Masteroppgave

NTNU
Norges teknisk-naturvitenskapelige universitet
Fakultet for ingeniørvitenskap
Institutt for bygg- og miljøteknikk

Erik Nordahl

Leak Localization with the Dual Model on a Real-World Water Distribution System

Masteroppgave i Bygg- og Miljøteknikk
Veileder: David B. Steffelbauer
Medveileder: Franz Tscheikner-Gratl
Juni 2022

NTNU
Kunnskap for en bedre verden

Erik Nordahl

Leak Localization with the Dual Model on a Real-World Water Distribution System

Masteroppgave i Bygg- og Miljøteknikk
Veileder: David B. Steffelbauer
Medveileder: Franz Tscheikner-Gratl
Juni 2022

Norges teknisk-naturvitenskapelige universitet
Fakultet for ingeniørvitenskap
Institutt for bygg- og miljøteknikk

NTNU
Kunnskap for en bedre verden

Acknowledgements

The document that lies before you is the final work of my M.Sc. in Water and Wastewater Engineering at the Norwegian University of Science and Technology. The thesis completes five years of hard work, and I am truly grateful for all the knowledge and memories I have acquired during this time.

First of all, I want to thank my primary supervisor, associate professor David Steffelbauer, for your invaluable expertise in hydroinformatics and programming. Your knowledge has significantly contributed to improving my skills in these subjects. In addition, I want to thank my second supervisor, associate professor Franz Tscheikner-Gratl, for our many professional discussions, your eyes for detail, and for your commitment and availability. I also want to thank professor Tone Merete Muthanna for arranging academic writing sessions, where your extensive feedback was greatly appreciated.

Furthermore, I want to thank all my fellow students with whom I have shared a reading room this spring. We have experienced this journey together, and we have pushed each other during this time, by providing each other with feedback and support. More importantly, you have reminded me of the importance of relaxation during stressful times. I wish you all the best for the future.

Moreover, I would like to give a special thank you to all the people who has indirectly contributed to my research. To my girlfriend, thank you for always being supportive, optimistic, and engaged in my work. In addition, I want to thank my parents and my sister for making all this possible. I will always be thankful for your unconditional love and support throughout my life.

Erik Nordahl

Trondheim
2022-06-11

Thesis Structure

Traditionally, a master thesis at NTNU results in an extensive report covering a specific subject. However, this master thesis is written in a manuscript format, following the requirements and structure of a research article. Therefore, the length of the text is limited to a maximum of 10 000 words running text from introduction to conclusion, with each figure and table accounting for 350 words.

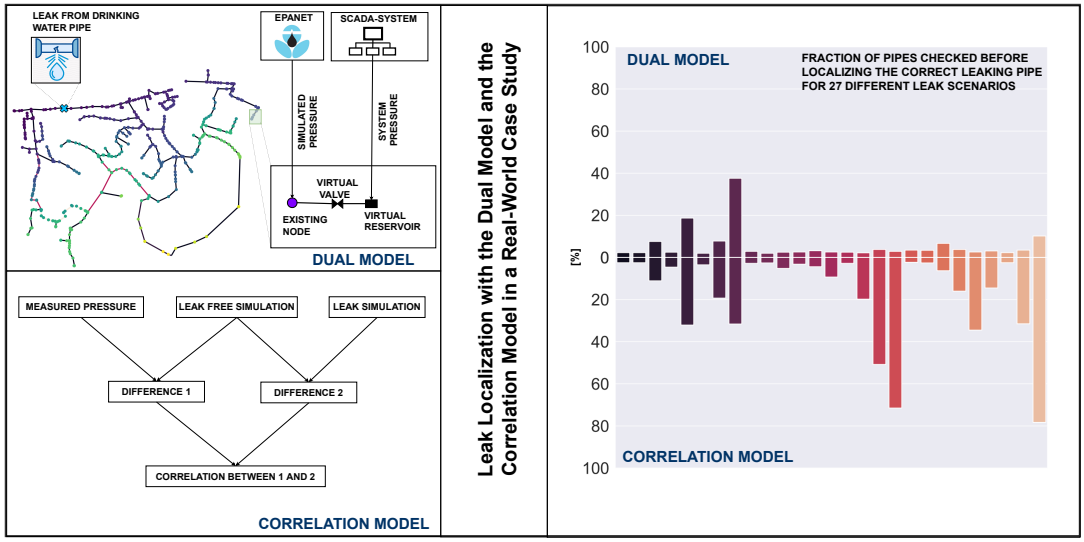
The author has chosen to write a manuscript thesis to experience the requirements and expectations of being a researcher. Furthermore, a manuscript thesis promotes accessible and available research for everyone interested in reading the thesis, which is both motivating and inspiring.

The thesis has been accepted for presentation at IWA World Water Congress & Exhibition in Copenhagen in September 2022. In addition, the thesis will (hopefully) be published in an international journal. At the moment, only David Steffelbauer (primary supervisor) and Franz Tscheikner-Gratl (second supervisor) are involved in the thesis. However, three other authors will also contribute to the manuscript after this thesis is delivered. The aim is to publish the manuscript as a paper by the end of 2022.

Graphical Abstract

Leak Localization with the Dual Model on a Real-World Water Distribution System

Erik Nordahl,
Primary supervisor: David B. Steffelbauer,
Second supervisor: Franz Tscheikner-Gratl



Research Highlights

Leak Localization with the Dual Model on a Real-World Water Distribution System

Erik Nordahl,
Primary supervisor: David B. Steffelbauer,
Second supervisor: Franz Tscheikner-Gratl

- The Dual Model can localise real leaks with different locations and magnitudes.
- The Dual Model can localise leaks without a well-calibrated model, which is a significant advantage for water utilities.
- The Dual Model shows better performance with three pressure sensors than the more commonly used Correlation Model obtains with 11 pressure sensors.
- The Dual Model's main limitations are that the model is sensitive to the leak's location in the water distribution network and that the nodal elevations must be adjusted.

