Deterioration modelling of small-diameter water pipes under limited data availability

Ben Ward[1&2], Andrew Selby[1], Simon Gee[1], Dragan Savic[2]

Abstract

High volume and low value buried infrastructure assets in the water distribution network are typically less well understood and often sub-optimally managed in comparison to more critical or higher value assets. This is despite attracting an estimated yearly expenditure from water utilities operating in the developed World in excess of £4.42 billion per annum.

To address this problem the authors have developed a comprehensive deterioration modelling framework founded on latest geospatial technologies and statistical analysis. The modelling framework is specifically applied to small diameter (25-50mm) water distribution assets, known as communication pipes, which connect individual customers to water distribution mains.

Reliability curves are developed from failure data provided by two UK based Water Companies that have captured specific communication pipe failure records since 2001. The deterioration modelling curves and supporting data are compared and contrasted to demonstrate the robustness of this modelling approach, which is shown to be capable of modelling failure rates to a high degree of accuracy. This was validated by comparing the predicted number of failures against three years of failure data not used in the model building process. The yearly failure counts were predicted to within +/-5% accuracy and the overall cumulative modelled failure count at the end of 2014 was predicted within 1%.

Key words

Deterioration Modelling, Water distribution, Asset Management

List of Notation

$f(T)$	is the probability of failure at time T
$F(T)$	is cumulative probability of failure at time T
$R(T)$	is the probability of an asset remaining in service at time T
T	is time in years
β	is the Weibull shape parameter
η	is the Weibull scale parameter in years
γ	is the Weibull location parameter

Introduction

Communication pipes are defined as the water carrying assets that lie between the water main and the boundary of the private property being supplied. If a stop-tap or water meter is fitted this usually represents the end of the communication pipe and beyond this point the pipework is referred to as the 'supply pipe', which, in the UK, is the responsibility of the property owner (Ofwat, 2014). By their very nature communication pipes are considered high volume and low value assets. With typically one asset for each property, the pipework is of small diameter (25 - 50mm) and the length of the asset lies somewhere between 2-3m or 6-9m depending on the nature of the connection to the distribution main. Shorter connections, which typically occur when the property is located on the same side of the road as the distribution main are known as "short-side" communication pipes. Where-as communication pipes that required a road crossing, due to the location of the distribution main to property, are known as "long-side" connections.

In contrast, distribution mains can serve hundreds of customers usually range in size between 50 and 300mm in diameter. These assets are typically laid within the road and have significantly higher failure costs, not only due to the more vigorous and challenging repair techniques but as a result of the number of customers impacted. It is, therefore, understandable that past research effort has centred around the proactive management of these larger assets as the economic and social costs of pipe failures continue to rise (Engelhardt et al., 2000). However, an analysis of water company replacement figures between 2002 to 2010 has revealed that on average over 130,000 communications pipes are replaced each year across England and Wales (Ofwat, 2010). An assumed replacement cost of £800 per communication pipe would yield a capital investment value of £104M. Water utilities in England and Wales are not obligated to report operational expenditure against individual asset groups, but it is suspected that the Operational expenditure associated with these assets is in the region of an additional 80-120% of the capital value; therefore generating a £200M total expenditure budget in England and Wales for communication pipes alone. Assuming equivalent investment levels for communication pipes in other developed regions of the World, approximately 17.8% of the global population, would yield an annualised total expenditure in excess of £4.42bn ($6.95bn) in these regions alone (United Nations, 2011). Therefore, when considered as a collective asset stock, the global investment in the more developed regions of the World is significant enough to warrant the use of a deterioration model to better understand and manage the performance of these assets in the future. This could be achieved by using the outputs from the deterioration analysis to proactively target the replacement of poorly performing assets, or, via an awareness of upcoming investment requirements and future failure rates.

Pipe deterioration modelling can take many forms. At the highest level, models are distinguished by their nature, either physical or statistical. Physical models, generally speaking, are built around the understanding of the underlying physical parameters that govern pipe failure. These require the acquisition of detailed information about the pipe(s) being modelled for accurate predictions to be established (Rajani and Kleiner, 2001). Currently, the collection of this data is costly and it is not widely available across the whole distribution network. Even when data is available, it is commonly limited to water mains of larger diameters which tend to be critical in nature. Therefore, the expenditure associated with the collection of this data is unjustifiable; particularly so for lower value assets like communication pipes.

An alternative solution is to use statistical based models which evaluate the relationship between water main condition and key pipe characteristics by the use of historical data, which is mined and statistically evaluated to find patterns that can then be used to formulate a deterioration modelling equation. The most cited and widely recognised publications at the forefront of statistical modelling in the water industry at the time were (Shamir and Howard, 1979), (Clark et al., 1982) and (O'Day et al., 1986). The early developments in these models took a linear form and a few fundamental draw backs have since been identified with linear based models. The main disadvantage is that these models were heavily reliant on the availability of sufficient data for each pipe class for the rate of deterioration to be established without the interference of third party influences that lead to failure. These third party causes are commonly referred to as 'noise' because of their undesired influence on the identification of failure rates. Another important and restricting factor is that these models did not include data from assets where failure had not yet occurred, even though in some instances failure was imminent.

In the early 1990's semi-Markovian models were used to describe the deterioration process (Li and Haimes, 1992). The semi-Markovian model is a simplification of the deterioration process by modelling the current condition of an asset in one of a number of states. A probability is then applied to each asset to account for the likelihood of it moving into another state over a given time period. These types of model are the more commonly applied where condition information, obtained through inspections conducted over-time, is available. For this reason, semi-Markov deterioration models developed for sewerage assets, where condition inspection information is plentiful, have been found to provide an accurate representation of the assets deteriorating behaviour, (Ward et al., 2014).

Although less well published, Artificial Neural Networks (ANN) have also been applied to distribution main networks with the aim of learning the pipe breakage frequency rate through the use of historical incident data and to subsequently predict the future. Sacluti (1999) demonstrated their effectiveness rather early on in development of ANN's for civil engineering applications in a case study in Edmonton, Canada. Jafar et al. (2010) had similar success by coupling an ANN with Multiple-Linear Regression in a distribution network in Wattrelos, France. The authors also both acknowledged potential benefits in the use of ANN's for establishing rehabilitation strategies. Evolutionary Polynomial Regression (EPR) is another machine learning technique used to predict deterioration and failure by discovering patterns in pipe failure data amongst homogenous groups (Giustolisi and Savić, 2006). The EPR model produces simple relationship equations between a number of variables and confirmed the importance of; pipe age, diameter and length; when considering pipe burst frequencies and occurrences. The benefit of EPR over more conventional data mining techniques is the simplicity of the relationship equations although it is reliant on the accuracy of the data modelling inputs (L Berardi et al., 2008).

The application of whole life cost modelling for investment planning is a natural progression from statistical deterioration models which usually form the foundations of the approach (Lei and Sægrov 1998). An aggregated statistical model was used in this study alongside a lifetime modelling tool to develop optimised planning regimes. Similar techniques have been used to try to determine the optimal time for pipe replacement by forecasting the service life of a water main under two intervention scenarios; replacement or rehabilitation, (Shamir and Howard, 1979). In advancement towards multi-objective optimisation, researchers have considered how to evaluate the problem of maximising network performance whilst minimising cost. The benefits of presenting the problem in a multi-objective framework are that solutions take the form of a trade-off between cost vs. benefit. Thus, cheaper solutions with lower benefit and lower cost implications are not overlooked; there-by giving the decision maker a broader range of the potential solutions available to them. Halhal et al., (1997) is one of the first cited examples in this field and work continues, (Berardi et al., 2008), (Engelhardt et al., 2002) and (Nafi and Kleiner, 2010).

