1st International WDSA / CCWI 2018 Joint Conference, Kingston, Ontario, Canada – July 23-25, 2018

LEAK DETECTION AND LOCALIZATION BASED ON SEARCH SPACE REDUCTION AND HYDRAULIC MODELLING

Sophocleous S.[1], Savic D.A.[2], Kapelan Z.[3], Gilbert C.[4], Sage P.[5]
[1,2,3]Centre for Water Systems, University of Exeter, North Park Road, Exeter, EX4 4QJ, UK
[4]Severn Trent Water Ltd, Stratford Road, Longbridge, Warwick, CV34 6QW, UK
[5]WITSConsult Ltd, Milton Rough, Acton Bridge, Northwich, CW8 2RF, UK
[1]ss694@exeter.ac.uk

ABSTRACT

Reducing the cost and time required to isolate leakages occurring in Water Distribution Networks (WDNs) is the main task for resilient and sustainable management of these systems. The paper presents a systematic model-based leak detection and localization framework using optimization. The approach prerequisites a well calibrated WDN hydraulic model. The leakage localization model splits into two stages: (a) the search reduction stage where the number of decision variables and the range of possible values are reduced, and (b) the leak detection and localization stage for isolating the fault. The leakage localization method is formulated to optimize the leakage node locations and their associated emitter coefficients, such that the differences between the model predicted and the field observed values for pressure and flow are minimized. The optimization problem is solved by using a non sorting genetic algorithm. A real case from a UK system is presented with the outcome showing that the method reduces the leak search space within 10% of the WDN, while contributing to earlier leakage hotspot detection and localization. The framework for predicting leakage hotspots can be effective despite the recognized challenges of model calibration and the physical measurement limitations from the collected pressure and flow data.

Keywords: Leak Localization; Hydraulic Modelling; Optimization

1 Introduction

Leaks often remain undiscovered regularly resulting in large water and revenue losses for Water Distribution Network (WDN) operators. With time, their impact grows and may result in catastrophic bursts, causing significant changes in the system's operation. Finding leaks early is critical to a water company for either economic, environmental or reputational reasons. Historically, finding leaks has been challenging because even a substantial event can potentially show no manifest signs [1]. A wide range of leak detection and location techniques exists, relying on various approaches [2]. However, there is no universally agreed methodology for leak detection and localization with the number of techniques currently used by practitioners being limited.

Automatic leak detection requires pressure and flow field measurements. The smaller the monitored area, the easier it is to detect a leak automatically. Leak localization is frequently performed using acoustic equipment such as listening rods, leak correlators, leak noise loggers and non-acoustic methods, such as gas injection, ground penetrating radar technology and infrared photography [3]. These methods are accurate, however, it takes intense labor effort and long time to find a leak, even in small District Metered Areas (DMAs). To expedite the leak localization process, software-based methods, such as mathematical modelling are needed in addition to the hardware-based methods.

Compared to a leak-free situation, a leak causes larger flow in a pipe, resulting in larger head loss and different pressures within the WDN [4]. In a well monitored WDN, this creates a unique "signature" on pressure and flow data, which can be used to find its size and location. Calibrated models can be used to perform reliable WDN simulations and compare against the field data. Such automatic leak localization techniques can reduce the area of interest and facilitate traditional district audits to find the leak. Besides large volumes of water, this can save time and money.

The literature on leak detection and localization methods for WDN focuses on how to prioritize areas for leak surveys and facilitate pinpointing of leaks. Such research started with a seminal paper [5] that formulates the leak detection and localization problem as a least-squares estimation problem. However, parameter estimation is not an easy task, due to the non-linear nature of WDN models and the limited availability of field measurements relative to the number of parameters to be estimated, resulting in an underdetermined problem [6]. Many of these approaches are based on transient analysis, which is mainly used on a single, grounded pipeline due to the high effect of the system uncertainty on results. Non-transient model-based techniques have been also developed in recent years. Through the use of optimization techniques, these approaches analyse the difference between measurements and estimated values from leaky scenarios to signal the probability of a zone experiencing leakage. In addition, some of these model-based methodologies hypothesize a single leak in a WDN [7]. To date, model-based leak detection methodologies have not reached the maturity required for mainstream adoption by the water industry.