Leak Localization with the Dual Model on a Real-World Water Distribution System

Erik Nordahl[a],
Primary supervisor: David B. Steffelbauer[b],
Second supervisor: Franz Tscheikner-Gratl[a]

[a]*Department of Civil and Environmental Engineering, Norwegian University of Science and Technology, S.P Andersens veg 5, Trondheim, 7031, Norway*
[b]*Hydroinformatics Group, KWB - Kompetenzzentrum Wasser Berlin gGmbH, Cicerostr. 24, Berlin, 10709, Germany*

Abstract

On average, 23 % of drinking water in Europe is lost before reaching the consumer. To reduce these losses and the negative impacts they entail, water utilities are transitioning from a passive approach where only reported leaks are fixed to actively look for leaks in their distribution systems. The model-based approach enables this transition by combining sensor data with hydraulic models and has been thoroughly investigated for the past thirty years, obtaining promising results in several different case studies. However, most of these studies have been performed either in a virtual environment or on small existing networks. This paper aims to display the possibility of reducing the leak localization time using the recently developed Dual Model on actual measurement data. The Dual Model is created by adding virtual reservoirs to existing pressure measurement nodes, transforming the pressure drop caused by leaks into virtual leak flows that amplify the leak signal. Results show that the Dual Model can localize most leaks, with 21 out of 27 leaks giving a false positive fraction below 2 %. Furthermore, the Dual Model manages to handle uncertain input parameters, obtaining promising results without a proper pipe roughness calibration and few pressure sensors. For water utilities dealing with fragmented information about the system's condition, this attribute is highly advantageous. Another model-based approach, the Correlation Model, was compared with the Dual Model for validation purposes and was generally outperformed. The main limitation of the Dual Model is the need to correct the elevation of the nodes, which was decisive

in localizing leaks smaller than 2 $\frac{L}{s}$ and that the model was sensitive to the leak location. Future work should explore why the Dual Model is sensitive to the leak's location and test the Dual Model on other real systems with different characteristics. In addition, an in depth-analysis comparing the different input parameters is needed.

Keywords: Drinking water leakages, Hydraulic modelling, Model-based leak localization, Pressure sensitivity, Water losses
PACS: 0000, 1111
2000 MSC: 0000, 1111

Sammendrag

I gjennomsnitt forsvinner 23 % av drikkevannet i Europa før det kommer frem til forbrukeren. For å redusere disse tapene og de negative konsekvensene de medfører, omstiller mange kommuner sin lekkasjehåndteringsstrategi fra å kun reparere rapporterte lekkasjer til å aktivt søke etter lekkasjer i distribusjonssystemet. Modellbaserte metoder tilrettelegger for denne omstillingen ved å kombinere data fra sensorer med hydrauliske modeller, og har de siste tretti årene blitt nøye undersøkt, samt oppnådd lovende resultater i flere forskjellige studier. Majoriteten av disse studiene har imidlertid blitt gjennomført på virtuelle nettverk, eller på små eksisterende nettverk. Denne masteroppgaven tar sikte på å demonstrere mulighetene for å redusere lekkasjelokalisasjonstiden ved å benytte den nylig utviklede Dual Modell på ekte måledata. Dual Modell består av å legge til virtuelle reservoar til eksisterende trykkmålingsnoder, med det formål om å transformere trykktapet forårsaket av lekkasjer til virtuelle vannstrømmer med tydeligere lekkasjesignal. Resultater viser at Dual Modell er kapabel til å lokalisere majoriteten av lekkasjene, med 21 av totalt 27 lekkasjer funnet med en falsk positiv fraksjon under 2 %. Det viktigste funnet er imidlertid at Dual Modell kan håndtere usikre input-parametere. Modellen oppnådde lovende resultater uten ruhetskalibrering, og med få trykksensorer, som er svært fordelaktig for kommuner med mangelfull informasjon om tilstanden på sine vannsystemer. I tillegg ble Dual Modell sammenlignet med en annen modellbasert metode, Korrelasjonsmodellen, der det kom frem at Dual Modell oppnådde bedre og mer stabile resultater. De viktigste begrensningene til modellen er behovet for å korrigere høydene på nodene, som var avgjørende for å lokalisere lekkasjer mindre enn $2\ \frac{L}{s}$, og at modellens prestasjonsevne var avhengig av lekkasjens beliggenhet. Fremtidig arbeid burde derfor undersøke hvorfor modellen er sensitiv til hvor lekkasjen befinner seg, samt teste modellen på andre vanndistribusjonssystemer med ulike egenskaper. I tillegg er det nødvendig med en grundig analyse av input-parameterne til modellen, for å bestemme og sammenligne viktigheten av disse.

Nøkkelord: Drikkevannslekkasjer, Hydraulisk modellering, Modellbasert lekkasjelokalisasjon, Trykksensitivitet, Vanntap

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CM	Correlation Model
cov	Covariance
DM	Dual Model
FP	False Positive
FSM	Fault Signature Matrix
MS	Maximum Span
TD	Topological Distance
WDN	Water Distribution Network
WNTR	Water Network Tool for Resilience

1. Introduction

European water utilities lose, on average, 23 % of valuable drinking water through leaks (EurEau, 2017), which negatively impacts the economy of the utilities. In addition, there are many environmental consequences, including the energy needed for pumping (Colombo et al., 2002) and flooding of urban areas (Hu et al., 2021). Moreover, leaks can lead to the intrusion of pollutants through holes and cracks, which can be harmful to humans (Colombo et al., 2002; Gibson et al., 2019). Nygård et al. (2007), for instance, found an increase in gastrointestinal disease following water main breaks. Furthermore, leakages might reduce the water utility's reputation as clean water with sufficient pressure often is taken for granted in developed countries (Bendz and Boholm, 2020).