The aforementioned literature reaffirms that the benefits of these technologies are already being realised for larger and more critical assets, e.g., distribution mains. Water companies are harnessing the power of these techniques to help them select optimal intervention timings and solutions, whilst also targeting maintenance towards poorly performing, high risk assets. In contrast, communication pipes form a typically unmapped and unattributed asset stock, which are often sub-optimally managed as a result of insufficient knowledge and un-optimised investment plans. Although these assets are deemed low value infrastructure in comparison to distribution mains, the high asset volume means that substantial yearly capital investment is needed to maintain serviceability, thus justifying the development and application of a deterioration model to help utility managers better understand the performance of communications pipe. Ward et al. (2015) have progressed the deterioration model further via the development of an asset level decision support tool used to trade-off between whole life costs (totex) and the prevention of future asset failures (serviceability).

Methodology

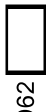
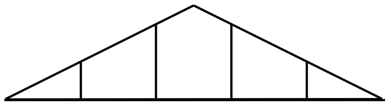
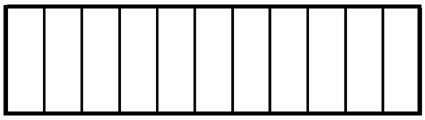
The starting position for the study is to address the lack of data availability and quality for communication pipes, which are often two of the governing factors presenting a barrier to the successful deployment of effective asset management techniques for high volume-low value infrastructure, i.e., where information surrounding the extent, attribution and condition of the asset stock is limited (Vangdal and Reksten, 2011). To overcome this knowledge gap a set of notional assets, which are arbitrary straight line connections formed between customer property address points and the nearest most relevant distribution mains (or trunk main) are generated using a Geospatial Information System (GIS). This approach also allowed for an approximation of asset length. However, with all geospatial processes of this nature some un-realistic asset lengths are created due to the variability of specific individual locations, i.e., where a property actually connects directly to a near-by trunk main but it is digitised as a connection to a distribution main considerable distance away. To account for these anomalies, all notional asset lengths are plotted on a frequency distribution by asset length graph, to help identify a threshold at which individual assets beyond a certain length become highly unlikely or unrealistic. Notional assets lengths exceeding this threshold are subsequently assigned the mean length for the group to avoid unrealistically long lengths biasing the model. For communication pipes, this approach is applied to two groups of asset; short and long side communication pipes. Where-by a long-side communication pipe is defined as a communication pipe connecting to the water main across a road, i.e., the water main is located in the opposite carriage way of the road to the property being supplied. Assets that do not cross a road centre line have been defined as short-side connections for the purpose of this study. The authors deemed this to be a reasonable approach and achieve the main aim of the analysis which is to distinguish between assets with higher associated failure costs, due to asset length and social distribution, e.g., road closures.

Determining and attributing pipe dates retrospectively cannot be precise. Therefore, a logical data hierarchical procedure is developed to take advantage of the most appropriate data sources available. Depending on data availability the process uses a mixture of: corporate communication pipe and distribution mains asset age data; property age estimates from the HM Revenue and Customs (HMRC) Valuation Office Agency (VOA); and historic mapping. Each of the data sets were used in the following manner: Distribution mains age was used as an approximation for communication pipe age by identifying which distribution main the communication pipe was connected to and using the installation date for this asset as a surrogate for its own age. This was only applied so long as the distribution main had not been rehabilitated which would have reset the age of the asset. Property age estimates from HMRC Valuation Office Agency were provided at postcode level and despite not being applicable to an individual property, the methodology assumes that the data could be applied with reasonable accuracy where the majority of properties within a postcode were of the same age. Finally, an analysis of historic mapping created development regions, formed by observing the growth of a town over time, Figure 1. For example, housing estates shown on maps produced in 1950 that were previously not visible on a 1940's map were attributed, via selections within polygons, as a 1940-1950's development.

Figure 1. Development regions

To handle the uncertainty associated with this type of methodology, unique distributions are assigned to each asset depending on the level of confidence surrounding the data source and/or the range of dates it covers. Therefore, the characteristics of a single asset are spread across multiple years according to the distribution, where-by the distribution is selected according to the confidence levels assigned to the different data sources. Whilst it may seem obscure to further segregate individual assets over age ranges, when the asset stock is considered at regional level this approach provides an improved representation of the network by spreading potential uncertainties in the data. Table 1 is prepared to help visualize the different distributions.

Table 1. Age data uncertainty distributions

Data Source	Visual representation	Justification
Installed year	 1962	Uniform distribution for a single year, where the date is known to be accurate and correct.
Mains Age	 1970 1971 1972 1973 1974	The mains installation year is reasonably reliable therefore only a small distribution is applied, centered around the installation year, i.e., 1972.
VOA & Historic mapping	 1950 1955 1960	Uniform distribution from the earliest to the latest date spanned by the range for each individual year.

After establishing a commissioning distribution for each communication pipe, a further procedure is developed to determine the likely material for the asset. This assessment was based on industry research publications and engineering experience from operational management teams who worked to identify unique operating zones where different materials were installed over-time, (WRc and UKWIR, 2005). Within these zones unique material age usage profiles were assigned to help account for the subtle differences surrounding the materials installed over time, often dictated by a local council's preferences or availability of materials in the region. Rather than using traditional binary indicators which would assign a single material to an individual asset, the methodology adopts a type -1 fuzzy inference system (Mendel, 2000). IF-THEN rule sets are used to define the fuzzy membership of an asset to different pipe materials, depending on its age. The use of this technique has helped to address the deficiencies that are inherent in the application of binary logic to imprecise systems by propagating uncertainties through the model (Kleiner et al., 2009). Membership functions are often applied to linguistic variables that describe a condition or state which is imprecise. However, in this instance the installation date for the asset is used as the antecedent to which the fuzzy inference system is applied, in order to derive the consequent which is expressed as a fuzzy membership of the asset to specific materials .There-by addressing the uncertainty of the material used at each installation date and also accounting for the phasing in and out of different materials over time. Figure 2 depicts a typical material usage profile showing the phasing in and out of materials overtime for a single operating zone that installed Lead (PB), Cooper (CU), Black Polyethene (BPE) and Medium Density Polyethylene (MDPE) pipework overtime. A five year phasing in out period was identified by both operational teams as the most likely representation of reality due to the fact that new pipe materials take time to establish themselves as the preferred technology and the stock of the material being replaced takes time to be depleted. As such, assets installed within a transition period are attributed with a percentage of both materials, e.g., 20% CU and 80% BPE for assets installed within this operating zone in 1965.