A prerequisite of accurate leak localisation is a well-calibrated model. However, the inverse problem is often ill-posed, characterized by the non-uniqueness of the identified parameters. Thus, multiple combinations of decision variable values can produce equally fit solutions, but inaccurate leak localisation. A solution to this problem can be to improve the accuracy of leak localisation by reducing the search space without losing optimum solutions. A study reported in [8] detected and localized leakage as a pressure driven demand (using emitter coefficients), providing a tool for assisting leakage detection engineers to identify leakage hotspots. The developed method reduced the problem dimensionality by specifying the maximum number of possible leaks within a system. However, the method did not narrow down the number of candidate leakage hotspot locations or the range of flows, prior to conducting the leakage hotspot detection via inverse modelling. Thus, it is easy for an algorithm to prematurely converge in a large search space, which further affects the accuracy of the leak localization. A novel optimization method called step-by-step elimination method combined with a Genetic Algorithm (GA) was proposed in [9] to calibrate the model and detect leakage. The staged approach eliminated node locations where leakages were not reported during the optimisation process, however, it was only tested in hypothetical and laboratory networks. Another approach [10], proposed a model pre-processing method based on sensitivity analysis to simplify the calibration problem for leak detection purposes. However, it did not ensure that the optimum solution remained in the search space.

This article presents an improved model-based approach for finding leaks in DMAs from pressure and flow data. The method is based on search space reduction, ensuring that the optimum solution is not lost. It exploits existing WDN analysis methods to narrow down the search space of the inverse modelling problem to highly sensitive decision variables. The method uses an optimization model for leakage localisation applications in real WDSs. A GA is used to solve the optimization problem searching for calibration parameters values, while minimizing discrepancies between observations and model predictions. The proposed methodology is applied to a real WDN from the UK.

2 Methodology

2.1 Overview

The proposed method detects and localizes leaks in DMAs using emitter coefficients, based on the simultaneous comparison of all available pressure and flow data captured from deployed sensors with the hydraulic model outputs, as part of a simulation-optimization framework. A well-calibrated model is a prerequisite for this approach. The methodology involves two main stages (a) the Search Space Reduction (SSR) stage and (b) the Leak Detection and Localization (LDL) stage. Simulations were run using EPANET [11] Programmer's Toolkit linked with a MATLAB optimization code. During the SSR stage, the number of decision variables and the range of possible values is reduced through the implementation of three main steps. The three steps include a reduction via: (a) Problem Simplification, (b) Analysis of Minimum Detectable Nodal Leakage (MDNL), and (c) Optimization Analyses. Then, at the LDL stage, an optimization problem of searching for calibration parameter values is solved to indicate the size and location of leaks in the WDN.

2.2 SSR via Problem Simplification

A node is considered a candidate leak location for adjustment during the inverse modelling process if there is uncertainty in its emitter value. Thus, all nodes in a WDN model (including pipes, valves, pumps and tanks) could be potential leak locations resulting in a vast combinatorial search space. In reality, the majority of leaks happen on pipes, but the model normally assigns aggregated demands to nearby nodes. In a WDN model, not all nodes represent areas of aggregated demand. For example, it is normal practice to add two nodes to a pipe section just upstream and downstream of a valve, without assigning any demand to them. Although it is true that valves can leak due to a weak stem, the losses are often insignificant relative to the inlet flow, or to the undetected leaks. Step 1 of the SSR stage reduces the number of decision variables, assuming that leaks only happen on pipes, thus, all nodes associated with non-pipe components are removed.

2.3 SSR via Analysis of Minimum Detectable Leakage

Depending on the sensor configuration there is limited observable WDN space, i.e., the length of pipes that can be monitored for leakage. Furthermore, all devices are accurate within a specified range. Thus, if the pressure change from a leak is below the device's accuracy range, the event will remain undetectable regardless the distance from the sensor. Based on the number of sensors, location and reading accuracy range, there is a minimum detectable flow for each location, which establishes a lower bound for the subsequent optimization analysis. The Minimum Detectable Nodal Leakage (MDNL) process (Step 2) starts by simulating the boundary conditions and analysing the resulting pressure at nodes where sensors are present. Then, a leak with a large emitter flow relative to the inlet is simulated at every candidate leakage node. The pressure residuals between a no-leak and leak scenario are determined for all possible leak nodes. Each time the leak size is systematically reduced until the residuals from the simulated leak flow across all sensors do not exceed the sensor accuracy range. This establishes the MDNL for each potential leak node in the WDN model.