An abundance of leak detection and localization strategies have been developed to reduce these negative impacts (Hu et al., 2021). The strategies can be divided into passive and active approaches (Puust et al., 2010). In a passive leakage control strategy, only reported leaks are fixed. On the contrary, an active approach involves monitoring and examining the network on a regular basis, or other proactive and predictive tools aimed at limiting the impact of leaks. Generally, active strategies lead to lower water losses, making them the preferred method in comparison with the passive approach (Farley and Trow, 2003).

Active leak localization is currently based on either acoustic methods, which utilize frequency or non-acoustic methods. The acoustic methods include listening rods, leak-noise correlators, and leak-noise loggers. The non-acoustic methods include ground-penetrating radar, gas injection, thermal infrared imaging, minimum night flow analysis, step testing, and radioactive tracers (Puust et al., 2010; Boulos and Aboujaoude, 2011; Farah and Shahrour, 2017). A common advantage of both methods is the high detection accuracy. However, these traditional techniques are considered time-consuming and labour-intensive because of the need for human operators (Adedeji et al., 2017). These methods are further limited by the weak performance observed during daytime (Puust et al., 2010). In addition, the acoustic methods are material dependent, demonstrating low efficiency in plastic pipes (Steffelbauer, 2018).

Model-based approaches try to circumvent these shortcomings in finding leaks by comparing measurement data with estimates obtained from hydraulic simulations. The approach was first introduced by Pudar and Liggett

(1992), who argued that it was possible to localize leaks using pressure measurements and formulated leak localization as an inverse problem. Since then, several different model-based leak localization methods have emerged (For a review of the methods see Hu et al. (2021)). These include; i) Error-domain model falsification (Moser et al., 2016, 2017) ii) Sensitivity matrix-based approaches (Pérez et al., 2011a; Pérez et al., 2014; Casillas et al., 2015) iii) Optimization-calibration approaches (Righetti et al., 2019; Blocher et al., 2020) and iv) Combinations of model-based and data-driven approaches (Ferrandez-Gamot et al., 2015; Soldevila et al., 2016).

The model-based approach has been thoroughly investigated because it is simple, cost-efficient, and performs well regardless of pipe material (Li et al., 2015). Furthermore, sensor technology is rapidly becoming more affordable, encouraging water utilities to install more pressure and flow sensors, which are favourable for model-based approaches (Steffelbauer, 2018). Increased computational power and rapid digitization should further motivate the transition to model-based and data-driven approaches. Another benefit rarely mentioned is the lower need for human interaction to decide if something is a leak or a normal fluctuation. Despite all these possible benefits, model-based approaches are rarely seen outside the academic environment (Steffelbauer, 2018).

There are several weaknesses with the model-based approach. First, the hydraulic model must be calibrated to ensure that the model represents the behaviour of the existing system. Model calibration is a costly, data-hungry process, (Pérez et al., 2011a), requiring expertise and knowledge (Hu et al., 2021). Second, consideration of time dependencies can be challenging. With time, roughness increases, internal pipe diameter decreases, and nodal demands continuously change (Kang et al., 2018; Adedeji et al., 2017). Therefore, hydraulic models must be updated regularly to give good results (Puust et al., 2010). In addition, the model must be reconstructed and recalibrated when the topological structure changes (Kang et al., 2018). Third, the measurements received from flow and pressure sensors can be unreliable because of outliers (Garcia et al., 2015) and noise (Steffelbauer et al., 2021).

Sensitivity-matrix-based approaches aim to overcome these obstacles by utilizing pressure measurements and sensitivity and have obtained promising results in several case studies (Pérez et al., 2011a; Pérez et al., 2014; Casillas et al., 2015). Pérez et al. (2011a) binarized the pressure residuals between leak and leak-free model simulations with measurements and leak-free model simulations. The results showed that 31 out of 42 leaks were detected in the

correct zone. However, the residuals' transformation to binarized numbers involves a loss of information. A comparative study by Pérez et al. (2011b) showed that it was more convenient to use the pressure residuals directly. The results were less affected by noise and boundary changes and were generally more robust. The method was later on tested in a real-world case study using measurement data from Barcelona, where the method proved to be effective and robust in detecting and localizing a single leak (Perez et al., 2014).

Another more recently developed model-based method is the Dual Model (Steffelbauer et al., 2020). The Dual Model introduces virtual reservoirs and valves connected to actual pressure measurement nodes (Steffelbauer et al., 2022). The pressure head in the reservoir equals the measured pressure, while the pressure in all other nodes is given from EPANET-simulations (Rossman, 1994). The pressure difference between the reservoir head and the connected nodes generates reservoir flows when the system is imbalanced to restore equilibrium conditions. These flows serve as a first indication of the leak's location and size. The Dual Model has previously obtained promising results on virtual data, achieving first place against 18 teams from all over the world at an international conference on leak detection and localization in China (Vrachimis et al., 2020). The advantages of the Dual Model are higher sensitivities to pressure changes than other models and that the leak and the system imbalances have the same unit of flow (Steffelbauer et al., 2022). However, it remains uncertain if the Dual Model can maintain these properties in a real-world case study with increased uncertainty and variability.

