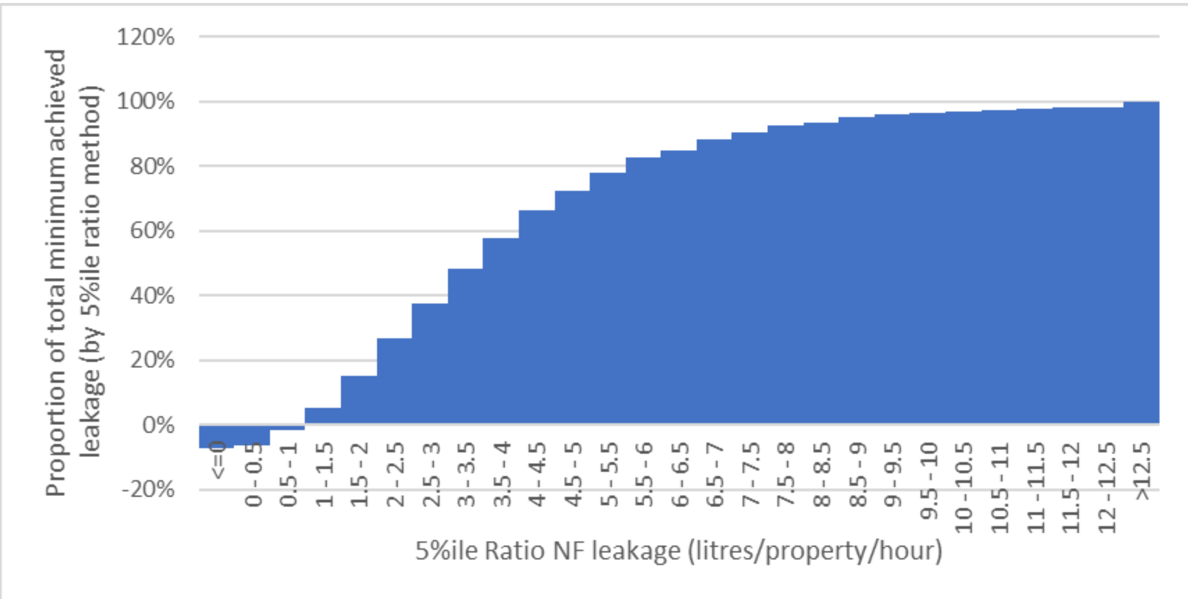


Figure 5: Cumulative proportion of lowest night flow leakage by litres per property per hour.

Figure 5 shows the contribution of the DMAs in each 5%ile MAL band to the total level of leakage. From this it can be seen that although there are over 400 DMAs with negative MAL values, they only contribute –7% to the total leakage level. At the other extreme those DMAs with MAL levels over 4.5 l/prop/hr which we are targeting contribute around 28% to the total leakage.

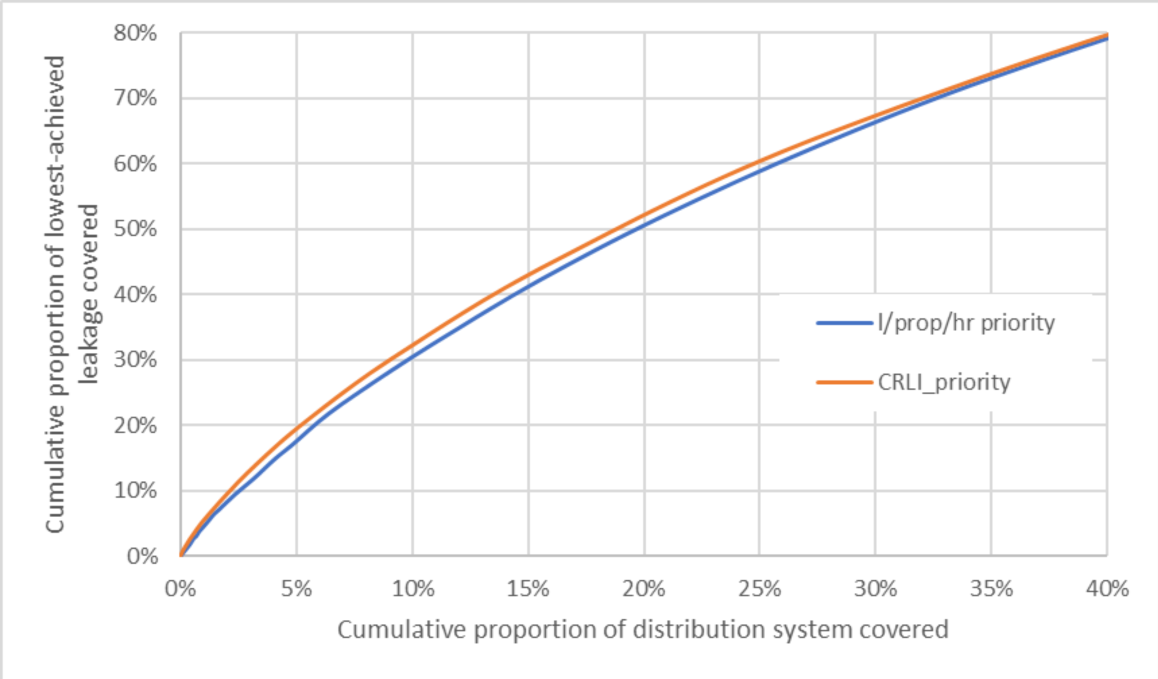


Figure 6:Cumulative proportion of lowest achieved leakage covered as a function of the proportion of the system investigated, prioritised by either litres/property/hour or CRLI.

Figure 6 shows the how much of the lowest-achieved leakage could be investigated by prioritising the highest lowest-achieved leakage DMAs. Two prioritisation methods are shown: litres per property per hour of night flow leakage and CRLI night flow leakage. CRLI takes account of both property count and length of distribution system. Both approaches produce very similar results. Reading off the graph it shows, for example, that 50% of the lowest-achieved leakage could be investigated by investigating just 18% of the distribution system using the CRLI prioritisation method (or 20% of the distribution system using l/prop/hr prioritisation).

3 Selecting the 25 DMAs for survey

This section describes the process used to select the 25 DMAs for survey and summarises the data for the DMAs chosen. The diagram below outlines the process used.

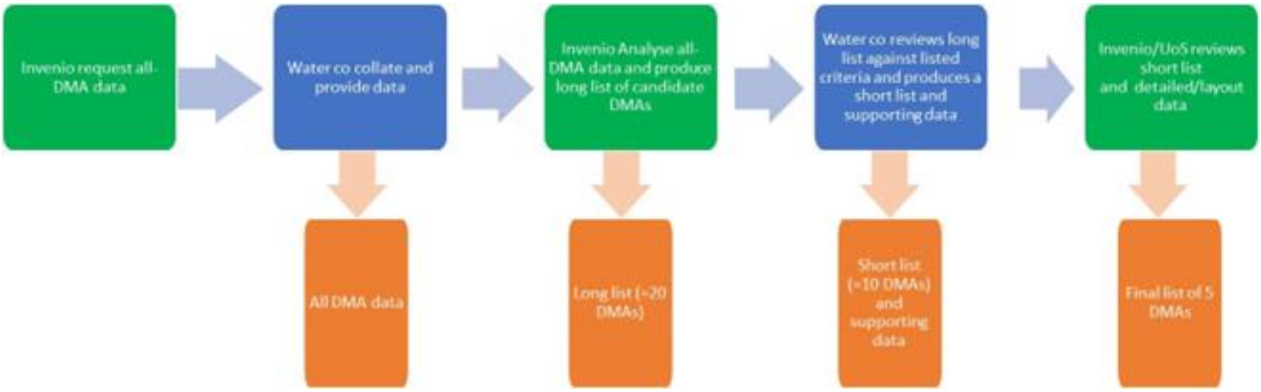


Figure 7: The DMA selection process

3.1 Selection process

From the DMA analysis outlined in section 2, Invenio produced a long list of potential DMAs for each of the five companies. The list was shared with each water company for review, following which a short list was developed. Discussions between Invenio and each water company led to the selection of the proposed 5 DMAs in each company taking account of the criteria in 3.1. The other DMAs on the short list are held as reserves, as it is likely some of the DMAs chosen for survey at this stage will be changed as work proceeds for operational and other reasons such as:

