THE CASE FOR FAST-TRACKING SERVICE CONNECTION RENEWALS TO REDUCE WATER LOSS

Julian Fyfe (formerly of Wellington Water), Cedric Papion (Stantec)

ABSTRACT

Renewal of assets approaching the end of their serviceable life is one of the four pillars of leakage management. However the pace of renewals can often be too slow to prevent increases in the rate of rise in leaks and bursts, particularly on a deteriorating network. In the Wellington Region, the aging drinking water network is exhibiting high and growing water losses as leaks and bursts form faster than they can be repaired. More than 75% of the observed faults have been found to occur on service connection valves and pipes. Under the current renewals model, service connections (and ridermains) are typically scheduled for replacement when the supporting water main is due for renewal.

Recognising that multiple repairs of a single service connection can quickly amount to the cost of replacement, Wellington Water commissioned Stantec to find a cost-effective approach to renewing service connections (reactive, proactive or both) outside of the BAU mains renewal model. The aim is to test the theory that dedicated service connection renewals could reduce losses and slow the rate of rise in leak formation in the short term and avoid or defer significant investment over the longer term. This paper presents the analyses related to identifying target service connections for proactive renewals, quantifying the water loss amelioration benefits, and the comparative economics of different renewal approaches.

The service connection leak history was first analysed against several factors known to contribute to leak formation such as age, pressure and material. Several machine learning algorithms were used to model the likelihood of leak for each connection. While the predictive power of the preferred model is insufficient to be meaningful at the level of individual connections, it provides a basis to determine geographical priorities and compare scenarios.

The model also allows the identification of areas where service connections renewals are the most likely to lead to a reduction in leak rate. This may be used to scope a programme of capital works dedicated to proactive service connection renewals. Unlike watermain renewals, service connection replacements are relatively straightforward and low-risk in that they cause minimal service disruption, involve shallow excavation, do not require heavy or expensive plant and equipment and have a limited footprint within the road carriageway. In theory, it should possible for dedicated crews to undertake the works efficiently at scale, particularly if unusually complex sites are omitted from scope.

The model was then used to quantify potential leakage reduction benefits and simulate various investment scenarios. The first considered reactively replacing faulty service connections instead of repairing them. Other scenarios consisted of combinations of reactive renewals and various levels of investment in proactive renewals.

The study's findings will allow Wellington Water to benchmark service connection renewals against leakage management options and inform the organisation's water demand reduction strategy.

KEYWORDS

Water supply, leakage, demand reduction, service connections, reactive renewals, proactive renewals

INTRODUCTION

Water loss is a pervasive problem for many water utilities across New Zealand that undermines water security and erodes Te Mana O Te Wai. In the Wellington metropolitan area (encompassing Wellington, Hutt, Upper Hutt and Porirua City Councils), historic under-investment, ageing infrastructure, high water reticulation pressure, limited customer metering and seismic activity are amongst a host of factors that have contributed to losses exceeding 40% of water supplied[1]. The increased demand on the water supply is compromising the resilience of the system in relation to both source water availability and maintaining adequate system headroom. Affecting the entirety of the distribution network, this puts the supply scheme at risk of shortfalls during peak demand periods and drought. It also makes the task of reducing the water take from the environment more difficult, which counters giving effect to Te Mana O Te Wai and complicates re-consenting of existing water takes and consenting of new supply-side measures.

Wellington Water, through its Sustainable Water Supply and Demand strategic programme, has characterised the nature and extent of water loss in the region, in particular from leaks on the public network, and has developed a water loss reduction plan to bring it under control. The plan, based on the four pillars of leakage management, includes renewals of specific assets known to be contributing significantly to network leakage – service connections.

SERVICE CONNECTION LEAKAGE

Leaks are identified either after being reported by the public or through active leak detection. Leak repair job records indicate that up to 85% of leaks occur on service connections in the Wellington metropolitan area. These may include service pipe breaks, service valve failures or leaking/failed joints. There is no standout causal factor behind service connection leaks, although common factors include age, quality and condition of materials and fittings, and workmanship.

On account of their leakage rate and relatively low risk to public safety and other infrastructure, service connection leaks tend to be de-prioritised in the triaging process, which results in them often persisting for extended periods of time. In Wellington, the median run time from reporting to repair for service connection leaks is around 30 days. The time a service connection will have been running before it is reported is unknown, but is likely to vary significantly depending on whether the leak surfaces, whether it is in a highly trafficked location and how large and disruptive the flow is.

Whilst a typical service connection leak flowrate may only be in the order of 0.25 to 0.5 L/second, the sheer number of leaks means that they contribute a significant proportion of overall network leakage flows. Figure 1 gives a breakdown of leak repair jobs count and

---

[1] https://www.wellingtonwater.co.nz/resources/topic/water-conservation/leaks

volume by asset type, with service connections comprising both service pipe and toby assets. The total contribution of service connection leaks is estimated to be as high as 75% of all leaks repaired.