The majority of leak localization research has been performed either in a virtual environment or on a smaller scale, which enables operating with a well (or even perfectly) calibrated model (Zaman et al., 2020). Although this is the first step towards actual leak localization, it is not directly transferable to the real world. In addition, it has been mentioned that model-based approaches require complex modelling procedures, which in combination with offline calibrations, make them too cumbersome to be implemented in a real-world water distribution system (Kafle et al., 2022). Therefore, this paper aims to investigate the applicability and performance of the Dual Model in localizing leaks in a real-world case study. In addition, the Dual Model performance will be compared with the results obtained with the Correlation Model developed by Pérez et al. (2011b). Given that the Correlation Model previously has been tested in a real-world scenario, it can be used to validate the Dual Model performance on actual measurement data. However, the

Correlation Model was only tested with a single leak scenario with significant leak outflow (5.6 $\frac{L}{s}$). In this paper, the two models will be tested on leaks with different magnitudes and locations. Furthermore, the two models will be tested with respect to different accuracies of model calibration to see how the models perform under increasing uncertainty. Finally, the models are simulated with different sensor numbers to evaluate how the models behave with less available information.

The following section (Section 2) contains the methodology and provides a fundamental understanding of the two models. Section 3 starts with a virtual leak outflow simulation with the Dual Model, before presenting the results from the real-world case study for both models. Afterwards, the model performance of the two models is compared under different circumstances. The section is completed with a discussion of the limitations and possibilities of the Dual Model. The final section (Section 4) presents the conclusions, as well as future research topics that can contribute to the further development of the Dual Model.

2. Methodology

This section starts with a mathematical derivation of the Correlation Model and the Dual Model. The section continues with a brief introduction about the case-study area before explaining the data collection process. After that, the assumptions are identified, and the performance evaluation criteria are given.

2.1. Mathematical derivation of the Correlation Model

The Correlation Model developed by Pérez et al. (2011b) compares pressure residuals between simulated leak and leak-free scenarios with measurements and leak-free scenarios. The residuals between the measured pressure, p, and the modelled leak-free pressure, $\hat{p}$, are stored in the fault indicator vector (ϕ) (Equation 1):

$$\phi(t) = \begin{bmatrix} p_1(t) - \hat{p}_{0_1}(t) \\ \vdots \\ p_{ns}(t) - \hat{p}_{0_{ns}}(t) \end{bmatrix} \tag{1}$$

Next, leaks are simulated for every single pipe, by adding a new node with an emitter coefficient to the centre of the pipe (see Subsection 2.8 for a more detailed description of the leak generation), which results in a pressure

response at the sensors. The leak matrix is obtained, containing modelled pressures for all leak scenarios. The pressure difference between the columns in the leak matrix and the modelled leak-free pressure is stored in the fault signature matrix (FSM)(Equation 2), with *nn* representing the number of nodes and *ns* the number of pressure sensors:

$$FSM(t) = \begin{bmatrix} \hat{p}_{1_1}(t) - \hat{p}_{0_1}(t) & \dots & \hat{p}_{nn_1}(t) - \hat{p}_{0_1}(t) \\ \vdots & \ddots & \vdots \\ \hat{p}_{1_{ns}}(t) - \hat{p}_{0_{ns}}(t) & \dots & \hat{p}_{nn_{ns}}(t) - \hat{p}_{0_{ns}}(t) \end{bmatrix} \tag{2}$$

Each column of the FSM are correlated with the fault indicator vector for each time-step using Pearson's correlation coefficient formula (Equation 3). The mean correlation value for the whole simulation period is computed for each leak location, creating a vector which contains scalar correlation values related to each pipe. Notice that the words location and pipe is used interchangeably, as only one location is assumed at the middle of the pipe section, resulting in an equal amount of leak locations and pipes.

$$\rho_{\phi,FSM} = \frac{\sum_{t=end}^{t=start} \frac{cov(\phi(t),FSM_i(t))}{\sigma_{\phi(t)}\sigma_{FSM_i(t)}}}{t_{\#timesteps}} \tag{3}$$

In Equation 3, i) i is the sensitivity vector of each node, ii) *cov* is the covariance, and iii) $\sigma_{\phi(t)}$ and $\sigma_{FSM_i(t)}$ is the standard deviation of the fault signature and the fault indicator.

A simple conversion is performed on each scalar value stored in the aforementioned vector to avoid negative correlation values (Equation 4) before ranking the values in descending order. The highest value in this vector represents the leaking pipe with the highest correlation with the actual leak. Hence, the vector is sorted from most probable to least probable leak locations.

$$\rho_{pipe,id} = \frac{1 - \rho_{\phi,FSM}}{2} \tag{4}$$

2.2. Mathematical derivation of the Dual Model

The Dual Model developed by Steffelbauer et al. (2022) is created by adding virtual reservoirs and valves connected to existing pressure measurement nodes (Figure 1). The head of the reservoir for each time step equals the measured pressure plus the node elevation. This shifts the boundary condition from the fixed-demand at the sensor nodes, to the fixed-head at the

corresponding virtual reservoir. As a result, the previous boundary condition node becomes a free variable available for modelled input.

When a new leak occurs in the system, the flow towards the leakage increases, creating a pressure drop in the system; a pressure imbalance appears between the virtual reservoir head and the pressure measurement node, which is based on the leak-free model. Consequently, water flows between the reservoir and the leak-free system to restore stable conditions. This flow will then act as an amplifier for the detection and localization of leaks.