- Leakage being reduced significantly by the water company before the survey is due
- Leakage being reduced significantly as a result of the pre survey ALC sweep
- Works being carried out in the DMA such as mains rehabilitation

3.2 Selection Criteria

The criteria for addition to the long list were:

- that there had been minimum achieved leakage (using the individual company's standard night use allowances) for at least three years above zero.
- that there had been minimum achieved leakage (using the individual company's standard night use allowances) for at least three years above 30 (l/(km.prop)^0.5/hr)
- that there had been minimum achieved leakage (using a ratio method) for at least three years above 30 (l/(km.prop)^0.5/hr)
- Between 500 and 900 total properties in each DMA

Note that all of the DMAs eligible for the longlist had already been selected to reject those where non-household night use is a significant part of the night flow (Section 2.2.2).

DMAs were selected for the shortlist from the list provided by each water company after their own checks. The shortlisted DMAs were selected to provide a range of:

- minimum achieved leakage when normalised by both property count and a measure that takes account of both property count and distribution system length.
- Average zone pressure

- Pressure management at the DMA inlet
- Mains length per property

3.3 Chosen DMAs

The DMAs chosen for survey are set out in Appendix 3 and summarised as follows.

The 25 DMAs contain 17,585 properties; an average of 703 per DMA which is in accordance with the funding submission that stated an average of 700 properties per DMA. The range is from 474 to 899 properties.

The total length of mains is 321km; an average of 12.8km per DMA with an average of 18.2 metres of main per property.

The total MAL leakage in the 25 DMAs is 137.5 M3/hr (around 3 Ml/d) and the average is 7.82 litres/property/hr.

4 The survey and analysis process

`This section describes the process to be followed in each DMA.

4.1 Initial ALC sweep

The first step is to verify the MAL level in the DMA using current BAU processes for active leakage control (ALC). In advance of the consumption survey, the water company or its contractors will carry out a sweep of the DMA and then arrange for any leaks located to be repaired. The resulting leakage level will be compared to the historic MAL of the DMA and if it is significantly higher a second BAU sweep will be carried out.

The aim is to generate an initial report for each DMA on completion of this stage. The DMA name and reference number will be anonymised to maintain confidentiality.

4.2 Consumption survey

On completion of repairs from the initial ALC sweep, Invenio will carry out a consumption survey of the DMA. The aim is to derive an actual consumption profile for every property supplied rather than using estimated consumption values. This will be achieved by analysing data from Smart meters, and meters already fitted with data loggers, and by logging of all other available meters. For properties and other supplies that are not metered, the Invenio Stop.Watch logging system will be used; this is a patented system that analyses frequent and accurate temperature measurements taken from the external stop tap.

The aim will be to log every property for a period of one to two weeks, and from that to sum these into a total consumption profile for the DMA. The consumption will be split into usage (based on intermittent use events) and customer side leakage (based on continuous flows past the customer meter or stop tap).

The consumption profile will then be compared to the profile of flow into the DMA from the logger fitted to the DMA meter. The difference is taken to be leakage on the water company network between the DMA meter and the customer supply points. The average 24-hour profile for the logging period is used for this purpose as shown in the figure below.

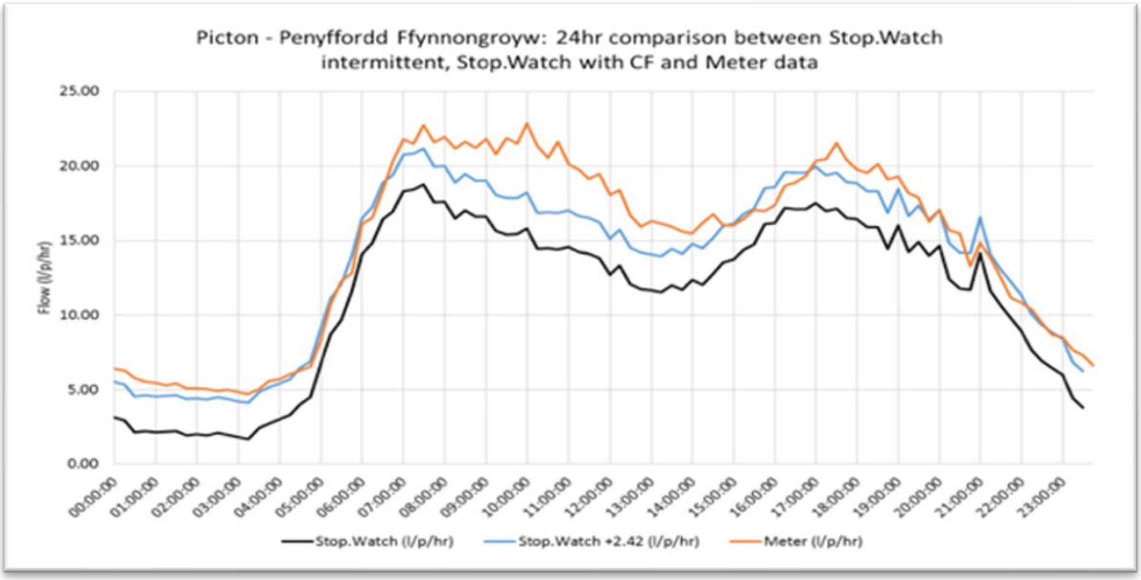


Figure 8: Average daily DMA flow profile showing components (low network leakage)

The black line is the total of the intermittent use events averaged over the logging period. The data is displayed on a per property basis as it is rare to obtain 100% coverage of good quality data for every property in the DMA. The blue line is the black line plus the continuous flows beyond the customer supply point i.e. customer side leakage (internal and external). The orange line is the flow into the DMA from the DMA meter. In this case it can be seen that at night there is very little difference between the DMA flow and the total flow into customer properties, so there is very little network leakage. It also shows that there is some unaccounted-for use in the DMA from around 7am to 3pm.

In other DMAs there can be a considerable network leakage. In the DMA below the difference between the blue and black lines represents a continuous loss of water upstream of the customer supplies.