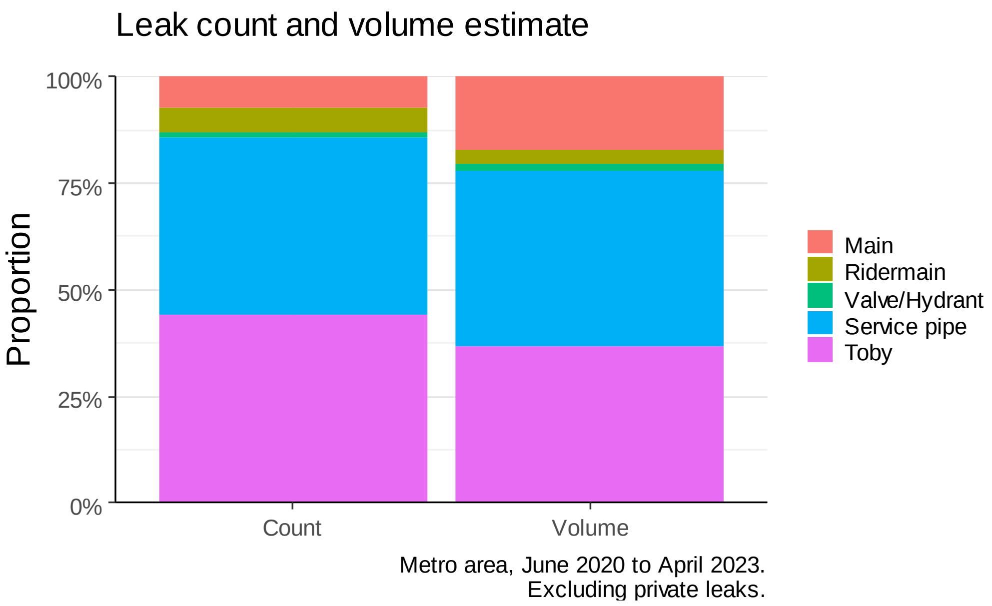


Figure 1 Proportion of leak jobs and estimated corresponding water loss by asset type.

STUDY OBJECTIVE AND SCOPE

In recognition of the scale of water loss occurring on service connections, Wellington Water initiated a study to assess the potential of service connection renewals to reduce demand in the Wellington metropolitan area (Stantec, 2023). The objective of the study was to characterise the water savings from different approaches to service connection renewals and produce cost estimates to allow comparison with other demand management strategies. The study attempted to quantify the benefits of different forms of service connection renewal interventions and to identify geographic clusters of candidates for service connections renewals. At the heart of the study is a numerical model, that estimates the likelihood of leakage from service connections based on various predictors such as age and pressure. In the first instance, the scope was limited to Hutt and Upper Hutt cities, but the rest of the Metro area (Porirua and Wellington cities) was considered where practical.

This paper focuses on the modelling methodology followed for the study, and how the analysis can be implemented in practice. It then presents a brief overview of the findings.

MODELLING APPROACH

Leak repair data was joined with GIS information to develop a combined dataset of predictors and leakage rate. This allowed to fit and evaluate a linear model that predicts the likelihood of leaks forming on any given service connections. The model outputs were combined with estimates of leak flow and intervention costs to assess different renewal investment scenarios. All data manipulation, analysis and visualization were undertaken using the programming language R.

References

Research specifically on service connection leaks is very rare (Gouveia, 2022); the closest field of study is the modelling of distribution main leaks which has been abundantly covered

(Vega, 2023), but this is not directly applicable to service connections. The smaller size and different material types make service connections vulnerable to different failure mechanisms compared to distribution mains. Additionally, the data available can be less complete and reliable.

Data sources

Repair and renewal costs were estimated using Wellington Water's maintenance work order database (Maximo) and through conversations with Operations staff. Pipe data was extracted from Wellington Water's GIS and augmented by pressure information calculated by hydraulic models. Soil conditions were downloaded from the GNS Science website.

Leaks were identified from Maximo. This has been reliably populated since June 2020, so the export dated April 2023 covered 34 months of data.

It is important to understand how the leaks used for the analysis get recorded in Maximo. Leaks appear through defects at joints, within the pipe wall or on fittings such as valves. The leak flow rate can increase over time and eventually the leak can be identified through specialist detection activities or simply through the naked eye and reported by the public. Many leaks, especially small ones, do not get detected/reported for many years. Once identified, the leak will be assessed and triaged by Wellington Water Operations, and eventually repaired and recorded in the Maximo completed work order register.

Since the leak detection activities target specific zones they can introduce a bias in the apparent leak frequency; consequently, detected leaks were not considered in this study. Customer-side leaks are not repaired by Wellington Water unless by agreement with property owners. This is relatively rare and private leaks were not also considered in this study.

What are we trying to predict?

To understand the relationship between leak rate and local conditions, it is essential to know which service connection the leak was on. If a leak record in Maximo could not be joined to a specific asset, it was not considered in the modelling. The study therefore attempts to model the reported public-side service connection leaks which could be joined to an asset in GIS. Valve leaks (in this case tobies) were distinguished from pipe leaks (Figure 2).

Importantly, the models trialed predict an average number of leaks, not a rate of leaks per km, as would be standard practice for leaks on distribution mains.

Data manipulation

The Maximo database had missing and at times contradictory information which required clean-up and processing, particularly to determine if the leak occurred on a service connection, rider main or distribution main, and which was the corresponding asset in GIS. This was done using (if possible) the Maximo pipe/valve asset ID, the Maximo address, the free text description of the maintenance work. All valve service connection leaks were joined to the closest service connection pipe to enable further analysis. In the rest of this paper, the term "pipe" is intended as "service connection pipe".

Is this a classification or a regression problem?

The relationship between leaks and predictors may be approached either as a regression problem (how many leaks per year may we expect for a service connection?) or a classification problem (what is the probability of a given service connection having at least one leak in a given period?). The classification approach was found to require a cumbersome preparation and to provide a less useful outcome than the regression approach.

This problem could also have been treated as a survival analysis (how long until the next failure?). The data is both left- and right-censored: we do not have failure history before 2020 and not all pipes have failed. This type of analysis is complex (Mailhot et al, 2000) and because of the short data period available, it is not obvious it would be successful; it was therefore not considered as part of this study.

A distinction between service connections and distribution mains is that the leak frequency does not increase at the same rate as the pipe length. This is because toby leaks are more frequent than service pipe wall failures, and there is generally only one valve per service connection, regardless of its length. The predicted variable is therefore the number of leaks over the data period for each service connection.