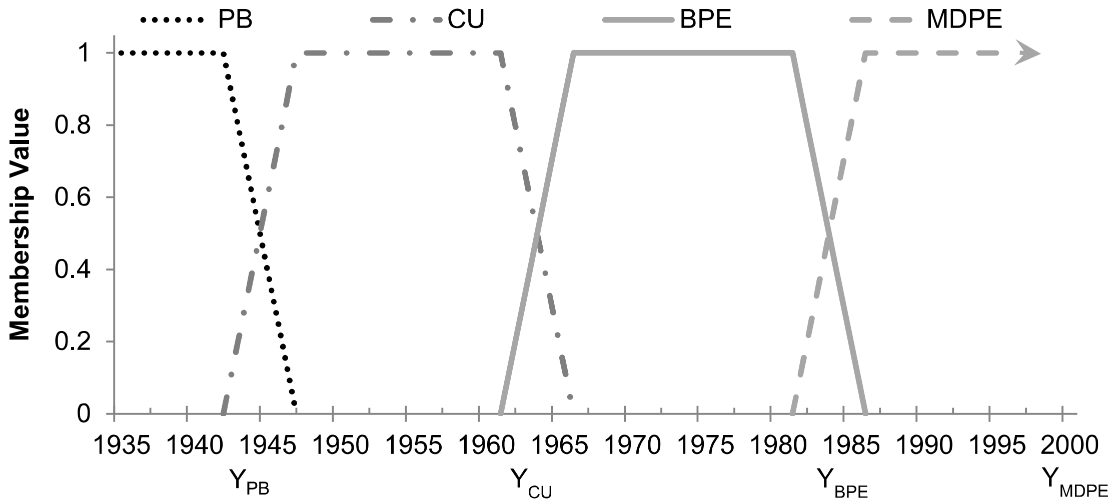


Figure 2. Material usage profile for an individual operating zone

The key assumption underpinning the deterioration modelling process is that all assets in the same material peer group and of the same age have the same probability of failure. The process fully accounts for the assets having different commissioning years and different failure years, whilst applying this age assumption. Hence a PB pipe commissioned in 1930 and reported as failing in 2000 is equivalent in age and hence behaviour to a similar pipe commissioned in 1938 and reported as failing in 2008. A diagrammatic representation of the full modelling process is presented in Figure 3 which is interpreted as follows. Starting with a random estimation for Weibull parameters, the asset stock and associated failure counts are modelled on a yearly basis from 1837 to present day, taking into account new assets that are installed in each year and the predicted failures by material type. The mathematical form of the 3-parameter Weibull probability density function (pdf) is displayed in Equation (1), together with the corresponding cumulative Weibull distribution function in Equation (2)

$$f(T) = \frac{\beta}{\eta}\left(\frac{T-\gamma}{\eta}\right)^{\beta-1} e^{-\left(\frac{T-\gamma}{\eta}\right)^{\beta}} \tag{1}$$

$$F(T) = 1 - e^{-\left(\frac{T-\gamma}{\eta}\right)^{\beta}} \tag{2}$$

In each year the simulated failures are recorded for each pipe material to allow for a comparison against the observed failures during the observation window, 2001 and 2011. The individual errors for each material are then aggregated to provide an overall assessment of error which is used as the measurement of accuracy during the calibration process. An aggregated error is used to ensure that the interaction and transition between the different pipe materials are fully accounted for, i.e., as one asset fails in one of the material groups it can become re-born as a new asset in the same or a more modern material group depending on the fuzzy rule set applied. This material transition is achieved in the model by the use of feedback, applied at each time step, to account for the effects of the historic repair and/or replacement activity. The repair and replacement rates are fixed over time but are variable by material type – this is necessary to reflect actual business policy whereby certain materials are replaced more than others. Upon replacement of the pipe, where this is by a different material, feedback is used to account for the reduction in the former material and the gain in the new material stock. Future failures of the replacement material(s), known as secondary failures, are modelled under the same process. Conversely, whilst repairs do not alter the original pipes material, an adjustment is made in the model to allow for the fact that a pipe failed at particular age and is re-born in the same material group – effectively a "new" pipe, but with the same probability of failure as its peers. Two examples are provided to demonstrate this concept for an asset commissioned in 1920 and attributed as 100% PB material.

1) *If* failure year = 1963 *And*, intervention type = Replacement

Then, PB failure count for pipe age 43 years = +1

And, PB asset stock count for pipe age 43 years = - 1

And, BPE asset stock count for pipe age 1 years = + 0.8

And, CU asset stock count for pipe age 1 years = + 0.2

2) *If* failure year = 1963 *And,* intervention type = Repair

Then, PB failure count for pipe age 43 years = +1

And, PB asset stock count for pipe age 43 years = +/- 0

The calibration process itself is an iterative procedure which seeks to minimise the aggregated error by solving for the Weibull parameters for each pipe material to give the most representative failure output. The calibration process solves for all three Weibull parameters using either: 1. The method of least squares, solved by using the non-linear generalized reduced gradient method (Lasdon et al., 1974); or, 2. Through the application of a multi-objective genetic algorithm considering the residual error in two dimensions, using the NSGAII algorithm in GANetXL (Bicik et al. 2008; Savić et al. 2011). Whilst the genetic algorithm offered a minor improvement in the overall calibration of the parameters, the computational requirement did not warrant its use. Therefore, for the purposes of expediency, the Weibull parameters are calibrated using the method of least squares.