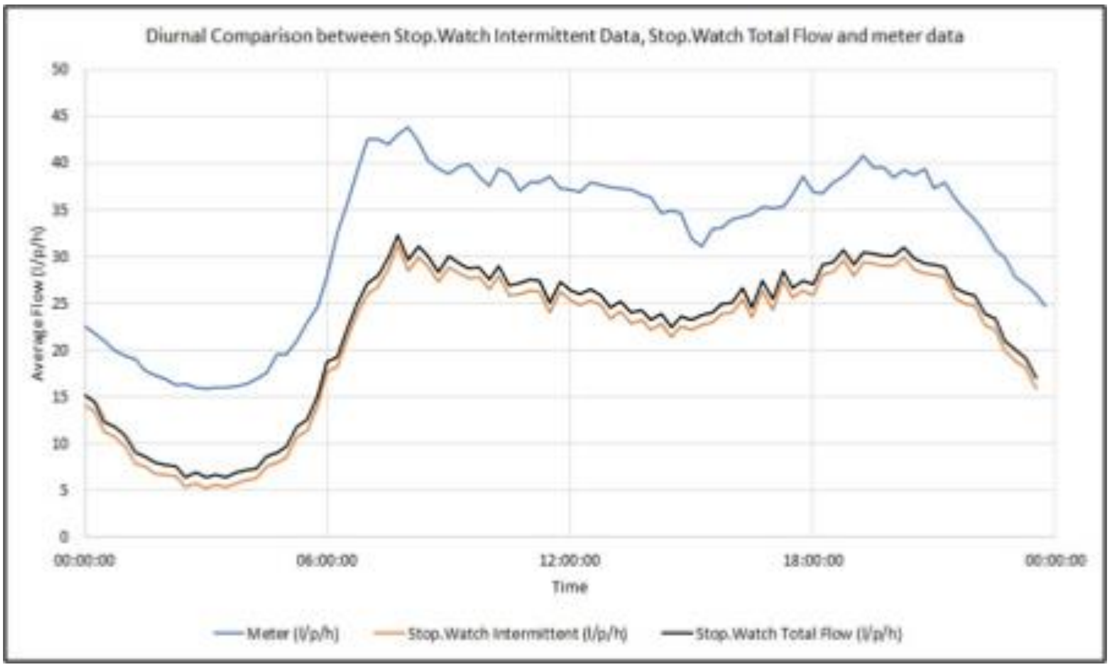


Figure 9: Average daily DMA flow profile showing components (high network leakage)

4.3 Pressure Survey

During the consumption survey period, Invenio will also fit fast frequency pressure loggers to hydrants in the network and take topographic measurements to obtain datum heights for each logged location. The data will be used in the modelling stage.

4.4 Modelling

The hydraulic modelling will be undertaken by University of Sheffield using static data about the DMA from each water company and dynamic data from the field surveys supplied by Invenio. The static data will be derived from GIS records and by extracts from network models where they are available.

The process for modelling is described in more detail in Appendix 4. The aim of the modelling is to identify anomalies between the logged pressure and the modelled pressure which could indicate the presence of leak, some other form of unaccounted for water, or a discrepancy between the static data used in the model and the actual physical situation on site e.g. a valve being in a different open / closed status.

To avoid the network being identified and to maintain confidentiality of the customer consumption data, the model will be anonymised as set out in Appendix 5.

Invenio will also analyse the temperature data from the inlet of the DMA to each property supplied and may install additional temperature logging points as required, again to determine anomalies between actual and modelled data.

4.5 Leak localisation

The aim of the modelling is to localise anomalies in the DMA to within one or two 'streets'. A street for this purpose is taken to be a length of main between two junction nodes representing around 10 to 15% of the DMA.

4.6 Leak pinpointing

The next step is to undertake intensive examination of the localised area using a variety of techniques depending on several factors such as:

- The length of main between fittings
- The diameter of the main
- The mains material
- The number and type of connections off the main
- The physical environment e.g. across fields, under roads etc
- The operating environment e.g. can the main be isolated for long periods without impacting customer supplies ?

The techniques to be employed include:

- Use of the ground microphone
- Using hydrophone acoustic loggers
- Tracing the route of the main
- Checking for possible connections not recorded on GIS
- Use of in pipe survey techniques

4.7 Repair and report

On completion of the leak pinpointing, any leaks found will be repaired, following which the individual DMA report will be updated to include the work carried out and the results.

4.8 Data Flows

In Phase 1, the project team spent a great deal of time discussing the flow of data between the various parties, recognising the agreement to publish data in an open source way at regular intervals. The diagram below summarises the flow of data and whilst it is too detailed to explain in this report, it is included to show the complexity of the process.

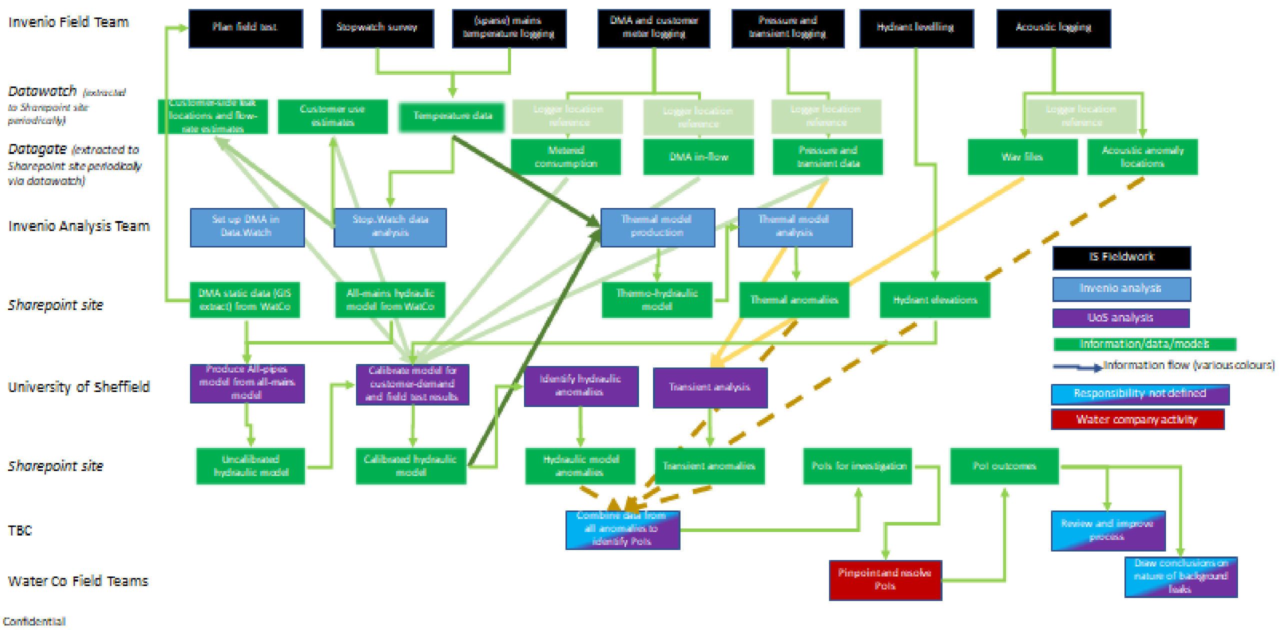


Figure 10: Data flows

5 Current Status

5.1 First two DMAs

The initial leak detection sweep has been completed in the first two Welsh Water DMAs and repairs are being carried out. In January 2023 the installation of data loggers will commence for the consumption survey.

While the consumption survey is in progress, University of Sheffield will create the network models for the two DMAs.

5.2 Forward programme

On completion of the two DMAs in Welsh Water, the aim will be to complete 2 DMAs in each of the other 4 water companies in Stage 1 following a similar process. Lessons learned from Stage 1 will be used to revise the methodology for Stage 2 with the intention of reducing the survey and analysis cost.

Then the field team will return to Welsh Water to complete 3 DMAs in Stage 2 followed by 3 DMAs in each of the other 4 water companies.

6 Summary

This report gives the overview of leakage levels across the 5 companies participating in the project and sets out the process by which the 25 DMAs for survey have been selected. It is inevitable that changes will be required as work proceeds due to operational and other reasons. Therefore, it is proposed that this report be updated at regular intervals.