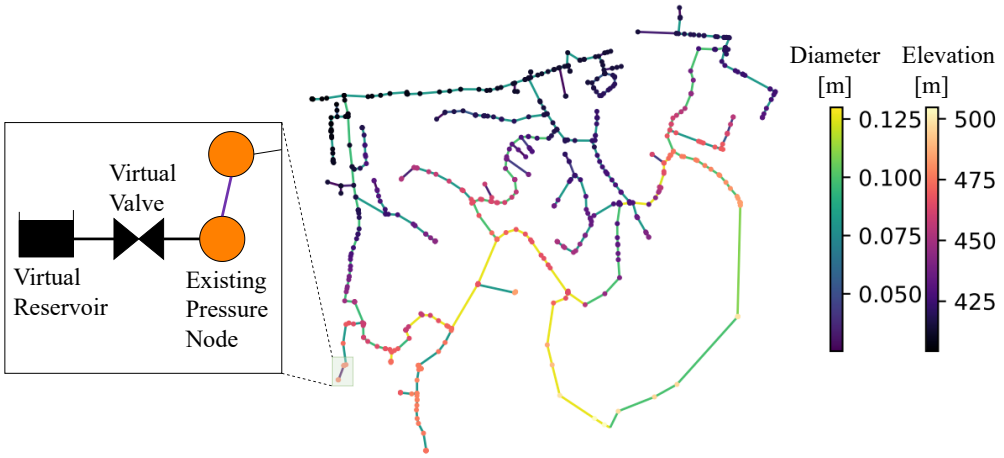


Figure 1: Dual Model Principle illustrated on the Graz-Ragnitz water distribution system (See subsection 2.3 for an overview of the network) including network characteristics for pipe diameter and node elevation.

The Dual Model is simulated with leaks at all pipes separately, with similar leak generation assumptions as for the Correlation Model. For each pipe simulation, the reservoir flow is stored for all time steps at all reservoir locations. Therefore, each pipe simulation results in a matrix containing the reservoir flow against time (Equation 5). The matrix size depends on the number of time steps (ts) and the number of virtual reservoirs (nr). The absolute value of each flow value is computed to avoid negative values because of bidirectional flow before adding it to the matrix. The number of matrices is equal to the total pipe count, with the subscript p indexing the pipe id of the current pipe.

$$\eta(t)_p = \begin{bmatrix} q_{1,1} & \cdots & q_{1,nr} \\ \vdots & & \vdots \\ q_{ts,1} & \cdots & q_{ts,nr} \end{bmatrix} \tag{5}$$

Afterwards, the average value for each column is computed, creating a single row containing the average reservoir flow for each reservoir. This procedure is conducted for each pipe, resulting in a flow matrix containing all leak-pipe scenarios (Equation 6). The matrix size depends on the number of pipes (l) and the number of pressure sensors or virtual reservoirs (nr), respectively.

$$\zeta(t) = \begin{bmatrix} q_{1,1} & \cdots & q_{1,nr} \\ \vdots & & \vdots \\ q_{l,1} & \cdots & q_{l,nr} \end{bmatrix} \tag{6}$$

The total flow for each leak location is the sum over the corresponding row. Hence, a flow vector is obtained, Q_t, which contains scalar flow values for each leak location (denoted with an i) (Equation 7).

$$Q_i(t) = \sum_{j=1}^{nr} \zeta_{i,j}(t) \tag{7}$$

The lowest value in Q_t represents the location with the highest similarities between the leak-pipe and the measurements. As a consequence, the most probable leak locations in Q_t are the pipe simulations generating the lowest total flow values. By sorting Q_t in ascending order, a list of the most likely leak locations is obtained. Note that an important distinction between the two methods is that the Correlation Model utilizes the pressure residuals directly, while the Dual Model transforms the pressure differences to virtual water flows.

2.3. Graz-Ragnitz WDN: a real-world case study

Graz-Ragnitz is a water distribution network (WDN) located in the rural areas surrounding Graz, Austria. The network is supplied from an inlet tank located south in the system. Two separate pipes are connected to this tank, but only the left pipe is open during the measurement period. The network mainly consists of plastic pipes, but some cast iron and steel pipes are present in the system. Demand data has been retrieved from billing information between 2012 and 2015 and the average consumer consumption during this

time has been allocated to the system's nodes. The water pressure varies significantly in the network because of large elevation differences between the south part (around 500 meters a.s.l) and the north part (around 400 meters a.s.l.). No pressure reduction valves have been installed in the case-study zone. Outflows are generated from 35 hydrants spread around the network (Figure 2) to simulate leak scenarios and facilitate calibration.