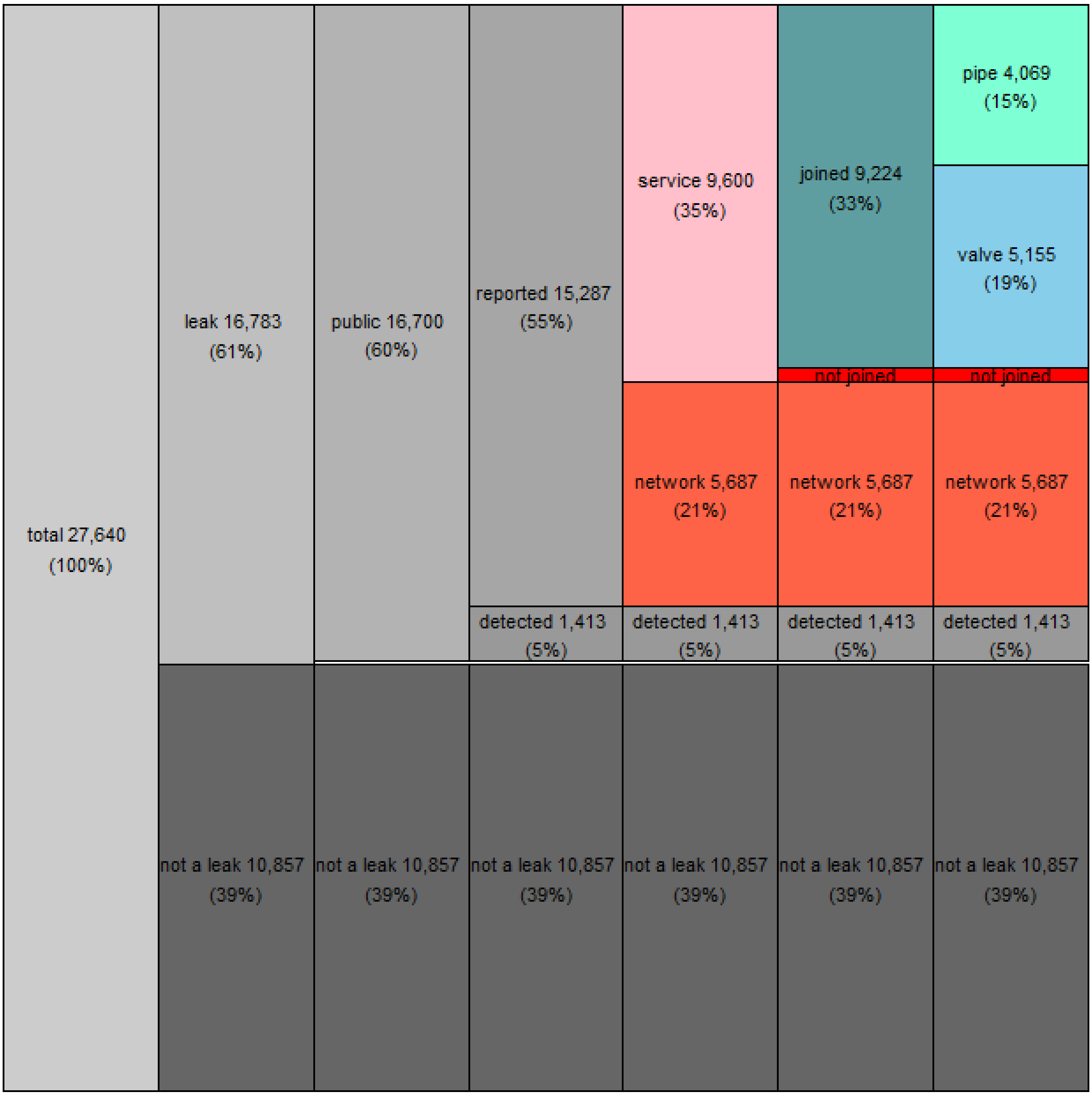


Figure 2: Breakdown of Maximo job records (metro area) over 34 months

Model family

As we want to estimate a count of events, a Poisson regression was applied. A Poisson regression is a linear model where the outcome is a rate of occurrence of a certain event (here a reported leak) over a certain period. The outcome is always zero or positive. Other machine learning approaches such as random forests and gradient boosted trees were trialed, but they were found to provide no significant improvement over the linear regressions in this case.

The vast majority of service connections have no leak on record over the data period. Most of the connections with a leak record had only one leak, and very few had two or more. To reflect this, a two-stage modelling approach could have been undertaken:

1. what is the likelihood of at least one leak over the data period?
2. if there is at least one leak, how many leaks?

These are known as zero-inflated models (Zuur et al, 2021) on account of the over-representation of zeros in the observed data. This approach would not change the predicted average rate of leaks for each pipe, but it would describe better how the actual leak count is distributed around this mean. It was found, however, that using zero-inflated distributions made no discernible difference to the modelled number of leaks predicted at the geographical scale of the study. The added complexity was therefore found to not be warranted.

Predictors

The study confirmed the well-documented positive correlation between leak rate and maximum pressure. The relationship between age and leak rate is not linear; to reflect this, the age predictor was used with an "age_15" interaction term. In essence, two coefficients were fitted to the age term, depending on whether the pipe is less or more than 15 years old.

Pressure range and diameter were not found to be significant predictors. While pipe length was found to be a significant predictor for the rate of pipe leak, on balance the added predictive power was not found to be sufficient to include the variable as a predictor in this case.