The report also provides background on the current understanding of background leakage and the hypothesis being tested through the project.

Appendix 1

Review of Current Knowledge

The concept of Background Leakage was introduced in the early 1990s, during the National Leakage Initiative that ran from 1990 to 1994, and which lead to the development of the original Managing Leakage reports. The BABE (Burst and background estimates) method is set out in a paper by Allan Lambert published in the IWEM (a forerunner of CIWEM, the Charted Institute of Water and Environmental Management) journal in April 1994.

Report E of Managing Leakage 1994 defined Background Leakage in a DMA as the level of leakage from the collective sum of minor leaks on fittings that could not be identified on night flow measurements. Background leakage was deemed to be "the collective sum of numerous minor leaks and seepages from valves, joints, hydrants, stop-taps, meters and boundary boxes on mains and services pipes; and from dripping taps and overflows from lavatory cisterns and roof tanks". It was believed that these would rarely exceed 100l/hr and would not be identifiable from DMA night flow measurements. The report also said that individually they would be below 500l/hr.

Background leakage is best assessed from night flow data in DMAs, and it was appreciated that the actual leakage level of background leakage would be a function of the attributes of the DMA in terms of length of mains, number of connections, number of properties and average pressure. It was considered that the level of detection could be expected to reduce in time as new technologies for leak detection were introduced.

Data on the minimum level achieved on a number of DMAs was obtained from some companies, and from this default values were suggested for unit losses per length of main and number of connections (communication pipes and supply pipes). A range of values were found and so the values were grouped as low, average or high. It was suggested in Report F of Managing Leakage 1994 that these could be interpreted as relating to the age and condition of the assets as 'good', 'average' and 'fair' respectively. All values were converted to be at 50m pressure. These values are reproduced in Table 5 below.

Background Loss Component	Units	Low	Average	High
C1: Distribution mains	l/km/hr	20	40	60
C2: Communication pipes	l/prop/hr	1.5	3.0	4.5
C3: Supply pipes	l/prop/hr	0.5	1.0	1.5

Table 3: Background night flow losses at 50m. Source: Managing Leakage Report D, Table 5.1

We have sought to understand the amount of data that went into the derivation of these values and how the analysis was undertaken. It is not clear, but it is our firm belief that it was significantly less detailed and involved far less data than has been analysed for this project.

The supply pipe component of 1.0 l/prop/hr included plumbing losses. *Managing Leakage* 1994 suggested that "50% of the background supply pipes losses should be attributed to the underground supply pipes and 50% to plumbing losses", i.e. 0.5 l/prop/hr each. At the time a default value for night use was 1.7 l/prop/hr which together with this plumbing loss value gave a night consumption value for households of 2.2 l/prop/hr.

A 2018 UKWIR report showed that the level of internal plumbing loss was likely to be considerably higher. Fast logging of consumption monitor areas and use of smart meter data has also shown that night use is also likely to be significantly more than was thought in 1994.

A 1997 UKWIR reported entitled *Updating Managing Leakage* introduced the term **Base** level of leakage defined as “the aggregation of sources of loss which are individually too small to be detected by active leakage control involving visual and acoustic inspection of all accessible fittings on mains and communication pipes, or by inspection of customer meter readings taken for normal billing purposes”. However, the 2012 update of Managing Leakage recommended that the term be discontinued.

A 2003 UKWIR report entitled “Background Leakage” defined background leakage as “Leaks too small to be found with current technology” and this is the terminology recommended in the 2012 update of Managing leakage. The report introduced new terms that remain relevant. This report introduced the term **policy minimum** leakage with the intention of making it clear that the actual minimum level of leakage could be changed by company action such as leakage detection techniques, skills, technology and rigour, by rehabilitation of the network, by the introduction, extent and type of pressure management and by changing company policies on external customer metering, frequency of reading, etc. Policy minimum levels of leakage (PML) should be assessed at a zonal level rather than individual DMAs. PML is effectively the level to which leakage will tend to asymptote as expenditure on active leakage control increases, with the ALC current policy, current pressure and a level of asset renewal which maintains the current network condition. Policy minimum will include the leakage from bursts which are “reported” and do not have to be detected by ALC operations.

A 2005 UWKIR report entitled *Towards Best Practice for the Assessment of Supply Pipe Leakage* suggested that the policy minimum could include leaks that are in their early stage of growth and therefore cannot be simply leaks from fittings. From this point of view policy minimum leakage could be seen as background leakage plus losses from leaks that will grow into detectable leaks at some later date. But even this does not take into account the nuances (and differences) in supply pipe leakage that would arise due to different company policies, on customer metering for example.

An UKWIR study in the 2009/10 programme looked at the factors affecting background leakage. A draft report was circulated, and we believe the title was changed to Factors affecting minimum achieved levels of leakage. However, the report was never published or accepted by the industry.

The 2012 update of Managing Leakage set out the current understanding of background leakage in Report 4 with several clarifications:

- The split between supply pipe loss and internal plumbing loss referred to above.
- that communication pipe losses are a function of the number of connections and not the number of properties, and therefore the units were amended.
- It was estimated that the average length of the underground supply pipe in the sample was 15m. It could be argued that the losses should be scaled for different lengths of supply pipe; this is discussed further in the 2005 UKWIR report on supply pipe leakage.

No specific research work has been commissioned to confirm or update these values since *Managing Leakage* 1994.

4 of Report 4 includes for these clarifications, together with the addition of the recommended standard for the Infrastructure Condition Factor which is the ratio of the actual MAL in a DMA to the level of background loss from the table.

Early work on ICF used the 'Average' values from Report D but *Updating Managing Leakage* 1997 used the values at 'Good' condition as an estimate of the base level of leakage for their work on the assessment of the Economic Level of Leakage. The 2003 UKWIR study on *Background Leakage* reported that "many companies have subsequently achieved minimum leakage levels well below the Managing Leakage estimates". However, UK practice is to continue to relate condition to the original Managing Leakage 'Average' condition.

4**Background Loss Component**	**Units**	**Condition**		
		Good	**Average**	**Poor**
Infrastructure Condition Factor (ICF)		0.5	1.0	1.5
C1: Distribution mains	l/km/hr	20	40	60
C2: Communication pipes	l/conn/hr	1.5	3.0	4.5
C3: Supply pipes – UGSP: either (av length 15m) or	 l/prop/hr l/km/hr	 0.25 16.7	 0.5 33.3	 0.75 66.7
C4: Supply pipes – plumbing	l/prop/hr	0.25	0.5	0.75

Table 5 Background night flow losses at 50m pressure (amended)