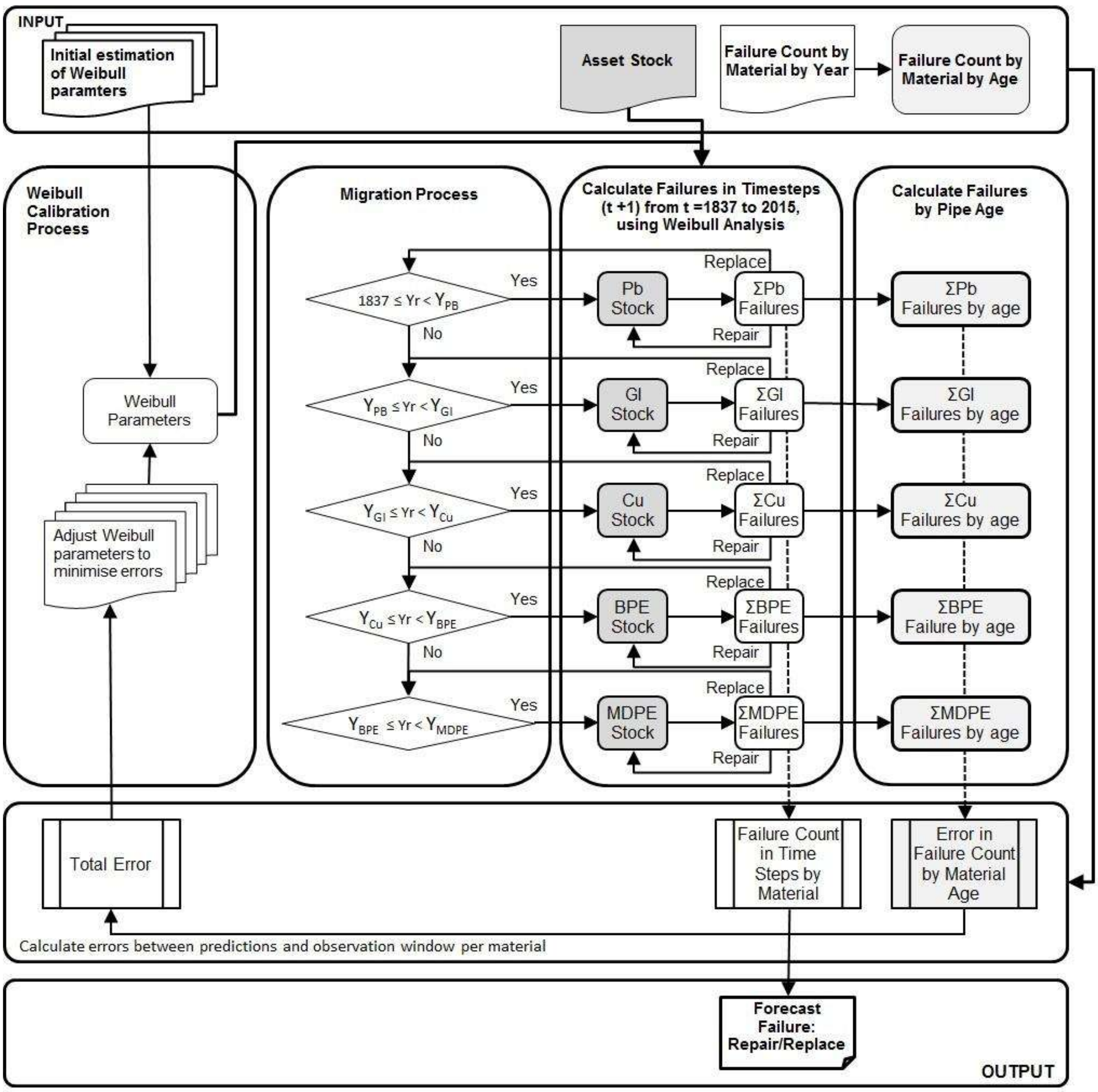


Figure 3. Material migration in a deterioration modelling framework

The output from the process is a calibrated set of Weibull parameters which describe probability of failure for each pipe material over its lifecycle. The probabilities of failure over time have been transformed into a more commonly used metric, reliability, for the comparison of asset performance in the case studies. Reliability values at yearly asset age time steps are calculated by subtracting the cumulative probability of failure from 1, to obtain a probability value between 0 and 1 which describes the percentage of asset stock likely to be in service at any asset age (T), Equation 3.

$$R(T) = \big(1 - F(T)\big) = 1 - \left(1 - e^{-\left(\frac{T-\gamma}{\eta}\right)^{\beta}}\right) \tag{3}$$

Case Study

The modelling approach has been developed to model the deterioration and failure of communication pipes assets which have previously presented a challenge to utility managers due to the typically unmapped and unattributed nature of the asset stock. The methodology has been applied independently to two English water utility providers, allowing testing of the accuracy and robustness of the approach. Both companies in this study own and operate a similar sized asset portfolio of approximately 800,000 communication pipe per Water Company, of which c.85% are unmapped in both organisations.

The previously described notional communication pipe asset digitisation and attribution process, developed in GIS, was run for all unmapped and/or unattributed assets across both regions. A frequency distribution by asset length graph was then used to identify a threshold at which individual assets beyond a certain length become highly unlikely or unrealistic in each region to account for local geographic conditions and typical arrangements. The outputs for Water Company 1 provided a threshold of 18m and 46m for short-side and long-side communication pipes lengths respectively. A mean length value of 2.5m and 7.3m for short and long-side connections was observed and these mean values were applied as corrections to all notional assets in excess of the threshold lengths. As expected, a greater standard deviation from the mean was witnessed for long side connections of 5.4m versus 2.7m for short-side connections. This is to be expected and reflects the non-standard property layouts for dwellings set back from the road, as opposed to those that follow more traditional construction arrangements that run parallel to the road alignment. The short and long side communication pipe length distributions, post correction, are shown in Figure 4 and Figure 5 for Water Company 1.

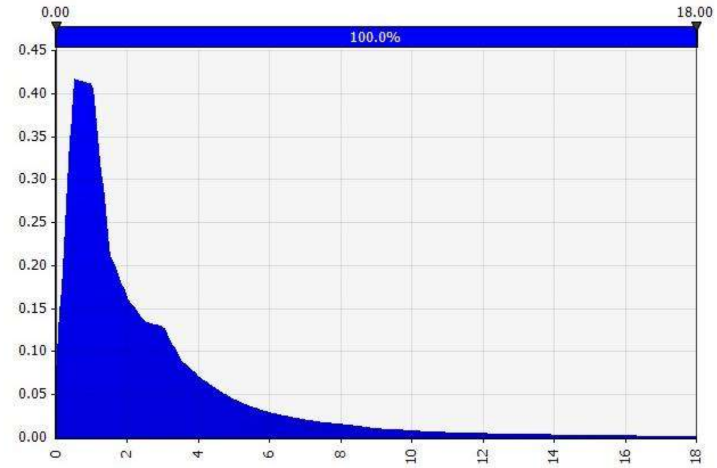


Figure 4. Short side communication pipe length distribution

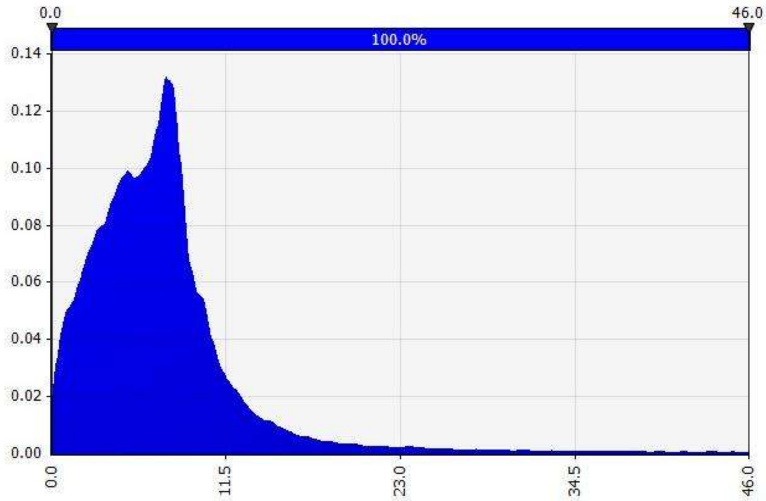


Figure 5. Long side communication pipe length distribution

Asset age and material were later inferred for the notional asset stock using the most appropriate available age data source for each asset and the associated fuzzy material rule sets according to the operating zone encompassing that asset. Eight and twelve distinctly different operating zones and accompanying fuzzy rules sets were identified by the two utilities companies, leading to the assignment of the assets within these zones to materials according to the fuzzy rule set for the zone, e.g., Figure 2. . For the purposes of deterioration model calibration a combined 122,649 communication pipe failure observations were supplied over an eleven year period between 2001 to 2012 by Water Company 1 (60,827 records) and Water Company 2 (61,822 records). A normalised failure rate calculation for each material and each observation year is plotted using Equation 4, Figure 6, to provide a comparison of the failure data supplied by each Water Company.

$$Failure\ Rate\ (No.per\ year) = \frac{\sum Failures\ Count}{\sum Assets} \tag{4}$$

It is observed in Figure 6 that the failure rates witnessed by both companies are broadly in agreement and identify higher failure rates for PB and BPE pipe materials and lower rates for Galvanised Iron (GI). Whilst the observed failure rates for PB assets during this window are higher than other materials, it does not indicate that PB is a poorly performing material. In fact, given the estimated average age for PB pipes, which were no longer installed past the 1960's, the failure rate is surprisingly comparable to the more modern pipe material, BPE.