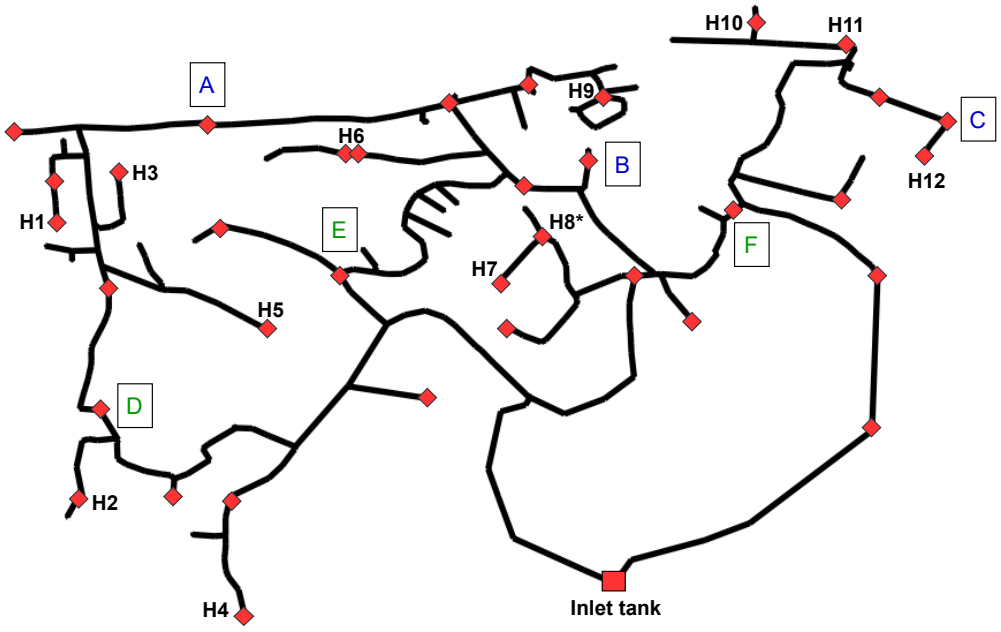


Figure 2: Graz-Ragnitz water distribution network with hydrants (red diamonds), pressure sensors (H1 to H12) and the inlet tank (red rectangle). Pressure sensor H8* malfunctioned during the measurement period. Leak locations are displayed with capital letters (A to F).

2.4. Flow and pressure monitoring system

The system inflow is measured at the tank with two devices, a Woltmann water meter and an Ultrasonic flow meter with a measurement frequency of one minute. The pressure is measured every second with twelve high precision pressure sensors of the type SEWAD 30, which has an accuracy of $\pm$ 0.2 % of the measurement value. The pressure sensor locations were determined in Steffelbauer (2018), which used optimal sensor placement algorithms to select the most appropriate locations. The tank level data is retrieved from the SCADA system of the water utility. The hydrant outflow was measured

with additional devices - Hydatlog 80mm with Storz B house coupling - magnetic inductive flowmeters with an accuracy of $\pm\,0.5\,\%$ and a measuring range from 1.5 $\frac{L}{s}$ to 60.3 $\frac{L}{s}$.

2.5. Leak scenarios

Leaks were generated by opening hydrants during the night between the $11^{th.}$ and the $12^{th.}$ of April 2016. Six different hydrants were opened during the night, marked with capital letters in Figure 2. Seventeen different leak scenarios were created; 14 of them only lasted a couple of minutes but had total leak outflows larger than 6 $\frac{L}{s}$. The three remaining scenarios had leak outflows slightly above 1 $\frac{L}{s}$, and lasted between 10 and 15 minutes. Some of the leak scenarios had two or three hydrants simultaneously opened. These leaks provided the possibility to increase the number of leak scenarios by dividing them into sub-scenarios, searching for each of these locations separately. The approach for these leaks was to add the outflow as extra demand for all leak locations except one, making them single leak scenarios with higher complexity. A complete list for all leak scenarios is given in Table 1.

Table 1: All leak scenarios simulated as single leak scenarios. The leak locations are shown in Figure 2. The asterix * indicates that the leak is simulated as extra demand at this location. To exemplify: leak S3a is located at leak location A with an outflow of 4.97 $\frac{L}{s}$, the other two leak outflows at leak location B and C are assumed to be known.

Leak scenario	Leak start time [H:M:S]	Leak end time [H:M:S]	A $\frac{L}{s}$	B $\frac{L}{s}$	C $\frac{L}{s}$
S1	01:37:45	01:40:30	15.03	–	–
S2a	01:47:30	01:49:45	8.06	4.11*	–
S2b	01:47:30	01:49:45	8.06*	4.11	–
S3a	01:58:00	01:59:30	4.97	1.14*	5.10*
S3b	01:58:00	01:59:30	4.97*	1.14	5.10*
S3c	01:58:00	01:59:30	4.97*	1.14*	5.10
S4	02:04:00	02:06:00	–	11.54	–
S5a	02:09:15	02:11:45	–	2.62	7.05*
S5b	02:09:15	02:11:45	–	2.62*	7.05
S6	02:15:45	02:18:15	–	–	12.03
S7a	02:22:15	02:25:15	6.93	–	7.06*
S7b	02:22:15	02:25:15	6.93*	–	7.06
Leak scenario	Leak start time [H:M:S]	Leak end time [H:M:S]	D $\frac{L}{s}$	E $\frac{L}{s}$	F $\frac{L}{s}$
S8	03:23:45	03:26:30	7.60	–	–
S9a	03:31:30	03:33:30	1.81	6.06*	–
S9b	03:31:30	03:33:30	1.81*	6.06	–
S10a	03:36:45	03:39:30	1.27	5.08*	5.05*
S10b	03:36:45	03:39:30	1.27*	5.08	5.05*
S10c	03:36:45	03:39:30	1.27*	5.08*	5.05
S11	03:44:00	03:45:45	–	15.97	–
S12a	03:51:45	03:54:45	–	6.00	6.00*
S12b	03:51:45	03:54:45	–	6.00*	6.00
S13a	03:58:30	04:01:00	2.20	–	6.02*
S13b	03:58:30	04:01:00	2.20*	–	6.02
S14	04:04:30	04:07:45	–	–	9.07
Leak scenario	Leak start time	Leak end time	Leak position		$\frac{L}{s}$
S15	02:30:00	02:45:00	A		1.18
S16	02:57:00	03:10:00	E		1.14
S17	04:12:00	04:23:00	F		1.34

2.6. Measurement adjustments

The pressure sensors report a value every second. These measurements are resampled to produce an average value with a given time step to reduce the impact caused by noise and consumption variations. It was found that resampling the measurements to produce one value every minute increased stability while maintaining a sufficient number of measurements. The tank level measurements report one value every minute, and was adjusted using second-order splines. It was also discovered that the internal clock installed in the sensors did not display the correct time. The sensors were therefore corrected individually. The errors ranged from a few seconds to $\pm$ 1 minute. Sensor H8 reported atypical values and was therefore neglected.