Pipe material is suspected to be a useful predictor in theory, but this could not be used in this study: the material information is generally missing for Upper Hutt and is almost exclusively recorded as PE for Hutt City. Similarly, soil type was tested as a predictor in line with other research in this field, but was not found to be a statistically significant predictor in the case of the Wellington region.

Leak history is particularly interesting and is a strong predictor of future leaks. A service connection that leaked in the first year of data has 3 to 4 times more chances to leak over the rest of the data period than a service connection with no leak history. There are, however, four issues with using this predictor.

The data period is relatively short and using leak history necessitates dividing it into even smaller parts: leak history on the first part, leak observation on the second part. From a numerical modelling perspective, it makes predictions harder to evaluate. From a user perspective, it makes the predictions more arbitrary because they come from such a short period of data. This will become less of an issue as more data becomes available in the future.

Moreover, leak history is strongly correlated to other predictors (e.g. age and pressure). The causes behind service past connection failures are often the same as those behind later failures. By using leak history, we are effectively double-counting those other predictors. Is it best practice to avoid such highly correlated predictors.

More importantly, using leak history does not lead to a practical outcome for large-scale planning of proactive renewals of service connections. There is a high degree of randomness as to where leaks occur. Including leak history as a predictor leads to high priority connections being highly dispersed geographically (essentially anywhere a leak has been recorded). Omitting leak history leads to much more clustered priority connections, which can be bundled into a practical programme of works.

The study includes an estimate of how many leaks may occur over an extended period of time under various renewal scenarios. It would be misguided to assume that the same connections that leaked in 2020-2022, and only these, will keep on failing in perpetuity.

The final predictor used is the locality the service connection belongs to. The probability of service connection leak may be influenced by the construction workmanship, the local standards in place at the time, where and what type of materials were sourced, or the thoroughness of the compaction - none of which can be known today but are aggregated in the “council” data field which was found to be very significant.

Several model specifications were trialed, using various combinations of predictors and interactions.

Model evaluation

Models were evaluated based on their ability to predict the number of pipe/valve leaks amongst unseen data from subsamples representing 10% of the population (10-fold cross validation). The models were found to be indistinguishable from each other using this approach. Indeed, the ability of a model to predict the total count of service connection leaks depends on the scale at which it is used: any model can predict accurately at city-level but none can reliably predict at pipe-level. Figure 3 suggests that, for a random 2% of service connections, the model predicts the correct number of leaks approximately +/- 5% in Hutt City and +/- 15% in Upper Hutt.

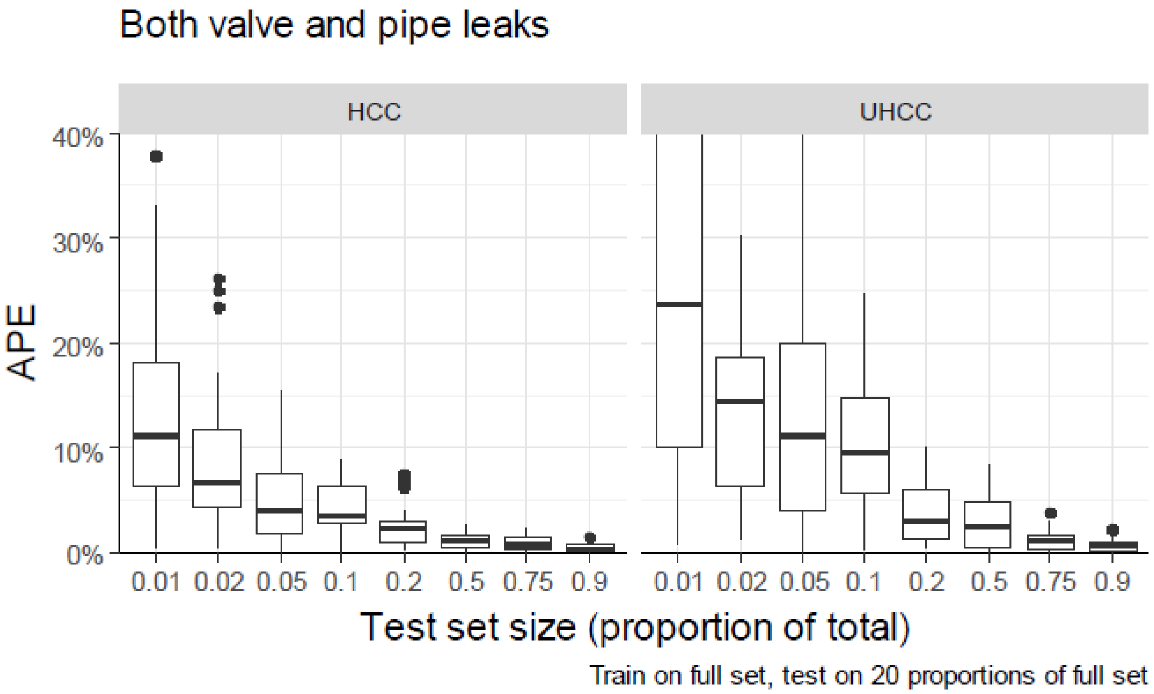


Figure 3: Diminishing model accuracy (Absolute Percentage Error for total leaks in the test set)

Models were also evaluated on their ability to replicate the response to key predictors, in particular the difference between young and old pipes – this is important for the renewal scenarios developed from the models. This was done both with unseen data and with the full dataset (Figure 4).