2.4 SSR via Optimization Analysis

The third step estimates the total losses and the approximate number of leaks in the WDN, to further reduce both the number of potential leak locations and the range of flows. A number of implicit non-linear optimization problems for parameter identification are formulated for a different number of potential leaks to test various scenarios for the WDN state. This is to emulate that leaks do not happen everywhere, but only on few pipes at a time. Solutions are evaluated by the discrepancy between simulated and field measured pressure heads and pipe flows. Step 3 divides into two parts. In Part I, the optimization analysis estimates the total water losses in the WDN and in Part II a second optimization problem is solved with the updated list of candidates and range of flows, to estimate the number of possible leaks. In both parts, optimization is formulated with a fixed number of decision variables. For any possible leak in the WDN two decision variables are defined, one for the leak location and one for the leak flow value [10]. The problems are subject to two sets of constraints: (1) the implicit constraints considering mass and energy balance equations; and (2) the explicit constraints used as bounds for the solution search space for each decision variable. The optimization is formulated as:

Search for: $X = (LN_i^n, K_i^n);$ $LN_i^n \in J^n;$ $n = 1,....,NLeak;$ $i = 1,...,NIndex^n$ (1)

Minimize: $F(X)$ (2)

Subject to: $0 \leq K_i^n \leq \overline{K^n}$ (3) $P_i^n \geq 0$ (4) $\sum_{n=1}^{NLeak} NL_{dup}{}^n$ (5)

where LN_i^n is the index for node i for the specified leak n, K_i^n is the emitter coefficient for node i for leak n, J^n is the set of candidate nodes for leak n, $NLeak$ is the number of specified leaks to be identified, $NIndex^n$ is the number of the candidates for the group of specified leakage nodes n to be identified, $\overline{K^n}$ is the maximum emitter coefficient for specified leak n, P_i^n is the pressure head at the detected leakage node i within group n and $NL_{dup}{}^n$ is the number of the duplicated nodes that are identified as leakage emitters in one solution for group n.

In Part I a single leak is assumed, and a total of two decision variables are considered in the optimization problem minimizing the weighted sum of squared flow differences, given by:

Minimize: $$F(\vec{X}) = \sum_{t=1}^{T}\Big[\sum_{nf=1}^{NF}\frac{w_{nf}\left(\frac{Qs_{nf}(t)-Qo_{nf}(t)}{Qpnt}\right)^2}{NF}\Big] \quad (6)$$

where $Qo_{nf}(t)$ is the observed flow of the nf-th link at time step t; $Qs_{nf}(t)$ is the simulated flow of the nf-th link at time step t; $Qpnt$ is the flow per fitness point, which converts flow differences into a dimensionless value based on the meter's reading accuracy; NF is the number of observed flows; Wnf is the normalized weighting factor for observed flows.

At each analysis K_i^n is allowed to systematically vary for ±50%, ±25%, ±10%, ±5% and ±1% relative to the specified emitter flow value. The search begins with an initial K_i^n equal to the maximum flow difference. When a fit solution is identified, the range of flow adjustments is consistently applied and K_i^n is updated with the optimal value. At the end of Part I the upper bound flow, $\overline{K^n}$, for all candidate nodes is established. If the estimated flow is below the MDNL value of a specific candidate node, then, the node is removed from the search space. During Part II a series of possible leak scenarios with fixed n is simulated based on the $\overline{K^n}$ value and the average MDNL of the remaining candidates. Optimization analyses are carried out, minimizing the weighted sum of squared differences between observed and simulated values for both heads and flows, given by:

Minimize: $$F(\vec{X}) = \sum_{t=1}^{T}\Big[\sum_{nh=1}^{NH}\frac{w_{nh}\left(\frac{Hs_{nh}(t)-Ho_{nh}(t)}{Hpnt}\right)^2}{NH} - \sum_{nf=1}^{NF}\frac{w_{nf}\left(\frac{Qs_{nf}(t)-Qo_{nf}(t)}{Qpnt}\right)^2}{NF}\Big] \quad (7)$$

where $Ho_{nh}(t)$ designates the observed pressure head of the nh-th junction at time step t; $Hs_{nh}(t)$ the simulated hydraulic grade of the nh-th junction at time step t; $Hpnt$ is the hydraulic head per fitness point, which converts head differences into a dimensionless value based on the sensor's reading accuracy; NH is the number of observed heads; W_{nh} is the normalized weighting factor for observed heads; All flow-related symbols are similar as in Part I formulation.