2.7. Creation and calibration of the water distribution system

The water distribution network model was created in EPANET (Rossman, 1994); it is partly skeletonised and consists of 650 nodes and 658 pipes. In comparison, the real network has approximately 1300 pipes and nodes. The total pipe length in the model is 10.2 kilometres, and the diameters of the pipes range from 70 to 400 mm and is unchanged from the real network. Hence, the model elements are pretty similar to the real network, but the model input parameters (e.g. demand and roughness) must be adjusted.

Hydraulic model calibration has for decades been one of the most important research topics in water distribution systems (Savic et al., 2009). The purpose of the model calibration is to estimate model parameters and analyse different types of uncertainty, providing the model with best-fit parameters (Jun et al., 2022). Previous research has suggested that a well-calibrated model is essential to localize leaks, as the results heavily rely on the model's ability to mirror the real-world system (Hutton et al., 2014). However, obtaining a well-calibrated model with low uncertainty can be difficult. First, municipalities often lack information about their water distribution systems (e.g. pipe age and pipe replacement history), increasing the complexity of the model calibration (Scheidegger et al., 2013). Second, drinking water pipes cannot be visually inspected, as is typical for wastewater systems (Kleidorfer et al., 2013). Third, measured values for pressure and flow contain outliers and noise, leading to a calibration bias (Walski, 1983). Therefore, a robust model able to handle uncertain input parameters is desired. To test the model's ability to handle uncertainty, three different model calibrations were created. All of these three model calibrations utilize the Darcy-Weisbach

formula to determine head loss. The three network calibrations contain different levels of information regarding the pipe roughnesses and the nodal elevations:

- **C1** is the least calibrated model. The roughness coefficient was set to 0.005 mm for all pipes, which is a common friction factor for plastic pipes (Porto, 2006). The elevations were unchanged from the information provided by the utility. No other model adjustments were conducted.

- **C2** includes nodal elevation adjustments but is otherwise utterly similar to C1. The nodal elevations for the 11 functioning pressure sensors were adjusted to produce zero reservoir flow in a leak-free situation, with the purpose of amplifying the leak signal.

- **C3** has the same nodal elevations as C2 but contains, in addition, an extensive pipe roughness calibration and estimation of a minor-loss coefficient caused by a partially closed valve. After the measurement period, it was confirmed that the water utility had forgotten to fully open a valve after maintenance work. This partially-closed valve would influence leak localization performance. Steffelbauer (2018) used differential evolution algorithms to localize and approximate the minor loss caused by this partially-closed valve. The minor-loss coefficient was estimated to be 1384, a substantial minor-loss coefficient. Lippacher (2018) found that it was convenient to group the pipes into five different groups related to their position in the system. Each pipe group were given a roughness value using optimization techniques (Appendix C.14 shows the roughness values found in Lippacher (2018)). Calibration C3 combines the findings from the two articles mentioned above, making it one of the best calibrations currently developed for this network.

The demand was calibrated by comparing the measured inflow with the simulated inflow between 1:22 and 1:31 a.m. on the $12^{th.}$ of April 2016. During this time, there were no leaks, making it possible to compare the minimum night flow in the model and the actual system. It was found that multiplying the demand retrieved from the billing information with 0.85 for all nodes gave the closest agreement between the model and the measurements, resulting in an error margin of 0.35 %. The demand calibration was equal for all three model calibrations.

In addition to calibration, another main challenge of model-based approaches is the difference between the number of pressure sensors installed compared to the number of possible leak locations (Steffelbauer et al., 2022). A simple solution to this problem would be to increase the number of sensors, but this is not a feasible solution as pressure sensors are relatively expensive. For water utilities with limited resources, this challenge can be an obstacle which prevents transformation from classic to modern techniques. Therefore, an essential task with model-based leak localization is to develop models that manage to function with fewer sensors and hence less available information. The Dual Model and the Correlation Model were first tested with information gathered from all 11 sensors before testing with 5 and 3 sensors. The sensor placements with 5 and 3 sensors did not involve reallocating the sensors, making these sensor placements sub-optimal.

2.8. Model leak generation, model assumptions, and model performance evaluation parameters

The Dual Model and the Correlation Model were programmed in Python (Van Rossum and Drake Jr, 1995). The python package WNTR (Water Network Tool for Resilience) was used to simulate the leaks with an emitter coefficient. The emitter coefficient was chosen because it gives a more realistic representation of the leak than modelling it as extra demand. With an emitter coefficient, the size of the leak is directly dependent on the pressure in the node, making it a variable instead of a constant size (Perez et al., 2014). The leaks were simulated at the middle of the pipes' total length by splitting the pipe into two separate sections connected to a new node. Hence, it was chosen to simulate the leak on the pipe section despite the actual leaks occurring at hydrants. The previously mentioned decision is justified because leaks usually occur at pipes or pipe fittings (Steffelbauer, 2018). The leak discharge is assumed to be known, as this value can be easily found by analysing the minimum night flow (Meseguer et al., 2014). Therefore, the new node was given an emitter coefficient producing a similar outflow as the measured leak. The leak scenarios are simulated as extended period simulations (a series of steady-state simulations) in EPANET (Rossman, 1994).

To evaluate the leak localization performance, several different metrics are used. The first is *TD*, the topological distance between the actual leak and the pipe with the highest leak correlation measured, computed using Dijkstra's shortest path algorithm (Dijkstra, 1959). The second is *FP*, the

fraction of false positive pipes. Finally, *MS* is the maximum shortest path distance between all FP-pipes.