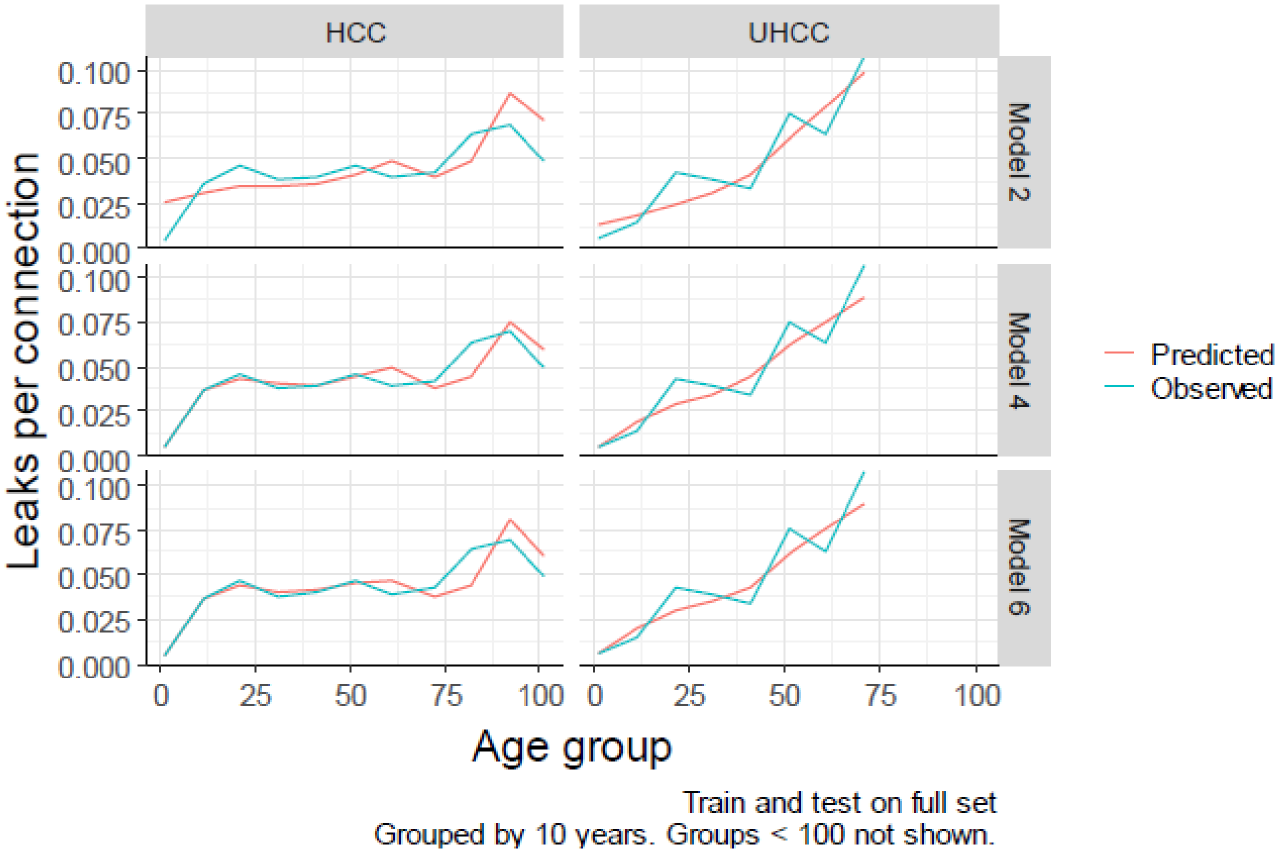
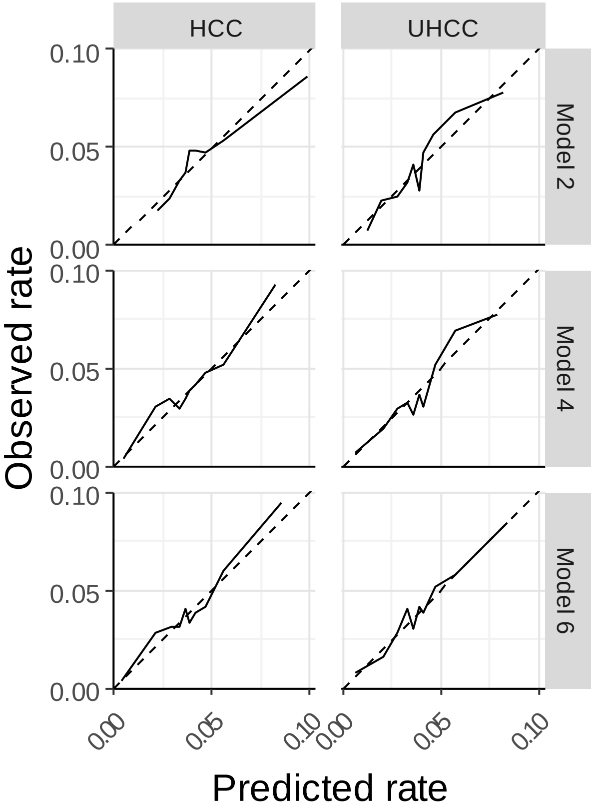


Figure 4: Example of model evaluation (by predictor)

Another useful approach to select models was to compare the mean predicted rate vs observed, for pipes grouped by quantile of predicted leak rate (Figure 5).



Rate in leaks per connection over the data period.
Train and test on full set.
Grouped by quantile. Groups < 100 not shown.

Figure 5: Example of model evaluation (by quantile)

Models retained

Four models were retained in total - one each for predicting leaks on pipes and on tobies in the two city council areas (Hutt City and Upper Hutt). All are Poisson linear regressions with the following predictors:

- council,
- maximum pressure,
- age,
- whether the connection is more than 15 years old.

Fitted model coefficients are provided in Appendix 1.

APPLICATIONS

Using the model formulation, we can assign a likelihood of failure to each service connection in the network. The consequences of service connection leaks do not vary hugely based on the local conditions: the repair cost and disruption will be higher in more heavily urbanised locations but the variability will be less than for distribution mains. Other aspects such as water lost and reputational risk for the utility can be considered constant across service connections. This lends itself to the proposition that the likelihood of failure is a good first approximation of the overall risk pertaining to service connection failure.