The fittest scenario establishes the lower bound emitter, $\underline{K^n}$, for the candidate leaks as well as the possible number of leaks in the WDN. If the estimated K^n is less than the MDNL value for a candidate node it is, again, removed from the search space.

2.5 Leak Detection and Localization

At the LDL stage, the problem is formulated similarly to Part II of the SSR stage based on optimization. The process detects the leakage flow for every node location minimizing the weighted sum of differences between both head and flow observations. However, n is fixed depending on the optimum solution in Part II of SSR, while the decision variables involve the reduced list of candidate leakage nodes and the corresponding range of flows as a result of SSR. This stage can be combined with model calibration to verify the status of valves and the state of pipes. Thus, during LDL it is possible to include additional decision variables, such as candidate valve components or groups of pipe roughness to calibrate the status/setting and the friction coefficients, respectively.

3 Case Study

The methodology was tested in a real DMA in the United Kingdom (UK), with a real leak event, considering that the WDN had historically high recorded leakage levels. The system has 45km of mains and serves a rural area of over 20km^2 and approximately 1,000 properties. Flow from the source node normally varies between 2.6 l/s at Minimum Night Flow (MNF) and 8 l/s during the morning peak demand. The hydraulic model is composed of 461 pipes, 601 nodes, while there is one inlet and three Pressure Reducing Valves. In Figure 1, the EPANET model of the system is presented. The DMA has flow and pressure sensors at the inlet, and 10 inner pressure sensors, whose placement is marked using a red circle symbol. The DMA has a large variation in the elevation with the lowest point being at 28m and the highest at 221m, which results in a pressure range between 20 and 148 meters of water head and an average of 53.8m. On November 7, 2016, a leak started at 15:30 hours and lasted for around seven days. Its exact location is indicated by a blue star in Figure 1, where it was found and repaired. The effect of the leak can be observed in Figure 2, which demonstrates a significant increase in the MNF. To assess the leak localization methodology a calibrated average day model was used, i.e. a model which simulates pressures at the sensor locations with an accuracy range of ±2m relative to the observations prior to the leak start, according to the company's standards. The observed data for the leak localization part involved the time series for flows into the system and the pressures at 11 locations. The training data involved the collected pressures and flows on November 8, in order to test whether the leakage localization methodology could report the leak location and flow at an earlier stage. A total of 96 field observation data sets over 24 hour period from midnight of November 8 to midnight of November 9, have been imported into the optimization modelling tool. Each data set represents a complete snapshot of system conditions, the observed inflow into the DMA and the 11 pressures used for evaluating the quality of leak detection and localization. A systematic flow difference between 2.2 and 2.55 l/s was observed at the MNFs throughout the seven days when the leak was running, from 2.65 l/s to 5.15 l/s. Interestingly there was a large variation in demand during the morning peak hours, as the leak only caused an increase of around 1 l/s compared to the day before the leak started.

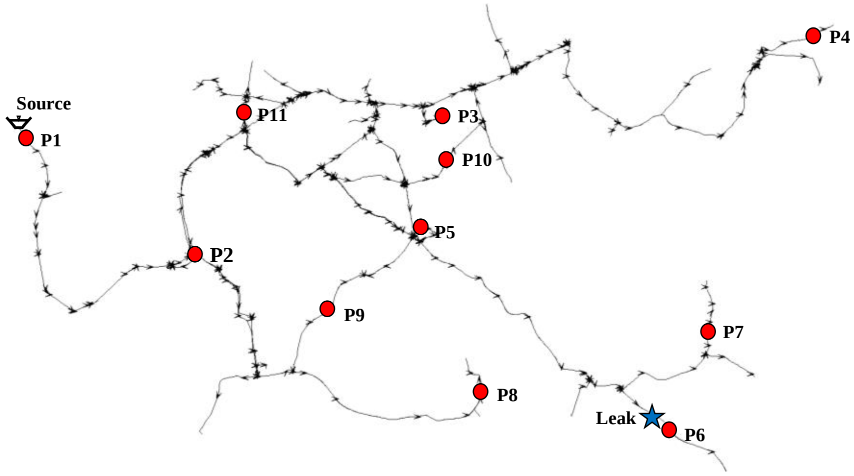

Figure 1. The District Water System and Sensor Configuration.