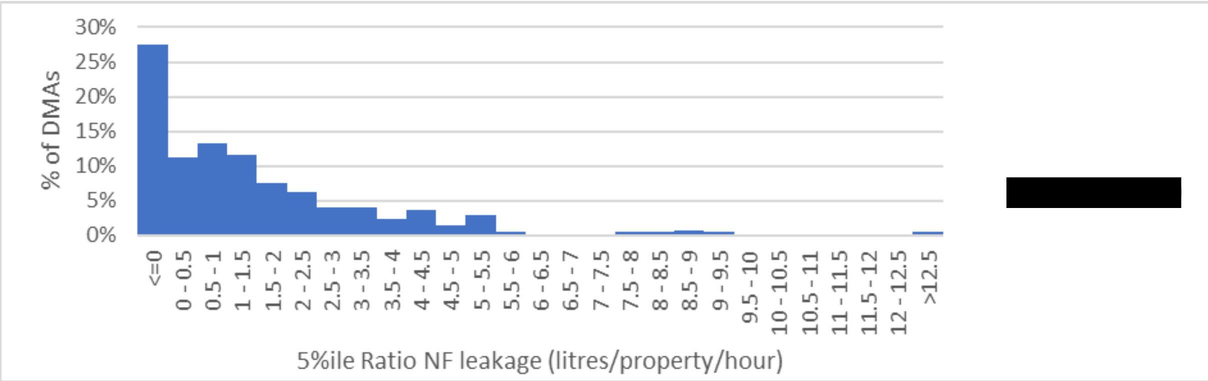
The 2003 study on background leakage found a strong relationship between estimates of the level of background leakage and the leak break-out rate on some of the data provided by companies. It was believed that this may indicate that background leakage is not only made up of minor weeps from gaskets but is made up of leaks in their early stage of development, before they are leaking at a rate at which they can be detected. Thus, an area that may have a high burst frequency will have a high number of leaks in this early stage and hence a higher level of background leakage. This concept of background leakage being made up of leaks in their early stage of development was taken further in a subsequent 2005 UKWIR report on supply pipe leakage. This report recommended the use of the 'Good' condition factors in *Managing Leakage* in all circumstances.

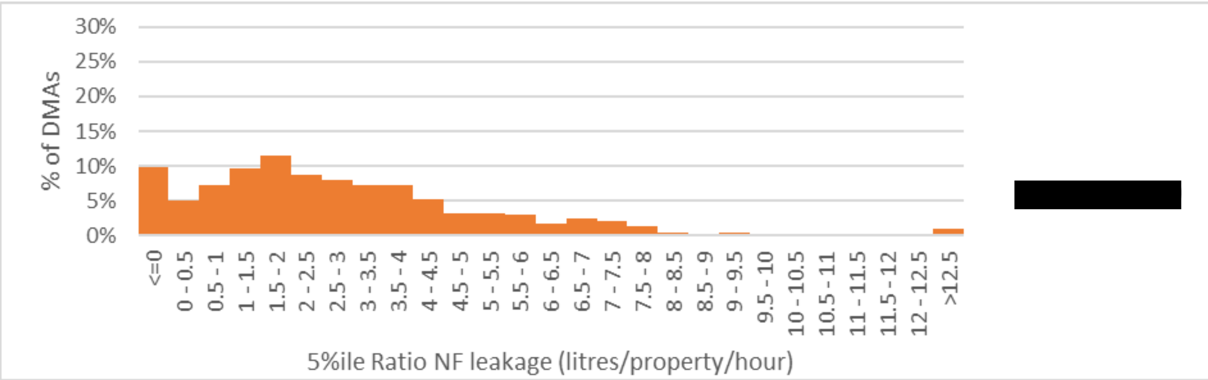
An alternative explanation is that areas with a high leak break out may also have a higher potential for leaks being undetected for a long time, which then add to the MAL level and appear to be background leakage.

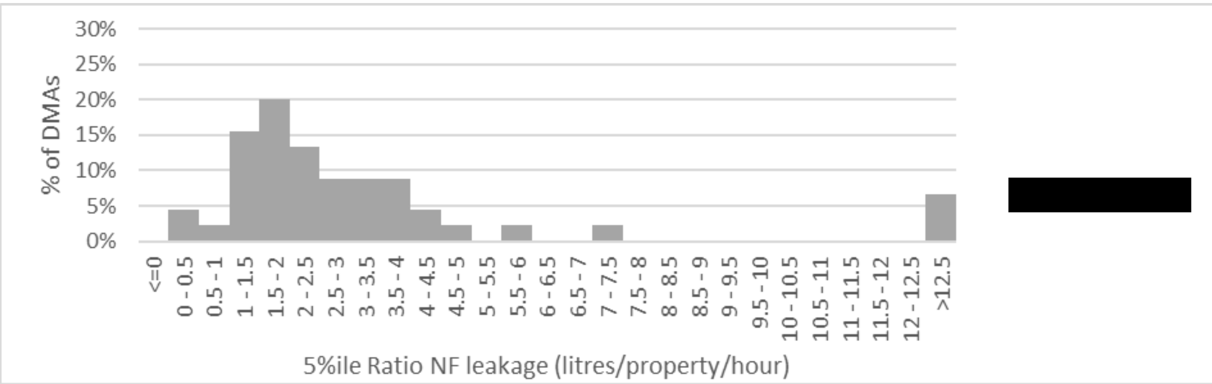
Appendix 2

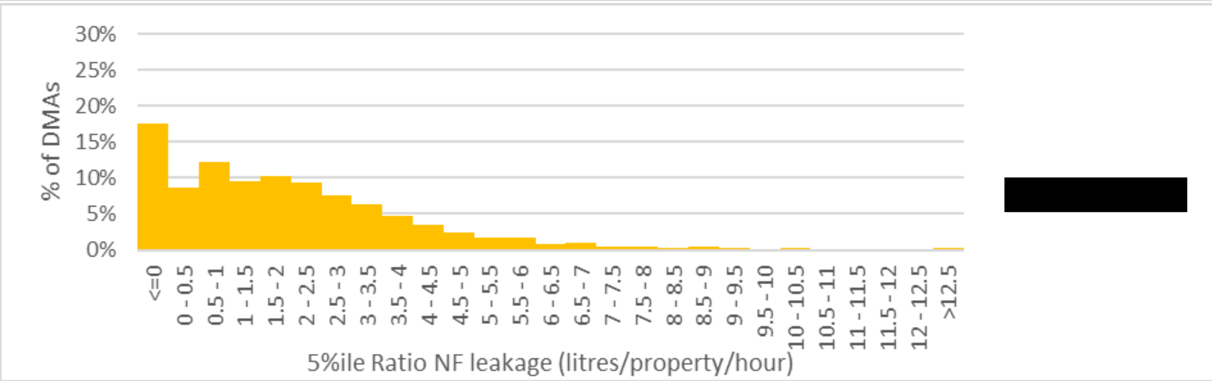
Results of the DMA data analysis

The DMA MAL levels in each company show a similar trend, with a proportion of DMAs exhibiting negative leakage (probably due to an over-assessment of night consumption); although [redacted] has very few. There is then a peak in the distribution at around 1 to 2 l/prop/hr followed by a long tail with each company having DMAs with MAL levels over 12.5 l/prop/hr.





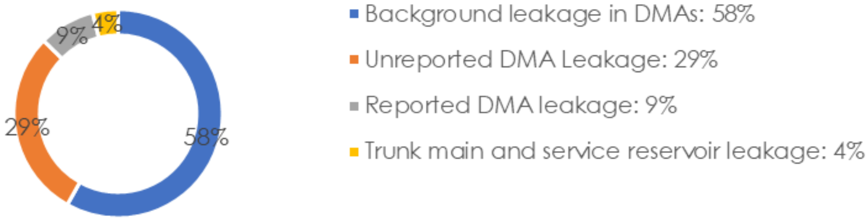




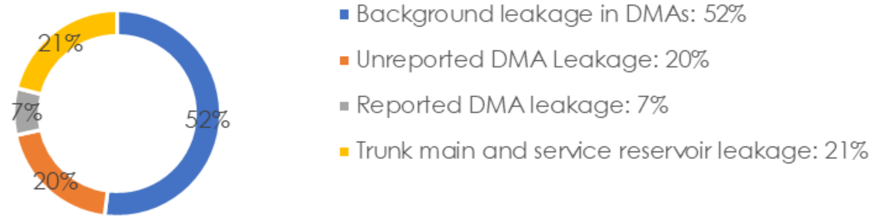
Results of Company-level analysis

These results are from company level reporting and show leakage components as a proportion of total. Three of the companies: Anglian, DCWW and Severn Trent have similar levels of background leakage, but Portsmouth has much higher background leakage as a proportion of the total and Affinity has slightly lower background leakage. The Portsmouth result can be explained by the fact that night-flow leakage reporting is carried out for the whole company, with no upstream losses reported separately. Therefore the "background" component actually includes what would be upstream (or "trunk main and service reservoir") losses in other companies. Affinity background leakage has been estimated directly from minimum achieved values in DMAs and may not be directly comparable.