Clusters of high risks

The calculated leak likelihood for a given locality is based on age and pressure. All the service connections of a given street are often of a similar age and share similar pressure, and the calculated leak likelihood is therefore spatially clustered, as evident in Figure 6. This can help plan a programme of proactive service connection renewals by targeting areas where renewals are the most likely to avoid future leaks.

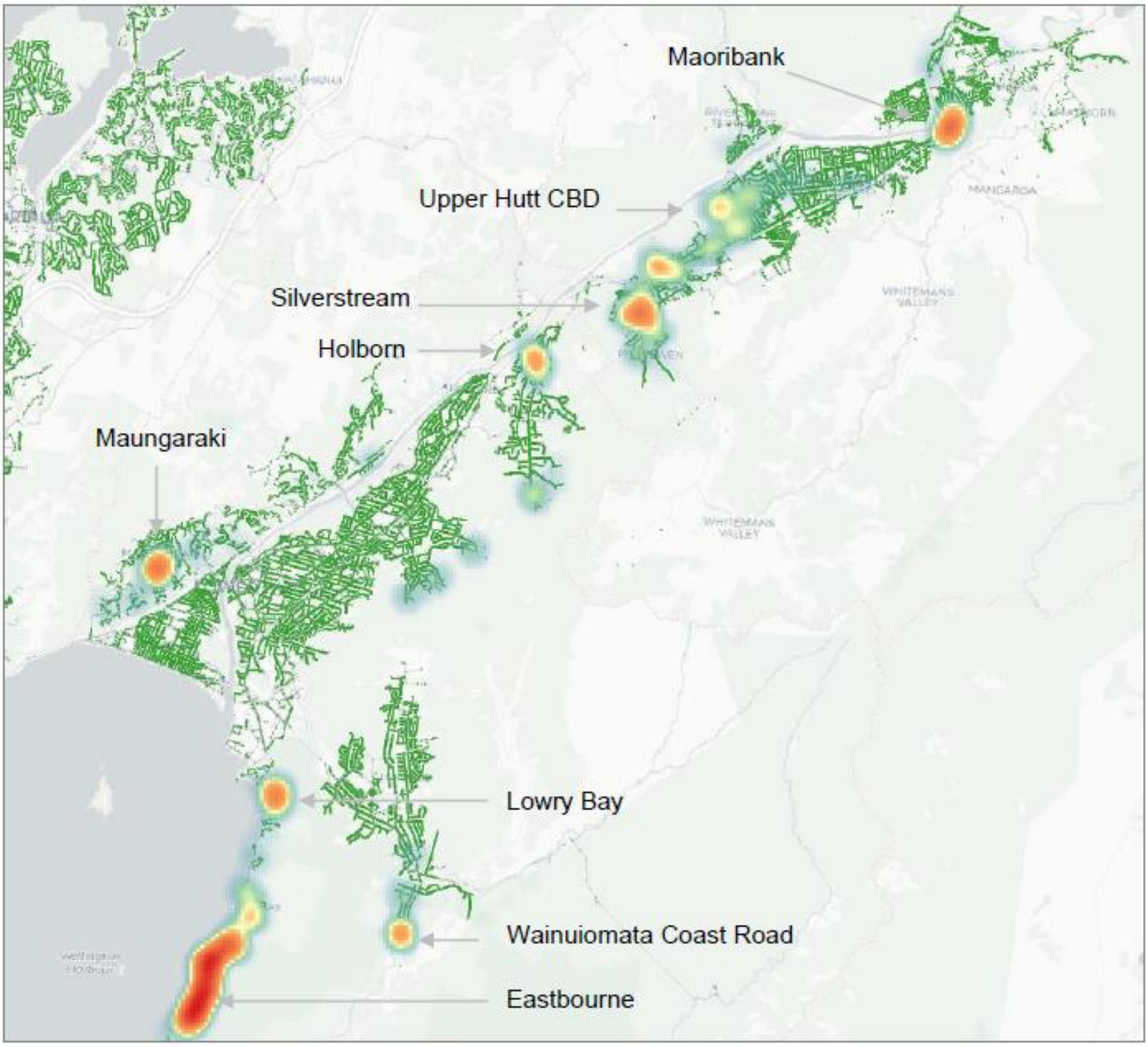


Figure 6: Top 5% of service connection pipes by estimated renewal benefit (arbitrary color scale)

Investment scenarios

The quantitative leak likelihood estimate can also support the testing of several asset management scenarios.

At year 0, each service connection can be simulated to leak a certain number of times, based on its calculated likelihood and the statistical distribution from the model. Depending on the scenario, this leak may be assumed to result in a repair or a reactive replacement. Depending on the scenario, the top X% of high-risk connections may be assumed to be proactively replaced. This process can be repeated over the course of the scenario, in this case 15 years.

While the median time between a leak report and repair is around 30 days, this does not include the period when the leak may have run undetected, and the estimate is not based on open leaks, some of which have run for extended periods of time. For lack of a better estimate, the actual average service connection leak runtime was assumed to be 60 days.

Wellington Water has installed Small Area Monitors (SAM) in recent years. SAMs are discrete, contained residential parts of the network where all consumption can be recorded by a single flow meter located on the incoming main or ridermain. These provide controlled environments to study water use, particularly night flow. A number of leaks have been identified and repaired in the SAMs, allowing comparison of night flows before and after the repair. The resulting leak flow rate estimates averaged 0.25 l/s per leak, with a high degree of variability. The SAM leaks are not, however, representative of the broader network as they are predominantly located in suburban areas, and were fixed relatively quickly (to maintain the integrity of consumption data monitoring). As leaks develop over time, the leakage flow often increases. Hence for this study, the average leak flow was considered to be double that of the SAMs, or 0.5 l/s.