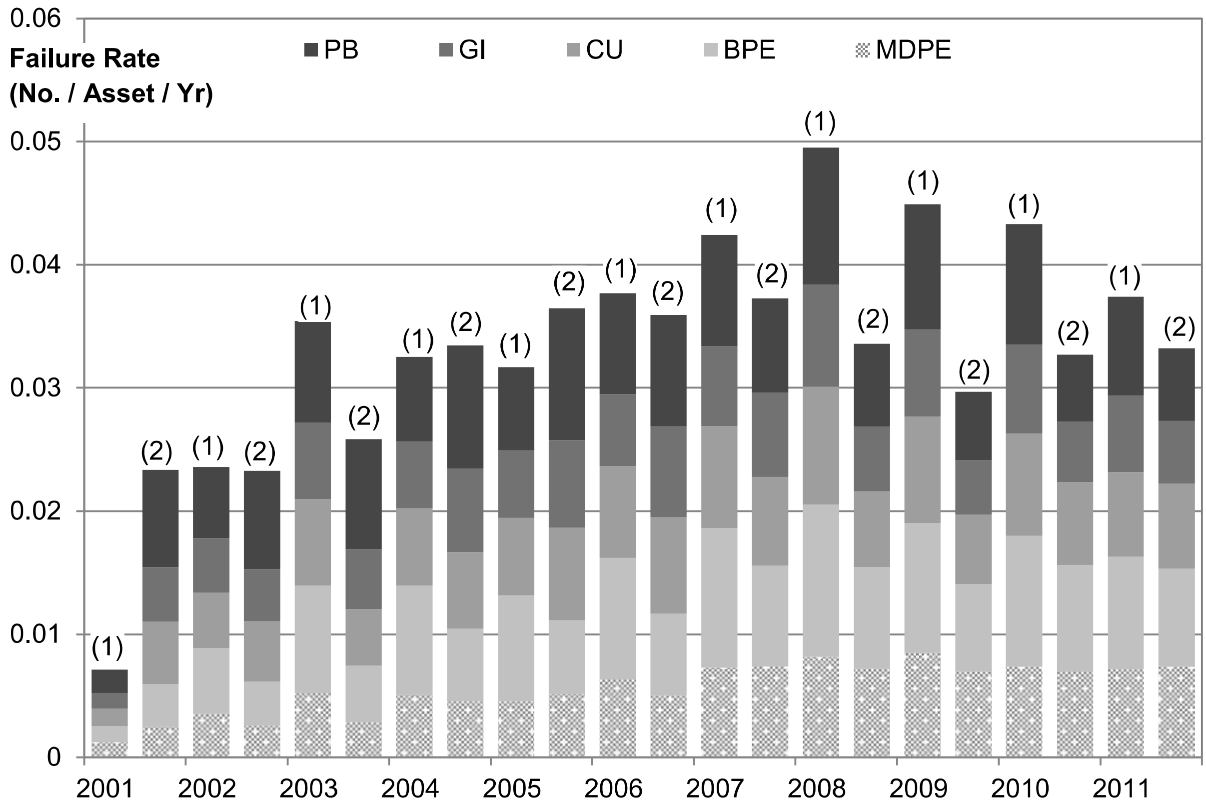


Figure 6. Comparison of failure rates by material type over-time between Water Company (1) and (2).

The failure observation records used for the calibration of deterioration curves for each pipe material are applied independently for each Water Company to account for any subtle differences in pipe material performance in each region, which could arise from differing ground conditions, climate variations and/or the quality of workmanship during installation. The observations used in the calibration of the deterioration curves were recorded over an eleven year period, 2001 to 2011 inclusive. The failure observations for Water Company 1 in 2001 were not used in the calibration process to prevent from biasing the model based on smaller than normal failure counts. The notably reduced failure counts in this year are thought to be due to missing records, captured on a previous work order management database as the company transitioned to a new system. Whilst a failure observation window of ten years is deemed to be relatively low for assets installed as far back as the 1800's, statistical modelling techniques have been shown to be successful when applied to buried water distribution network assets with similar failure histories (Røstum, 2000), (Le Gat and Eisenbeis, 2000), (Mailhot et al., 2000) and (Pelletier et al., 2003). Given that communications pipes are smaller and less critical assets than those modelled in the aforementioned literature, the availability of failure data is surprising.

A Weibull based statistical modelling approach was selected due to the relatively low failure records in respect to the overall asset stock (8%) and for the usefulness of the graphical Weibull survival curve outputs in an engineering and decision making context (Weibull, 1951), Figure 7. It is observed that each pipe material tends to follow a typical S-Curve reliability distribution which can generally be described as follows: Failure at commissioning due to deterioration is generally low, with asset reliability values in excess of 90% for assets under 40 years of age. However, after the initial onset of failure the asset reliability decreases at its fastest rate for a period of time and until it reaches a more steady state of reducing reliability between approximately 75 to 125 years. After the steady state of reducing reliability, incremental change in probability of failure lowers as the asset progresses towards "old age", when the majority of assets in the same class have failed leaving a small residual remaining in a serviceable state for a notable time period.