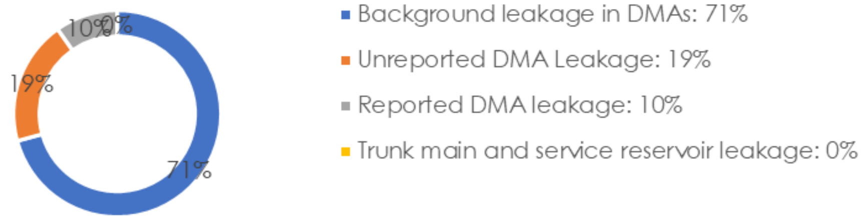
[redacted]



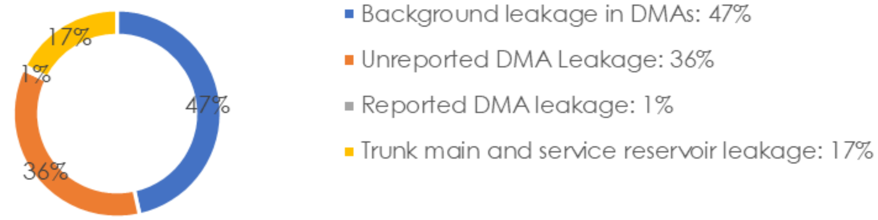
[redacted]



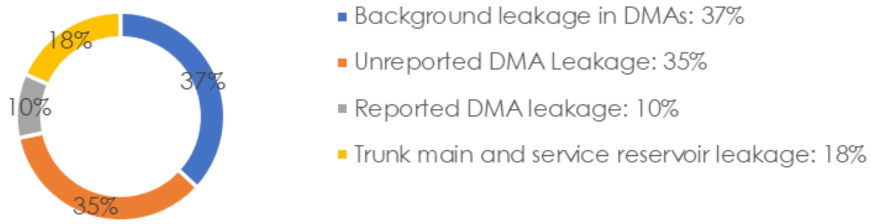
[redacted]



[redacted]



[redacted]



Appendix 3

The DMAs chosen for survey

Company	DMA Company code	Total property count	Length of main (km)	Min Night Flow leakage (l/hr)	Min Night Flow Leakage (l/prop/hr)
A	A_1	597	21.520	6806.5	11.40
	A_2	698	5.710	8419.3	12.06
	A_3	537	18.440	5115.8	9.53
	A_4	608	16.320	4214.5	6.93
	A_5	843	4.840	5593.5	6.64
B	B_6	601	8.027	4203.3	6.99
	B_7	742	11.405	5557.9	7.49
	B_8	677	10.089	5412.0	7.99
	B_9	683	11.520	5053.6	7.40
	B_10	736	26.898	12904.5	17.53
C	C_11	648	28.016	2832.0	4.37
	C_12	551	33.502	8171.3	14.83
	C_13	749	24.902	2909.2	3.88
	C_14	474	18.703	3548.9	7.49
	C_15	661	10.792	1583.4	2.40
D	D_16	857	4.378	4914.6	5.73
	D_17	792	6.379	3754.2	4.74
	D_18	876	5.365	4999.6	5.71
	D_19	828	4.765	6558.3	7.92
	D_20	899	5.532	12111.7	13.47
E	E_21	503	13.964	5262.4	10.46
	E_22	577	6.005	4491.9	7.78
	E_23	840	15.633	5159.5	6.14
	E_24	768	4.896	3958.5	5.15
	E_25	840	3.567	4016.9	4.78
Totals		17585	321.169	137553.4	
Averages		703.4	12.847	5502.1	7.82

Appendix 4

Proposed Modelling Approach for Background Leakage Detection and Localisation

Supplied by Prof Joby Boxall and Dr Richard Collins University of Sheffield

One dimensional network modelling provides a unique opportunity to simulate and hence understand the operation of complex drinking water distribution systems. The primary use of such models in practice is to confirm continuity of supply and to ensure minimum pressure criteria are met. This can be achieved with relatively low fidelity models, i.e. demand driven, demands grouped to nodes and assigned in a top-down manner (taking the DMA inlet pattern and distributing that as a function of numbers of users across the model), the model is then calibrated to a few selected points of measured pressure data. Leakage can be assigned to these models, however the pattern (its size and spatial distribution and either constant or greater at night) is usually assumed to be uniform across a DMA. As the effects of Background Leakage are by definition discrete and small, this top down, aggregated, modelling approach will not facilitate the accurate detection or localisation of background leakage.

We will instead undertake the ultimate bottom-up approach to model building and calibration. We will use the unique high fidelity customer demand data for this project to accurately simulate (spatially and temporally) each and every demand. Differences to DMA inlet (a mass balance difference for the DMA, which we will also perform outside of the model) will then enable an estimate of the total background leakage (+/- accuracy and uncertainties, will also be rigorously explored). We will then undertake a multi-parameter optimisation to allocate this leakage across the model, calibrating the model to the pressure data collected at each and every hydrant within the DMA. This approach will require optimisation of both the leakage (size, number and location) parameters and other model uncertainties (primarily pipe resistances). It should be noted that even with the unprecedented data of this project, the solution space will be vast and probably non-unique hence outputs will be presented as best estimates or approximations, with listing and ranking of potential solution scenarios. The modelling will also consider and include the pressure dependencies of demand. We will perform modelling with both standard lumped modelling and 'all connection' (i.e. every household connection simulated as a node) models to explore how this impacts the estimate of leakage. Later stages of modelling will then explore how degrading or reducing the input data set impacts the estimate of background leakage provided.

In addition to the extended time series modelling of the 1D networks, the Background Leakage will be assessed using the transient pressures and acoustic measurements collected from the network. Transient based leak detection, using time domain reflectometry and frequency domain damping of naturally occurring fluctuations in pressure of the system will be deployed to explore how the detection and localisation of leaks can be improved with this high resolution data. Noise correlation methods will be used with the high spatial density of acoustic loggers to further pinpoint leakage. These approaches will utilise the same bottom-up hydraulic network models, but run with different solution techniques.

Appendix 5

Anonymising Water Networks for Open Access Publication

Supplied by Dr Richard Collins University of Sheffield

Introduction

As part of the managing background leakage project there is a requirement to publish all the collected data and scripts. For this to be useful for future researchers this needs to be accompanied by network hydraulic data from the associated real systems. There are concerns with publishing network data as it may contain individuals data and that there may be issues if it is identifiable. As a result any published network data needs to be anonymised and not identifiable to a specific real world location, but retain the hydraulic properties to make it useful. Here I am proposing that the EPANET .inp file format will be the most ideal format to use for this process.