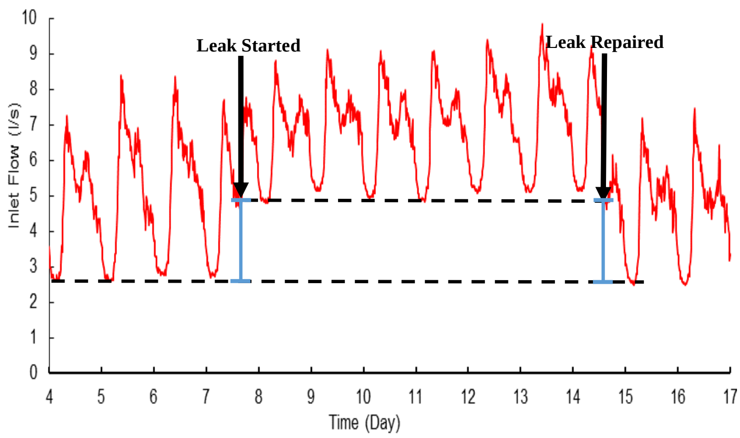

Figure 2. Inlet flow before and after leak reparation.

4 Results

Simplification reduced the search space to only 47% of network nodes (i.e. from 602 nodes to 283 nodes). Prior to any optimization run for identifying possible leakage hotspots, the system evaluation was conducted by comparing the field observed inflows with the simulated inflows of actual consumptions over 24 hours. This provided a starting point for defining the MDNL flows for each candidate leak location. The difference between average observed inflow and the average modelled inflow corresponded to losses of 27%. Considering the large variation and inconsistency in demand during the morning peak hours and throughout the day, an increased weight was given to the pressure and flow differences during the hours of low demand, i.e., between 00:00 and 06:00. Following Part I of the optimization-based search reduction a leak of 2.49 l/s was detected and the average MDNL, after eliminating the nodes with MDNL values higher than the detected leaks, was approximately 0.14 l/s, corresponding to 18 different leak scenarios. Part II showed that the most likely leak scenario was the existence of a single large leak in the WDN as this was the fittest scenario, as shown in Figure 3. The final list of candidates involved only 27% of the WDN nodes as potential leak locations (i.e. 162 nodes). LDL was, then, run for finding a single leak, while the setting of the three PRVs was also included in the optimization analysis and was allowed to adjust ±10%, to emulate the response of a PRV's setting to a real leak. The optimum solution reported a leak on the branch where the true leak occurred and 800m upstream of the true leak location (Figure 3). Considering the length of the WDN mains this represents an error of 2% (by length). Interestingly, the leak was found on a pipe section where there is no demand node and, thus, the closest node to the leak within the candidate list was 61m away from the true leak location. On the other hand, the search for the leak was narrowed down between sensors P6 and P7, which corresponds to a search space of only 10% of the total mains length. In addition, the leak was found seven days after its occurrence, which means that the approach can contribute to a much earlier detection and localization of the leak. An extended period simulation analysis was completed with the updated emitter in the hydraulic model and the inflow comparison for the leak period, between the observed flow data, and the simulated outputs before and after LDL is illustrated in Figure 4.

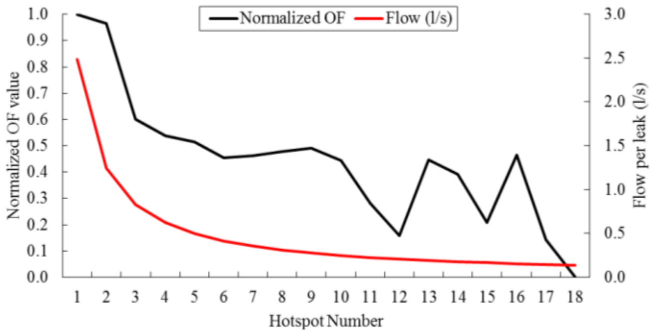

Figure 3. Part II optimization analysis outcome for the approximate hotspot number in the WDN.