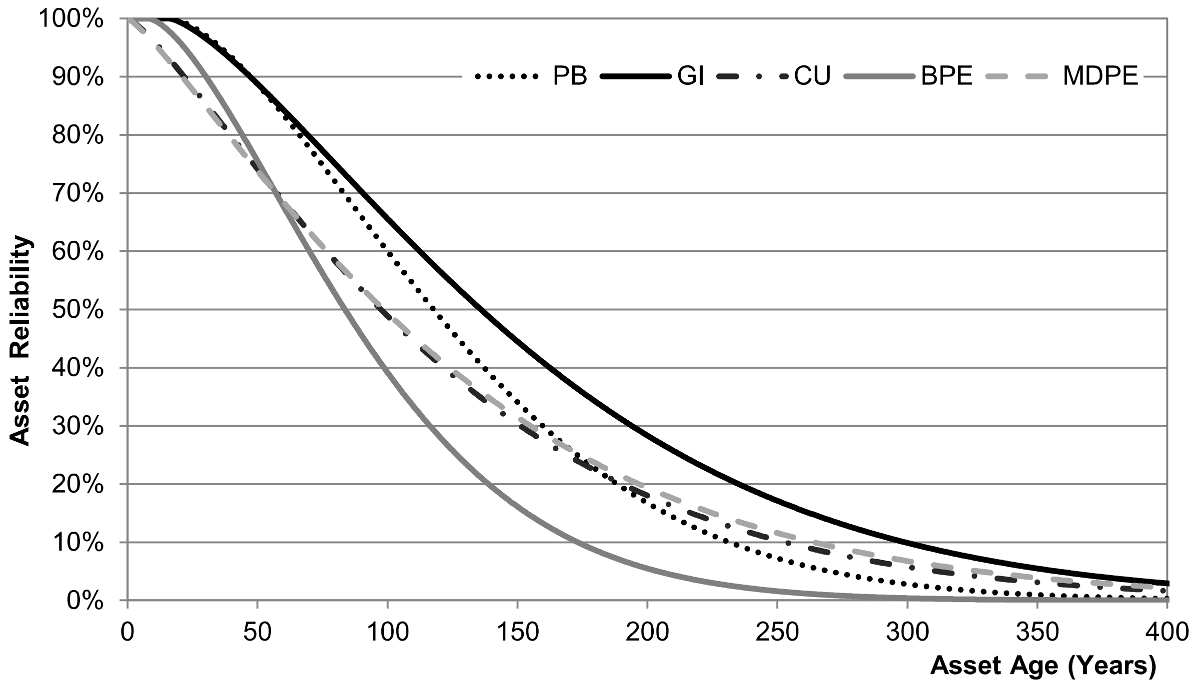


Figure 7. Deterioration profiles by material for Water Company 1

It can also be seen that whilst the reliability of GI assets is consistently higher than other materials over-time and conversely BPE is consistently lower, the reliability of the other materials is interchangeable over time. For instance, it is observed that PB assets outperform CU and MDPE pipework up until an asset age of approximately 115 years at which time PB continues to have a lower reliability. However, when considering an even more useful comparator for material performance, asset half-life, it shows little difference in performance: PB (101 years), MDPE (98 years), CU (97 years). Where-as it can be seen that 50% of GI assets are expected to be in service past 120 years of age in comparison to only 75 years for BPE. The accuracy of all the deterioration modelling curves against the failure observations for Water Company 1 and Water Company 2 are shown by the coefficient of determination, R-Squared, in Table 2 and Table 3. In support of this analysis, Figure 8 is produced to demonstrate the accuracy of predicted total failure counts for PB pipes by asset age versus the actual failure observations.

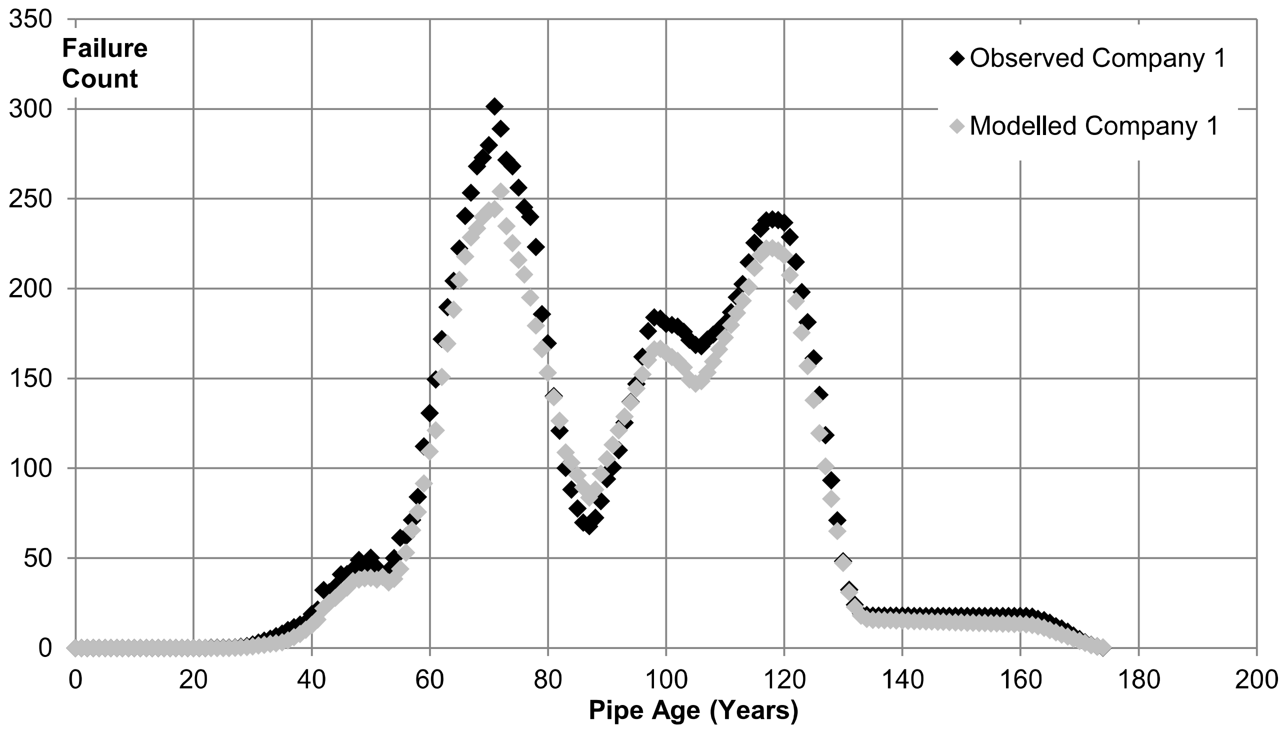


Figure 8. Modelled vs. Observed Failure Counts for Lead (PB) Pipes by Age

Figure 8 clearly demonstrates two distinct pipe ages where the failure count is highest; pipe age 70 years and 120 years. The 70 year peak coincides with steepest rate of failure for PB, Figure 7, and the highest asset stock attributed to post World War II construction. The accuracy of the data fit between the modelling outputs and the observed failure data is evaluated using a commonly accepted coefficient of determination, R-Squared (Gupta and Guttman, 2014) . The R-Squared is calculated by looking at the individual errors between predicted and observed failure counts for specific pipe ages and materials, e.g, 42 year old PB pipe. Therefore, each point on the graph represents a single pipe age and the associated predicted failure counts for this pipe is plotted against the observed failures, Figure 9. Whilst this is a fairly onerous assessment of model accuracy it helps to identify the individual age bands at which the deterioration model either over predicts, 81 – 100 years, or under-predicts, 61 – 80 years. For PB pipes it is observed that the model fairly consistently under predicts, albeit only marginally but more notably around the peak failures. Reassuringly, this is not a common characteristic of the model. In fact, BPE is the only other material with a slight under prediction. Whereas the other pipe materials experience marginal over predictions at the peak failure rates.