EPANET .inp File Format

The original EPANET was created in the 1990's as one of the first widely available network hydraulic solvers. Since then it has been updated relatively frequently with the latest version being released in 2020 (EPANET 2.2). EPANET was created by the US EPA and since 2014 been maintained by the Open Water Analytics group (https://github.com/Open-Web-Analytics), EPANET has always been freely available and for most of its history has been open source. EPANET is widely used around the world by water utilities as their primary network hydraulic tool, in addition most other commercial network hydraulic software either directly uses the EPANET hydraulic engine or have produced reworked versions of the code. All commercial network software is able to open or import EPANET .inp files, as a result the .inp file has become the de-facto method of transferring data between different software solvers.

The .inp file format is a human readable text file that contains information about the network nodes, links and hydraulic objects (pumps, valves, controls, demand patterns etc.) and the configuration of the hydraulic (and water quality) solver.

The EPANET format has a couple of key features that will allow good anonymisation of the network data. The primary one is that the network hydraulics are computed using the link (pipe, valve, pump etc.) properties which are stored directly in the link elements, for example pipes objects have a recorded specific length which is used for calculation of headlosses. Nodes have recorded positions, however this is only used for visualisation of the network. A pipe lengths is independent the distance between the nodes to which it is attached. As a result the network hydraulics can be preserved by maintaining the link properties, but we can move the node positions arbitrarily to “disguise” the network and ensure that it cannot be mapped back to the “real world location”.

In addition the .inp file allows for the storing of additional link and node information in a way that will ensure that it can still be opened by original software, but can the additional information can be accessed by any scripts that are written as part of the project.

It should be noted that exporting networks from modern commercial network software to EPANET format does sometimes loose some information (usually about complex control options not available in EPANET solver), and we will need to verify that the EPANET export representation of the network is appropriate for the 25 DMAs used in this project.

Example

Original Network

As an example I will demonstrate the network topology anonymisation on a real world network. The network was provided to the University of Sheffield from Yorkshire Water as part of a previous research project.

Network Layout

In Figure 1 we can see the original layout of the network, with the pipes closely following the real world road layout, even without the nodal coordinates directly relating to the real world location (Figure 2) it would be possible with suitable effort to identify this network.

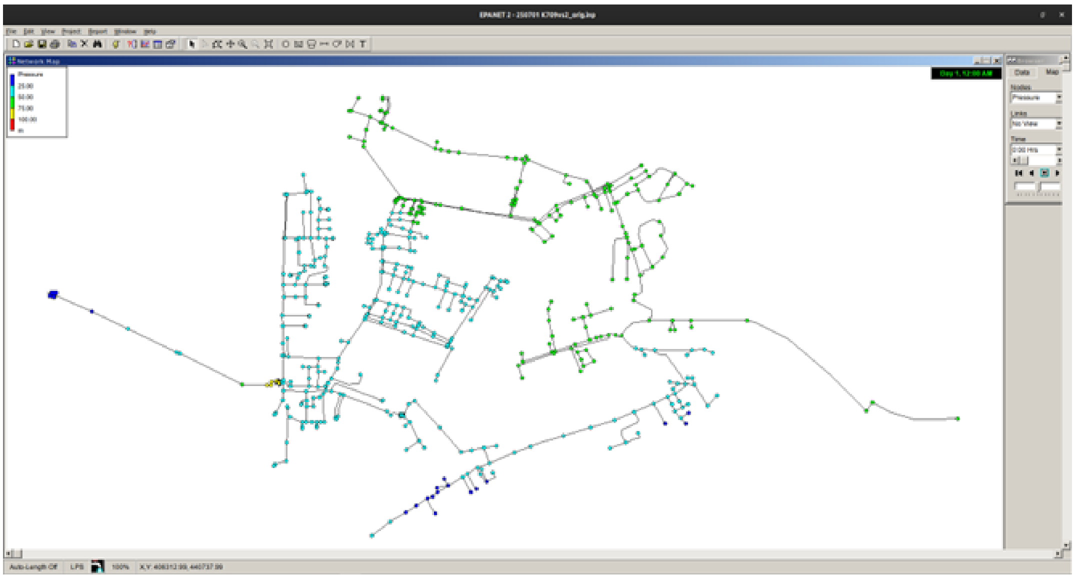


Figure 1: Original Network Layout

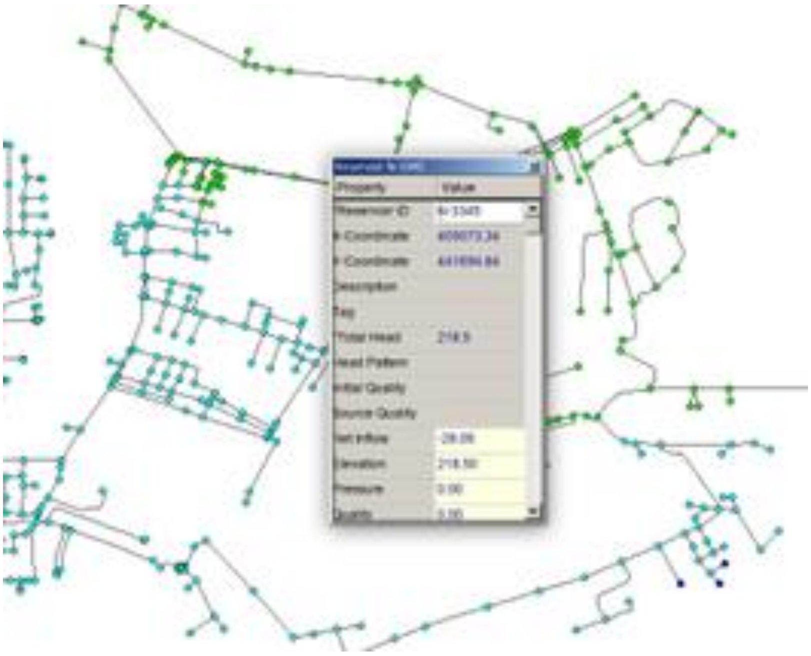


Figure 2: Coordinates of source reservoir in original network

Network Results

To ensure that hydraulic solution is unaffected by the network anonymisation in Figure 3 we can see the pressure profile of the node on the far right of the system A5383

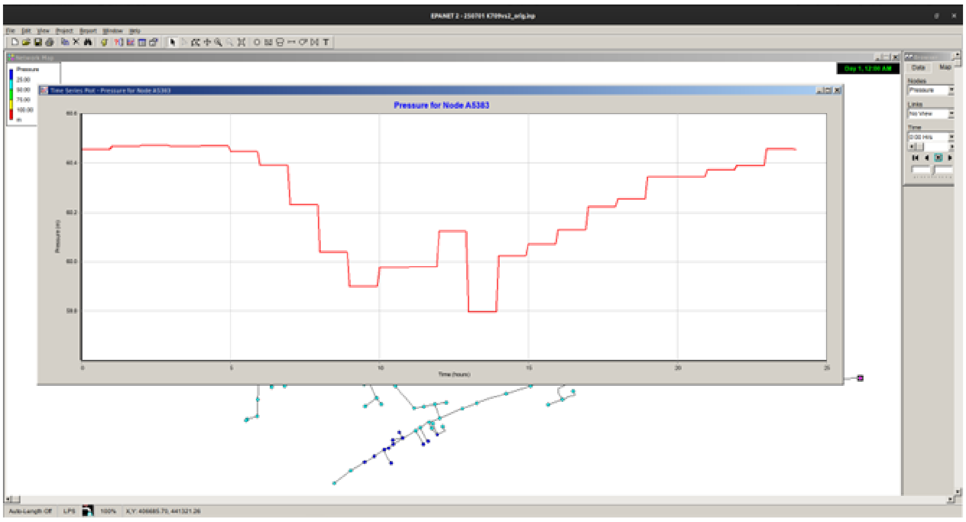


Figure 3: Pressure profile from Node A5383 in original network

Anonymised Network

A graph "morphing" algorithm was applied to the nodal locations to modify them in a way that, as close as possible, maintains the visualisation of the pipe lengths and any crossing points, but puts the nodes in new positions that no longer have any relation to their real world locations, Figure 4. In addition the network coordinates are now centred around an arbitrary point (here chosen to be 0,0), Figure 5.