The model simulated flow with the detected leakage emitter matches much better than the original modelled flow over 24 hour period. Between the hours of low demand and before the morning peak, the optimized model completely matches the observations of November 8 used to train the model. Interestingly, although the training dataset involved measurements for November 8, by comparing the updated flow data with the rest of the flows between November 9 and November 14, before the leak was found and repaired, an even better match

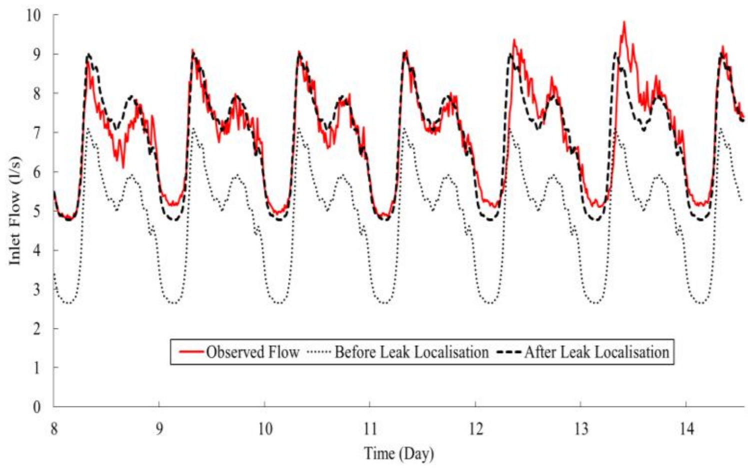

Figure 4. The Flow Differences before and after leakage detection.

is produced. The larger difference on the November 12 and 13 is due to the weekend demand, which follows a different pattern, relative to the weekdays.

5 Discussion

5.1 Industry Benefits in Leak Search Space and Time

Minimizing costs and time in finding leakages occurring in WDNs is a challenging task for water utilities. The approach used in this study helps demonstrate the benefits that could be achieved by developing an optimization-based approach. Furthermore, the next generation of models for operational uses can improve the custom-and-practice methods that have now remained broadly the same for 20 years, in both calibration and leakage localization. Although both models used prior and after leakage localization do not represent fully calibrated models, the approach shows success in narrowing down the leak search space and contributing to earlier leak localization. The search reduction methodology was able to narrow down the leak search space to within 27% candidate leakage hotspots and after LDL to within two pressure sensors, which allows for a maximum search space of 10% of the WDN (by length). Furthermore, considering that the leak flow was on average 2.49 l/s and running for a week with the associated cost of water of £1.5/m^3, there were losses of around 1,500m^3, corresponding to more than £2,000. If the offline approach was implemented to detect the leak using the data of November 8 and allowing the leak to be found and repaired by the Leakage Technicians on the afternoon of November 9, there would have been volumetric and monetary savings in water of more than 70%. The resulting model serves as a very good baseline for further calibration, following the localization of the leak in the field. A point to raise, considers the fact that PRVs follow a profile variation at their setting throughout the day. Such functionality is currently not supported by the EPANET hydraulic simulator used for leak localization. This implies that the leakage localization method developed in this paper not only facilitates earlier leakage detection but also contributes to model calibration.