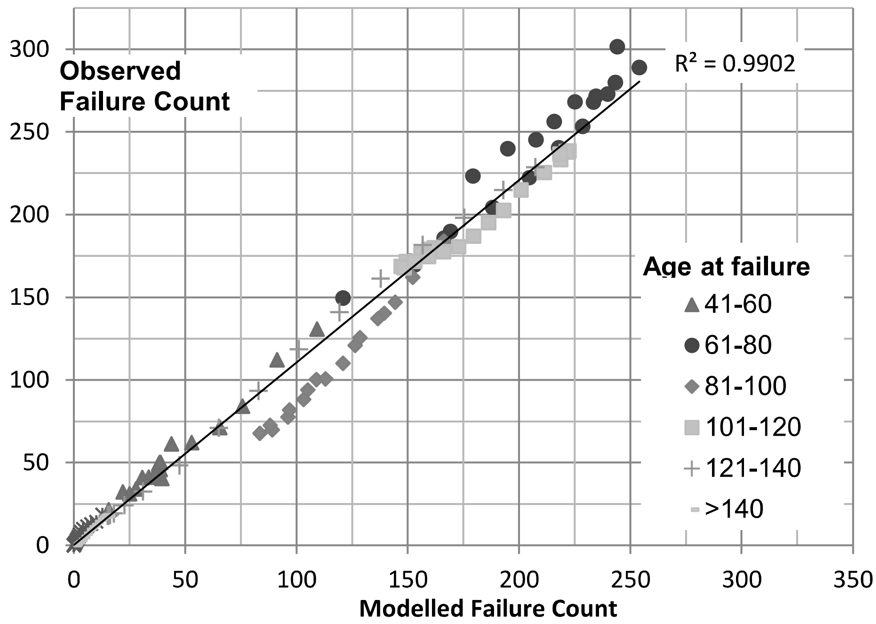


Figure 9. Data Fit for Lead (PB) Failures by Pipe Age

A full analysis of the accuracy of all pipe material deterioration models for Water Companies 1 and 2 is presented in Table 2 and Table 3, showing R-Squared values greater than 0.959 for all material types across both organisations. The failure observation counts used in the calibration of the Weibull parameters are also shown. It is observed that the Weibull parameters and associated deterioration curves are similar in nature between Water Company 1 and Water Company 2, despite the modelling framework being customised for each company to account for the different historic material usage profiles and failure observations. This is reassuring as it is expected that communication pipes of the same material and those installed in the same year, would deteriorate in a similar manner regardless of where they were installed. The reason for the location parameter (γ) only being used for older pipe materials is because failure observations in the dataset were not observed in these materials for ages less than 16 and 29 years. This is due to the phasing out of the PB and GI pipework between the 1950's to 70's and our failure observation window starting in 2001. As a result, all failures occurring in the early ages of asset life, which is often referred to as infant mortality, are missed by this model. However, this will not hinder the accuracy of the model for future predictions because PB and GI assets are no longer installed and infant mortality failures counts are typically very low.

Table 2. Water Company 1 Modelling Data

Material	Failure Observations	β	η	γ	R^2
PB	13,972	1.599	127.968	16	0.990
GI	3,090	1.398	155.875	16	0.990
CU	14,074	1.259	130.434	0	0.988
BPE	15,609	1.536	90.005	0	0.988
MDPE	13,454	1.218	133.136	0	0.996

Table 3. Water Company 2 Modelling Data

Material	Failure Observations	β	η	γ	R^2
PB	26,094	1.458	103.479	29	0.973
GI	15,247	1.158	152.056	29	0.960
CU	5,161	1.149	130.707	29	0.996
BPE	8,210	1.804	110.162	0	0.988
MDPE	7,110	0.922	343.824	0	0.974

Weibull deterioration curves are particularly useful for predicting future failure rates over a given time horizon, whereby the time horizon for failure projections can be extended for more accurate models (Røstum, 2000). To test the accuracy of the calibrated model in this study the predicted future failures are tested against the observed failures for the three years immediately following calibration, 2012 to 2014. A cumulative yearly failure count is produced to show the accuracy of the modelled output versus the observed failures during the calibration window (2002 to 2011) and the three year model verification period (2012 to 2014). A ten year failure rate projection is also plotted, Figure 10. The cumulatively failure count is a useful comparator for model accuracy because it is not influenced by yearly extremes which can be caused by unpredictable external conditions such as the weather. The final cumulative failure count from 2002 to 2014 produced by the model is within 1% of the observed failures. It is also reassuring that on average the absolute error for the modelling predictions during the verification window are within 5% for each individual year. This is compared to 8% for the calibrated data when yearly outliers are removed. Where-by outliers are defined as those with values +/- 1 standard deviation from the mean. Without this exclusion the models performance against the validation and calibration data is 18% and 14% respectively.

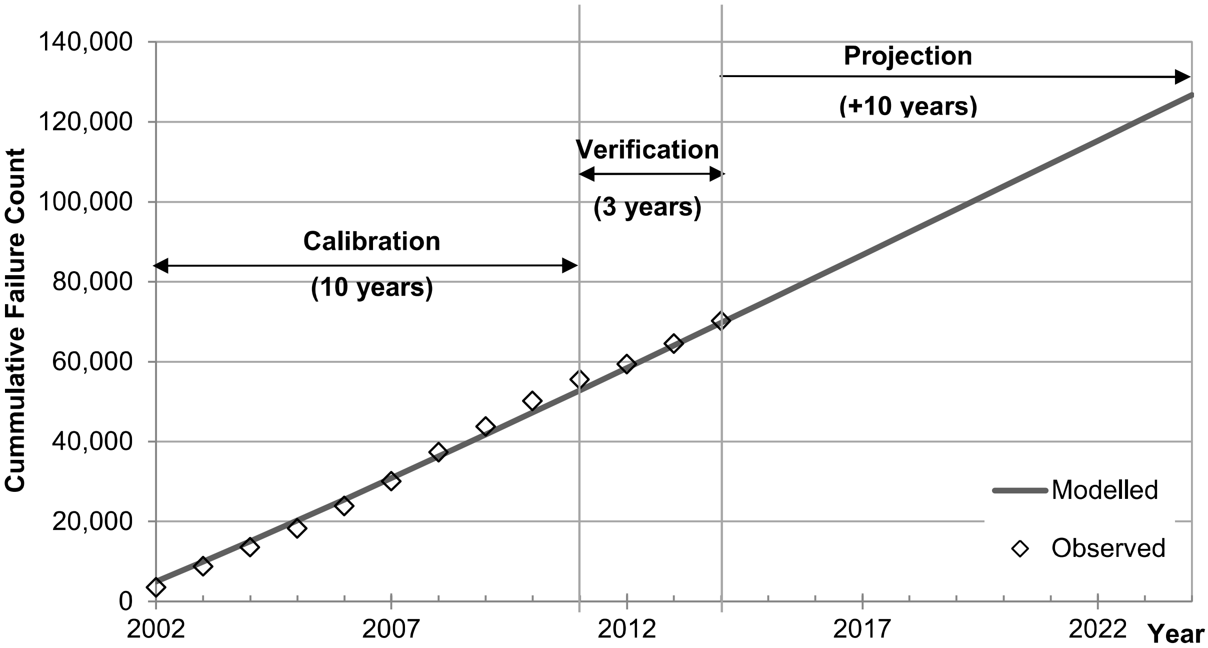


Figure 10. Model calibration, verification and projection

Conclusions

This paper demonstrates the effectiveness of a bottom up deterioration modelling framework for small diameter water distribution assets which connect individual customers to water distribution mains. The low asset value and low consequence of failure, for individual assets of this nature, is thought to be the main reason why previous literature has been focused on larger and more critical assets. However, when considered as a collective asset group, investment across the developed World is estimated to be in excess of £4.4bn per year, demonstrating the potential benefits of adopting more comprehensive asset management techniques founded on deterioration models of this nature.

The modelling approach developed by the authors uses a series of geospatial algorithms and fuzzy rule sets to infer the asset stock and assign basic attribution, prior to calibrating Weibull reliability curves for the five most commonly installed pipe materials; PB, GI, CU, BPE and MDPE. A three parameter Weibull function is used to simulate failures for each time step so that they can be compared against actual failure records within a ten year observation window for each material. Feedback processes that incorporate fuzzy rules sets have been introduced within the modelling framework to account for the fact that these assets are either repaired or replaced upon failure and that different materials were used when replacing these assets over time. The output is a calibrated set of Weibull parameters for each pipe material which can be used to calculate yearly reliability values that describe the probability of an asset remaining in service at any given age.

All five pipe materials were observed to behave differently over time and the modelled failure rates showed good correlation to the observed failures with R-Squared values exceeding 0.97. Similar reliability curves were observed for each of the materials between two UK based Water Companies that this modelling framework was applied to independently. Finally, the accuracy of the modelling approach was validated by comparing the predicted number of failures against three years of failure data not used in the calibration process. The yearly failure counts were predicted to within +/-5% accuracy and the overall cumulative modelled failure count at the end of 2014 was predicted within 1%.