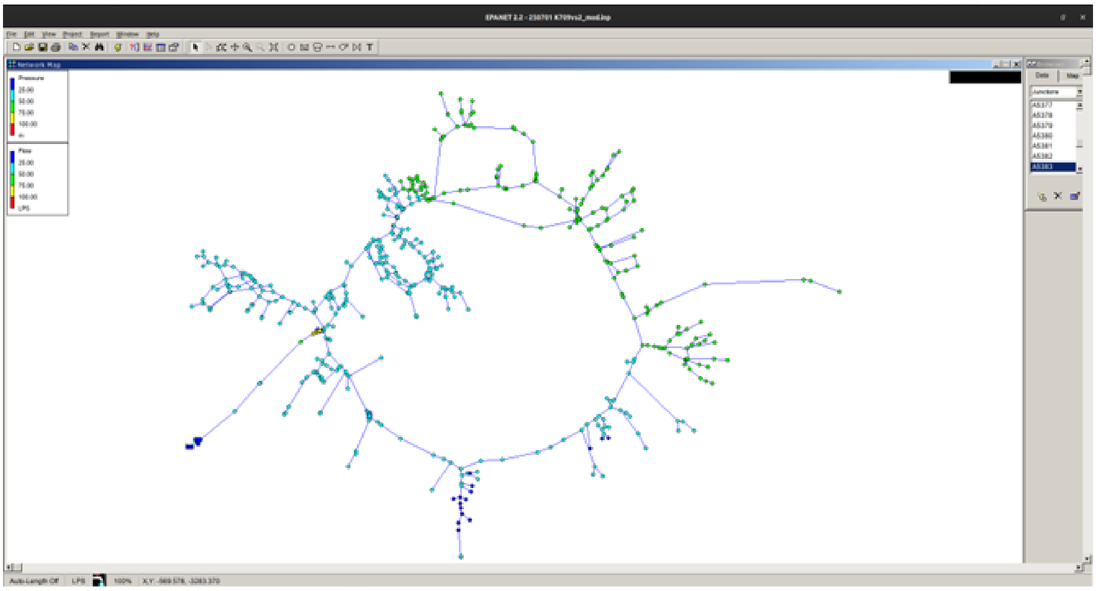


Figure 4: Anonymised Network Layout

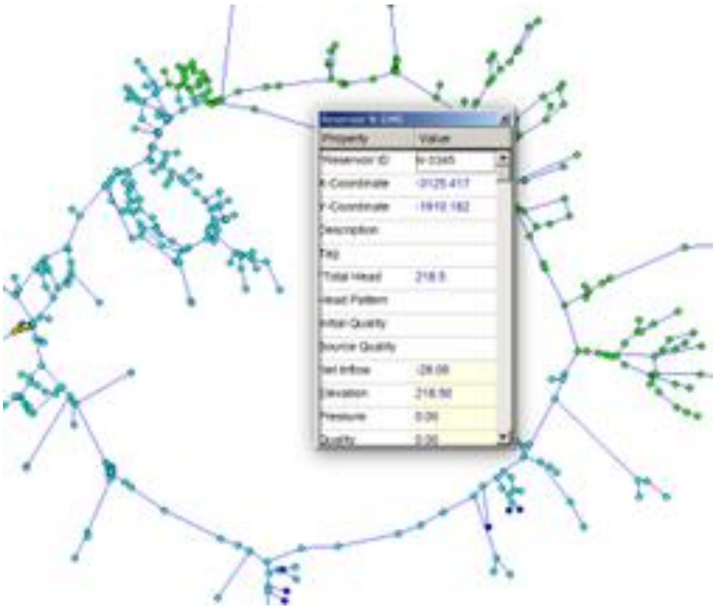


Figure 5: Morphed nodal position of source reservoir

Network Results

To confirm that the network hydraulics are unaffected by the anonymisation compare figure 6 with Figure 3 for the node at the right hand side of the network A5383.

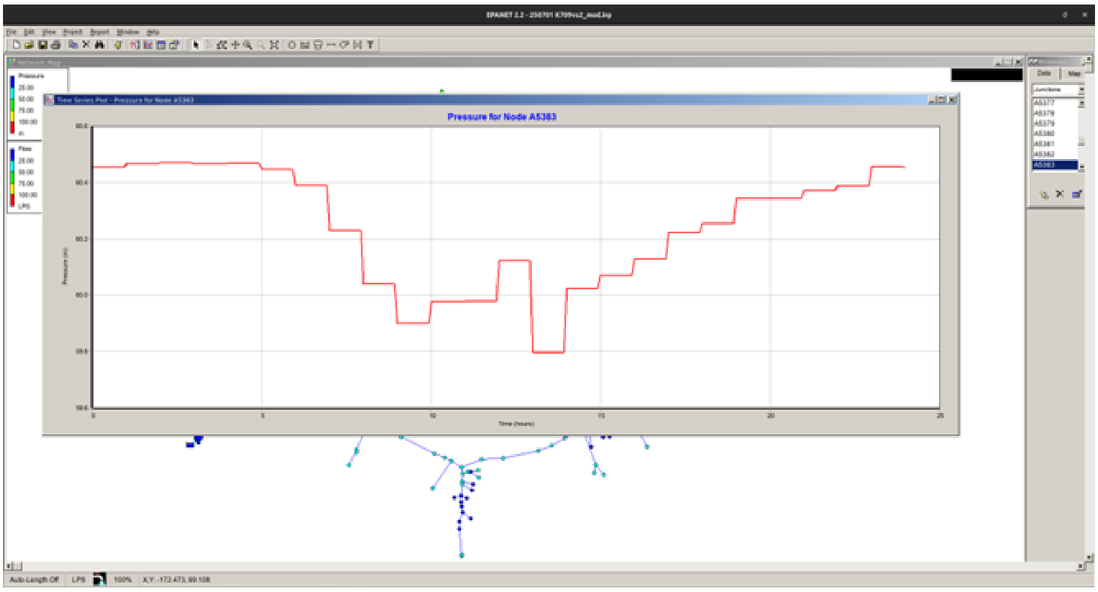


Figure 6: Anonymised *n*etwork *p*ressure profile for Node A5383

Mapping

In addition to the output anonymised network file this process also produces a “mapping” file that relates the original to anonymised locations. This file will be kept internal for the project but will allow us to add other data (customer meter data, pressure logging locations etc.) in a way that will ensure that it can be related to the correct points of the anonymised network.

Conclusions

This short document has demonstrated that it is possible to suitably anonymise a network topology to ensure that any open access published data is not able to be related back to the real world location. It should be noted that the above has not converted node or link names to allow for easy comparison between original and anonymised networks, but this process is easy undertake.

This document also proposes that the EPANET .inp file format should be used for open access publication as it is an open access format that is widely used by the water industry and is able to be extend depending on the requirements of this project.