The scenarios tested were:

1. Business as usual - current proactive replacement rate from mains renewals, no reactive pipe replacement, no reactive valve replacement (which is a simplification as some reactive valve replacement does occur, leading to a slight over-estimation of future leaks under BAU).
2. Each valve leak leads to a reactive valve replacement instead of a repair.
3. Each valve leak leads to a reactive valve replacement instead of a repair; each pipe leak leads to a full replacement of both the pipe and the valve.
4. Same as [2] plus proactive replacement of 2% of the service connection valves each year.
5. Same as [3] plus proactive replacement of 2% of the service connection valves each year.
6. Same as [3] plus proactive replacement of 2% of the full service connections each year.

The rate of 2% renewal per year was selected arbitrarily to demonstrate the process (Figure 7).

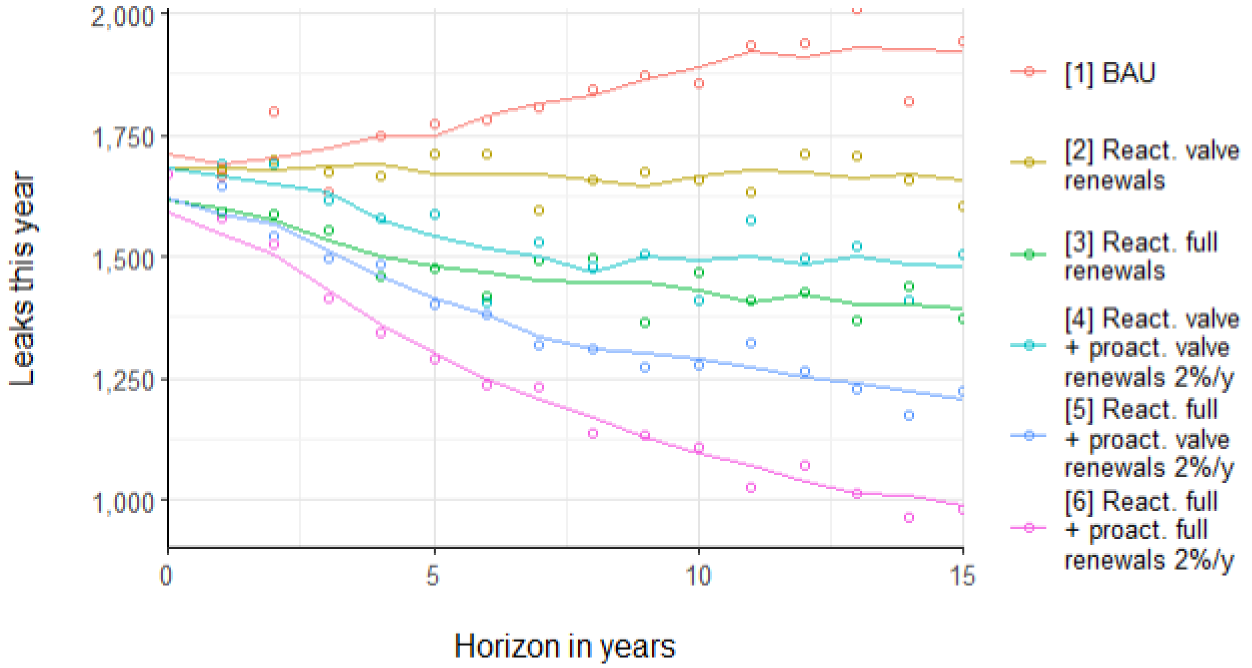


Figure 7: Estimated yearly service connection leaks

The number of leaks avoided can be converted into a volume of avoided water losses (Table 1).

Table 1: Estimated water loss savings of investment scenarios against business as usual

Scenario	**ML / year saving from BAU**[1]
[2] React. valve renewals	412
[3] React. full renewals	950
[4] React. valve + proact. valve renewals 2%/y	764
[5] React. full + proact. valve renewals 2%/y	1,225
[6] React. full + proact. full renewals 2%/y	1,619

[1] Average over 15 years

Combined with the estimated cost of the various interventions, this provided Wellington Water with a conservative high-level cost / benefit estimate for each strategy. The levelised cost of the best blend of reactive and proactive renewals is estimated to be in the order of \$0.70/m^3. Acknowledging a significant uncertainty around the cost estimates, this places service connection renewals at the lower end of cost for water supply/demand interventions (WSAA, 2022), suggesting they are an effective and economic form of demand management worth factoring into strategic water resource planning.

DISCUSSION

Service connection renewals are often relatively straightforward, low risk and generally require less traffic management compared to full watermain renewal projects. Packaging targeted service connection renewals for external contractors can free up constrained internal operations resources as well as design and delivery teams that could focus on more complex projects and other works. It therefore provides Wellington Water more options in how renewals are delivered and funded.

This analysis could be extended to include rider mains, which are a middle ground between service connections and water mains. Rider mains are more complex to renew, require more traffic management and service disruption, but including them enables renewal of greater lengths of potentially leaking pipework than service connections alone. Ridermain renewals also offer an opportunity for localised pressure reduction (Stantec, 2023), consisting of small diameter pressure reducing valves at the edge of the rider main. This setup has yet to be tested in practice but it has the potential to reduce water losses even further.

CONCLUSION

The analysis outlined in the paper demonstrates the value of implementing a dedicated service connection renewal programme together with traditional watermain asset renewals, highlighting the benefits of renewing service connection assets known to be leaking in addition to proactive service connection renewals in location considered to be at high risk of leaking in the future. It also shows how targeted renewals can play an effective role in reducing the incidence of service connection leaks, which not only reduces water loss, but frees up operation crews and contractors to focus on responding to larger, more serious network faults and the maintenance needs of larger and more critical assets.