The benefit of developing an accurate modelling framework to simulate future deterioration and failure of Communication pipe assets is that future maintenance expenditure can be modelled, which in turn provides Water Companies with a better understanding of their future capital and operational budgetary requirements. Additional benefits could be realised through the application of a whole life cycle cost optimisation model to identify the most cost effective maintenance strategy with respect to total expenditure. Similarly, maintenance strategies could also been developed with the objective of improving serviceability if a relationship between asset degradation and performance can be established. For example, by understanding the impact of GI pipework degradation on discolouration events and/or low pressure, a targeted proactive replacement programme could be implemented with the aim of improving customer satisfaction. However, this is a complex phenomena to model due to the number of additional contributing factors that lead to reduced serviceability, some of which are associated with the water distribution network itself and not necessarily the communication pipe assets. Therefore, the delivery of an effective proactive communication pipe replacement programme should give consideration to the combined structural and serviceability performance of all components in the distribution network.

Acknowledgments

The authors gratefully acknowledge the continued support from EPSRC through their funding of the STREAM Industrial Doctorate Centre, and from the project sponsors (AECOM).

References

Berardi, L., Giustolisi, O., Kapelan, Z., and Savic, D. A. (2008) Development of pipe deterioration models for water distribution systems using EPR. Journal of Hydroinformatics, 10(2), 113. [online] http://www.iwaponline.com/jh/010/jh0100113.htm.

Bicik, J., Morley, M. S., and Savić, D. A. (2008) "A Rapid Optimization Prototyping Tool for Spreadsheet-based Models". Proceedings of the Water Distribution Systems Analysis conference. South Africa, 472–482.

Clark, R., Stafford, C., and Goodrich, J. (1982) Water Distribution Systems: A Spatial and Cost Evaluation. Journal of the Water Resources Planning and Management Division, 108(3), 243–256.

Engelhardt, M. ., Skipworth, P. ., Savić, D. ., Saul, A. J., and Walters, G. (2000) Rehabilitation strategies for water distribution networks : a literature review with a UK perspective. Urban Water, 2.

Engelhardt, M., Skipworth, P., Savic, D. a., Cashman, a., Walters, G. a., and Saul, a. J. (2002) Determining maintenance requirements of a water distribution network using whole life costing. Journal of Quality in Maintenance Engineering, 8(2), 152–164. [online] http://www.emeraldinsight.com/10.1108/13552510210430026.

Le Gat, Y. and Eisenbeis, P. (2000) Using maintenance records to forecast failures in water networks. Urban Water, 2(3), 153–181.

Giustolisi, O. and Savić, D. A. (2006) A symbolic data-driven technique based on evolutionary polynomial regression. , 207–222.

Gupta, B. and Guttman, I. (2014) Statistics and Probability with Applications for Engineers and Scientists, John Wiley & Sons.

Halhal, D., Walters, G. A., Ouazar, D., and Savić, D. A. (1997) Water Network Rehabilitation with Structured Messy Genetic Algorithm. Journal of Water Resources Planning and Management, 123(3), 137 – 146. [online] http://link.aip.org/link/JWRMD5/v123/i3/p137/s1&Agg=doi.

Jafar, R., Shahrour, I., and Juran, I. (2010) Application of Artificial Neural Networks (ANN) to model the failure of urban water mains. Mathematical and Computer Modelling, 51(9-10), 1170–1180. [online] http://linkinghub.elsevier.com/retrieve/pii/S0895717710000051

Kleiner, Y., Rajani, B., and Sadiq, R. (2009) “Drinking Water Infrastructure Assessment: The National Research Council of Canada Perspective”. World Environmental and Water Resources Congress 2009. Kansas City, Missouri, American Society of Civil Engineers, 1–10. [online] http://ascelibrary.org/doi/abs/10.1061/41036(342)67.

Lasdon, L. S., Fox, R. L., and Ratner, M. W. (1974) Nonlinear optimization using the generalized reduced gradient method. RAIRO-Operations Research-Recherche Opérationnelle, 3(8), 73–103.

Lei, J. and Sægrov, S. (1998) Statistical approach for describing failures and lifetimes of water mains. Water science and technology, 38(6), 209–217.

Li, D. and Haimes, Y. (1992) Optimal maintenance-related decision making for deteriorating water distribution systems: 1. Semi-Markovian Model for a water main. Water Resources Research, 28(4), 1053–1061.

Mailhot, A., Pelletier, G., Noël, J.-F., and Villeneuve, J.-P. (2000) Modeling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: Methodology and application. Water Resources Research, 36(10), 3053–3062. [online] http://doi.wiley.com/10.1029/2000WR900185.

Mendel, J. . . (2000) Uncertain Rule-based Fuzzy Logic Systems: Introduction and New Directions, Prentice Hall.

Nafi, A. and Kleiner, Y. (2010) Scheduling Renewal of Water Pipes While Considering Adjacency of Infrastructure Works and Economies of Scale. Journal of Water Resources Planning and Management, 136(5), 519–530.

O’Day, K., Weiss, R., Chiavari, S., and Blair, D. (1986) Water main evaluation for rehabilitation/replacement. American Water Work Association.

Ofwat (2010) June Return. Past information on company performance (June returns). [online] http://www.ofwat.gov.uk/regulating/junereturn/jrhistoricdata/.

Ofwat (2014) Responsibility for water and sewerage pipes. [online] http://www.ofwat.gov.uk/consumerissues/rightsresponsibilities/supplypipes/

Pelletier, G., Mailhot, A., and Villeneuve, J. (2003) Modeling Water Pipe Breaks — Three Case Studies. Journal of Water Re, 129(1), 115–123.

Rajani, B. and Kleiner, Y. (2001) Comprehensive Review of Structural Deterioration of Water Mains : Physically Based Models. Urban Water, 3(3), 151–164.

Røstum, J. (2000) Statistical Modelling of Pipe Failures in Water Networks.

Sacluti, F. R. (1999) Modeling Water Distribution Pipe Failures Using Artificial Neural Networks.

Savić, D. a., Bicik, J., and Morley, M. S. (2011) A DSS generator for multiobjective optimisation of spreadsheet-based models. Environmental Modelling & Software, 26(5), 551–561. [online] http://linkinghub.elsevier.com/retrieve/pii/S1364815210003087.

Shamir, U. and Howard, C. (1979) An Analytic Approach Pipe Replacement. American Water Work Association, 71(5), 248 – 258.

United Nations (2011) Population Division. Department of Economic and Social Affairs, World Popu(ST/ESA/SER.A/313).

Vangdal, A. C. and Reksten, K. (2011) "Experience from inspection of water pipes – hard fact based AM - opportunities and consequences". Proceeding of the 4th Leading Edge Conference on Strategic Asset Management. Mülheim an der Ruhr, Germany.

Ward, B., Selby, A., Gee, S., Rosser, S., and Savić, D. (2014) Assessing impacts of the private sewer transfer on UK utilities. Infrastructure Asset Management, 1(2), 23–33. [online] http://www.icevirtuallibrary.com/content/article/10.1680/iasma.13.00003.

Ward, B., Smith, D., Roebuck, J., Savic, D.,and Collingbourne, J., (2015) Optimised investment planning for high volume- low value buried infrastructure assets. In Review.

Weibull, W. (1951) A Statistical Distribution Function of Wide Applicability. Journal of Applied Mechanics, 293–297.

WRc and UKWIR (2005) A Guide to Water Service Pipes. 3rd Edition. WRc.