5.2 Model Calibration and Data Challenges

A key requirement for accurate leak localization and message from this work is the use of well calibrated models of the WDN. With the advent of cheaper telemetry and monitoring devices, there are opportunities to further exploit the information captured, used in WDN modelling for operations at mains level. Optimisation can be a powerful tool for leakage hotspot detection. Thus, systematic approaches that leverage hydraulic models along with optimization techniques can be beneficial for WDN operations, if accompanied by good quality data. A large amount of accurate data is a necessary step for estimating calibration parameters with sufficient confidence, taking into account the increased complexity of large city WDNs and the ill-posedness problem in WDN modelling. Significant improvements in the accuracy of the model calibration and leak localization process can be secured with the inclusion of additional flow measurements captured from key flow routes in the WDN. Furthermore, the impact caused by small unknown leaks, or the local effect caused by unknown closed valves can be often insufficient to allow detection due to the measurement noise levels compared to model accuracy. However, this is a characteristic of optimization analyses when dealing with traditional WDN models. Incorporated unknown closed valves that may have been left open due to data anomalies, can be carried over during the calibration process, resulting in false positive leakage hotspot detection. This may be associated with incorrect pipe group roughness values that also exist in same flawed models in which unknown closed valves were assumed open during the model calibration process. Current WDN models are calibrated to simulate observed pressures within ±1m, whereas field pressure transducer accuracy lies within an order of magnitude less. Thus, hard-to-find leaks and topological anomalies can remain undetected due to small head losses. It would be unfortunate to forego the opportunity to move from calibration criteria of ±1.0m to ±0.1m, the latter being similar to the specified 0.1% full scale deflection for the 10bar transducers often used in the field work. By grasping this opportunity, the way will be open to the new generation modelling tools to provide far superior model calibration with increased opportunities for more successful detection of those previously undetected model anomalies.

6 Conclusions

An offline model-based leak search-space reduction approach was presented that reduces the complexity of the leak detection and localisation problem formulated as a calibration problem. Based on the case study results obtained here, the approach minimises the possibility of the optimum solution being eliminated and contributes to earlier leak localisation using an optimisation method. The method systematically reduces the number of decision variables and the range of possible values, considering the error in observations, while avoiding unnecessary simulation of solutions that do not cause any impact on model fitness. The approach has been tested on a real leak event, where the search was reduced to only 10% of the WDN (by length) and could lead to volumetric and monetary savings in water of more than 70%. A discussion is also provided on the major issues with the traditional calibration of WDN models and the ill-posedness of the calibration problem. In practice, the promising approach can leverage the use of hydraulic models for network operations and lead to significant benefits for the water industry.

7 Acknowledgements

This work is part of the first author’s STREAM Engineering Doctorate project sponsored by the UK Engineering and Physical Science Research Council, Severn Trent Water and WITS Consult Ltd.

8 References

[1] M. Ponce, L. Garza and V. Cayuela, "Model-based leak detection and location in water networks considering an extended-horizon analysis of pressure sensitivities", *J. Hyd.*, vol. 16, p. 649, 2014.

[2] R. Puust, Z. Kapelan, D. Savic and T. Koppel, "A review of methods for leakage management in pipe networks", *Ur. W. J.*, vol. 7, pp. 25-45, 2010.

[3] R. Li, H. Huang, K. Xin and T. Tao, "A review of methods for burst/leakage detection and location in water distribution systems", *W. Sc. & Tech.*, vol. 15, p. 429, 2015.

[4] R. Pérez, V. Puig, J. Pascual, J. Quevedo, E. Landeros and A. Peralta, "Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks", *Con. Eng. Pr.*, vol. 19, pp. 1157-1167, 2011.

[5] R. Pudar and J. Liggett, "Leaks in Pipe Networks", *J. Hyd. Eng.*, vol. 118, pp. 1031-1046, 1992.

[6] D. Savic, Z. Kapelan and P. Jonkergouw, "Quo vadis water distribution model calibration?", *Ur. W. J.*, vol. 6, pp. 3-22, 2009.

[7] J. Goulet, S. Coutu and I. Smith, "Model falsification diagnosis and sensor placement for leak detection in pressurized pipe networks", *A Eng Inf*, vol. 27, pp. 261-269, 2013.

[8] Z. Wu, P. Sage and D. Turtle, "Pressure-Dependent Leak Detection Model and Its Application to a District Water System", *J. W. Res. P. Man.*, vol. 136, pp. 116-128, 2010.

[9] A. Nasirian, M. Maghrebi and S. Yazdani, "Leakage Detection in Water Distribution Network Based on a New Heuristic Genetic Algorithm Model", *J. W. Res. Pr.*, vol. 5, pp. 294-303, 2013.

[10] S. Sophocleous, D. Savic, Z. Kapelan, Y. Shenand P. Sage, “CCWI2017:'A model pre-processing approach for improving calibration-based leakage detection using a genetic algorithm'”. figshare.

[11] L. Rossman, “EPANET2 users’ manual”, Risk Red. Eng. Lab. O. Res. and Dev., U.S. EPA, 2000.