This study provides the basis for an economic analysis of service connection renewals as a tool in leakage management and broader water resource management. It will help Wellington Water decide on the right mix of water mains-driven renewals versus targeted service connection renewals.

REFERENCES

Gouveia C, Soares A. (2022). *Machine learning classification models applied to water service connection leakage data: contributions on understanding factors involved in failure and insights for infrastructure management*. Environmental Sciences Proceedings. 21(1):83. https://doi.org/10.3390/environsciproc2022021083

GNS Science. *NZL GNS 1 250K geological units (scale 1-500k).* https://doi.org/10.21420/QF82-7D42

Mailhot, A., Pelletier, G., Noel, J-F., Villeneuve, J-P. (2022). *Modeling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: methodology and application*. Water Resources Research. https://espace.inrs.ca/id/eprint/9465/1/P979.pdf

Papion, C. (2023). *Proactive renewals of water supply service connections*. Stantec for Wellington Water.

Papion, C. (2022). *Localised pressure management – Stage 1*. Stantec for Wellington Water.

Vega, L. (2023) *Analysis and modelling of leaks in the Auckland water distribution system* [Doctoral dissertation, Auckland University].

Water Services Association Australia (2020). *All options on the table – Urban water supply options for Australia*.

Zuur, AF., Ieno, EN. (2021). *The World of Zero-Inflated Models*.

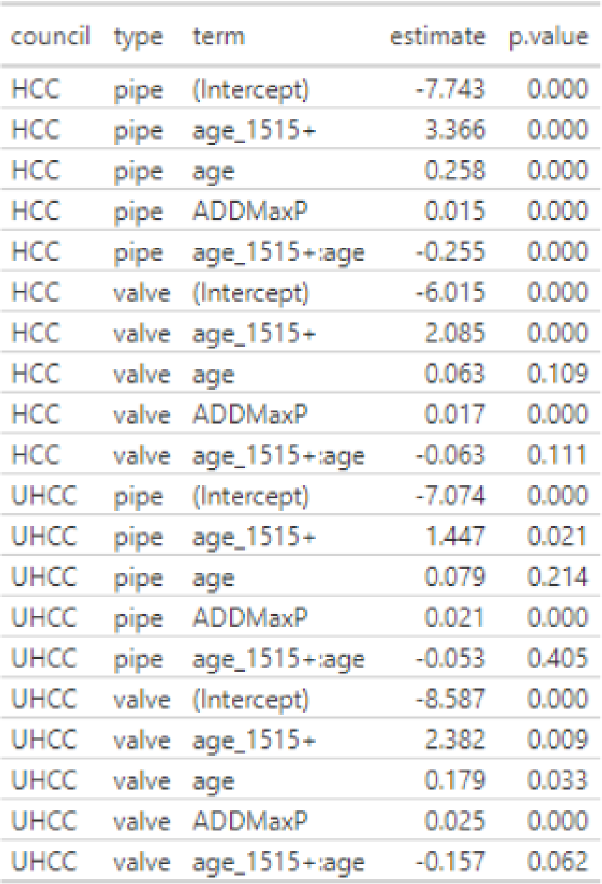
Appendix A: Fitted model coefficients

Both the pipe leaks and valve leaks models are in the form:

glm(nb_leaks ~ age_15 * age + ADDMaxP, family = quasipoisson)

The fitted coefficients for each model council are presented below.

Table A1: Fitted coefficients



council	type	term	estimate	p.value
HCC	pipe	(Intercept)	-7.743	0.000
HCC	pipe	age_1515+	3.366	0.000
HCC	pipe	age	0.258	0.000
HCC	pipe	ADDMaxP	0.015	0.000
HCC	pipe	age_1515+:age	-0.255	0.000
HCC	valve	(Intercept)	-6.015	0.000
HCC	valve	age_1515+	2.085	0.000
HCC	valve	age	0.063	0.109
HCC	valve	ADDMaxP	0.017	0.000
HCC	valve	age_1515+:age	-0.063	0.111
UHCC	pipe	(Intercept)	-7.074	0.000
UHCC	pipe	age_1515+	1.447	0.021
UHCC	pipe	age	0.079	0.214
UHCC	pipe	ADDMaxP	0.021	0.000
UHCC	pipe	age_1515+:age	-0.053	0.405
UHCC	valve	(Intercept)	-8.587	0.000
UHCC	valve	age_1515+	2.382	0.009
UHCC	valve	age	0.179	0.033
UHCC	valve	ADDMaxP	0.025	0.000
UHCC	valve	age_1515+:age	-0.157	0.062

The model uses a log transformation so the coefficients cannot be used directly. For HCC pipe leaks, as an example, this should be interpreted as:

- if age and maximum pressure are 0, the mean rate of leakage over the data period is $e^{-7.743}$ = 0.00043.
- a 1m increase in maximum pressure multiplies this mean by $e^{0.015}$.
- a 1 year increase in age multiplies this mean by $e^{0.258}$.
- if the pipe is 15+ years old, the mean is multiplied by $e^{3.366}$.
- if the pipe is 15+ years old, a 1year increase in age multiplies this mean by an additional $e^{-0.255}$.

Overall, the p.values reject the null hypothesis that these predictors have no effect on the leakage rate at 5% confidence level. There are predictors with large p-values such as the interaction of age with age being 15+ for UHCC pipes, suggesting that this particular situation does not improve the predictive power of the model. These predictors were considered to strike an acceptable balance of accuracy and simplicity and were therefore retained.