3. Results and Discussion

3.1. Dual Model virtual leak outflow simulation

The Dual Model was first simulated without any leaks present (Figure 3 and 4). The pressure difference between the nodes in the leak-free model and the measured pressure in the reservoirs leads to reservoir flows. The Dual Model should obtain similar total reservoir flow as the system inflow in a leak-free situation if the calibration is perfect. Hence, the reservoir flow in the Dual Model can be used to improve model calibration by changing the nodal elevations and consequently changing the reservoir head, which in hand changes the flow pattern. Note that the Dual Model reservoir flow and inflow differ in a leak-free scenario. The inflow stays relatively constant at around 1.2 $\frac{L}{s}$ during the night (Figure 3a and 4), while the Dual Model flow fluctuates around zero in a leak-free situation at first (Figure 3b). This was done intentionally (in calibration C2 and C3) to amplify the leak signal, as the Dual Model could not localize leaks with smaller outflow than the minimum night flow consumption. However, Figure 4 shows that the Dual Model later on in the night give the same minimum night flow pattern as the inflow does. Hence, the Dual Model no longer give zero flow in a leak-free situation. It is probable to believe that this change is caused by increasing water demand in the early morning hours.

Note that some reservoirs give negative flow patterns to some of the leaks (Figure 3c and 4c), meaning that water flows from the reservoir to the system. It was found that these negative flows did not negatively affect model performance but rather provided useful information contributing to the leak localization. Figure 3a, Figure 3b, and Figure 3c, show similar behaviour between the Dual Model and the system inflow. In addition, these patterns illustrate how the Dual Model amplifies signals, as can be seen by comparing the sharp spike seen in Figures 3b and 4b for each scenario with the slight pressure drop seen in Figure 3d and 4d.

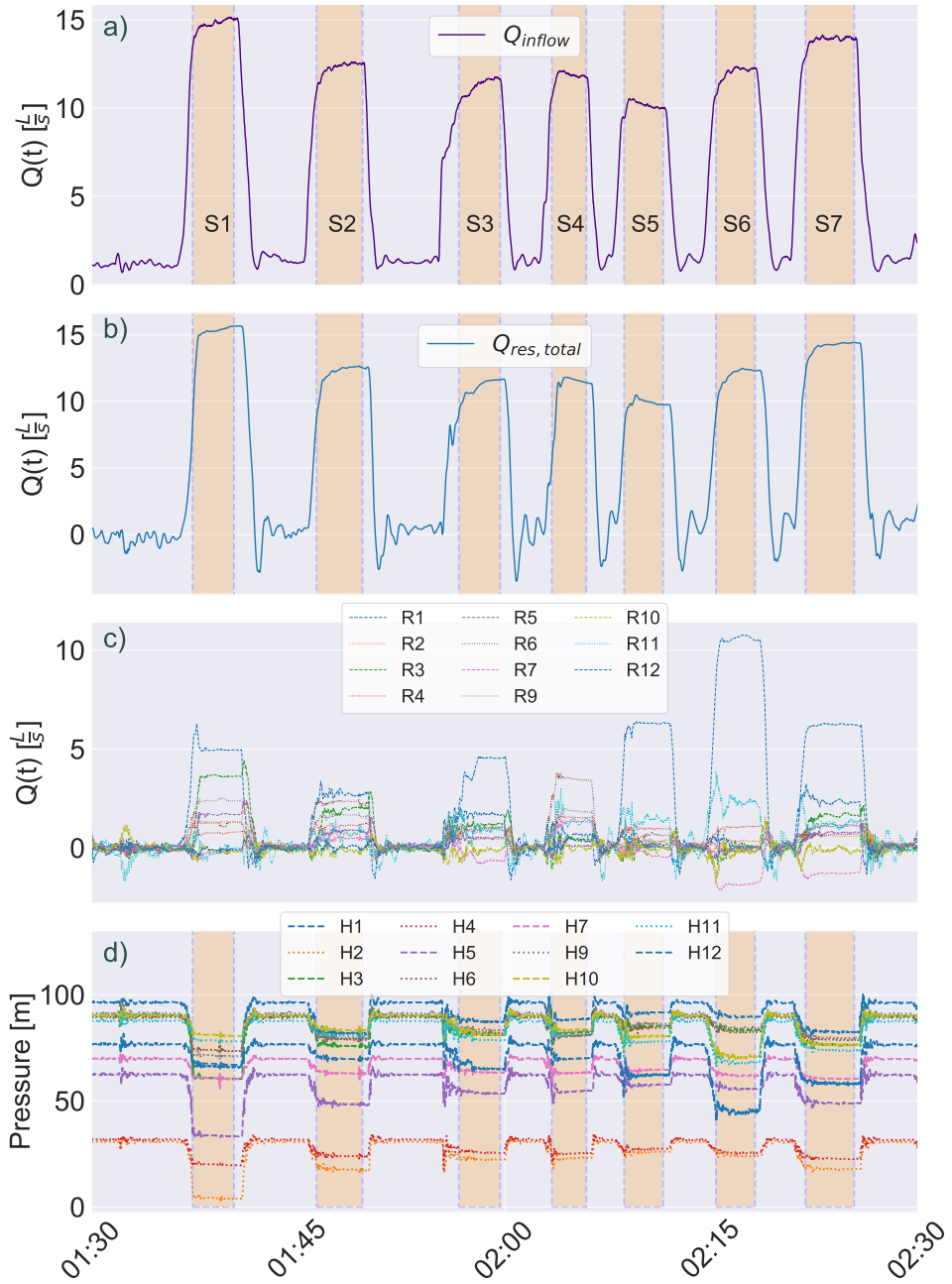


Figure 3: Dual Model leak-free simulation for leak scenario 1 to 7. a) Inflow measured at the system inlet b) Total reservoir flow for all virtual reservoirs simulated with the Dual Model c) Dual Model flow for each virtual reservoir d) Measured pressure.

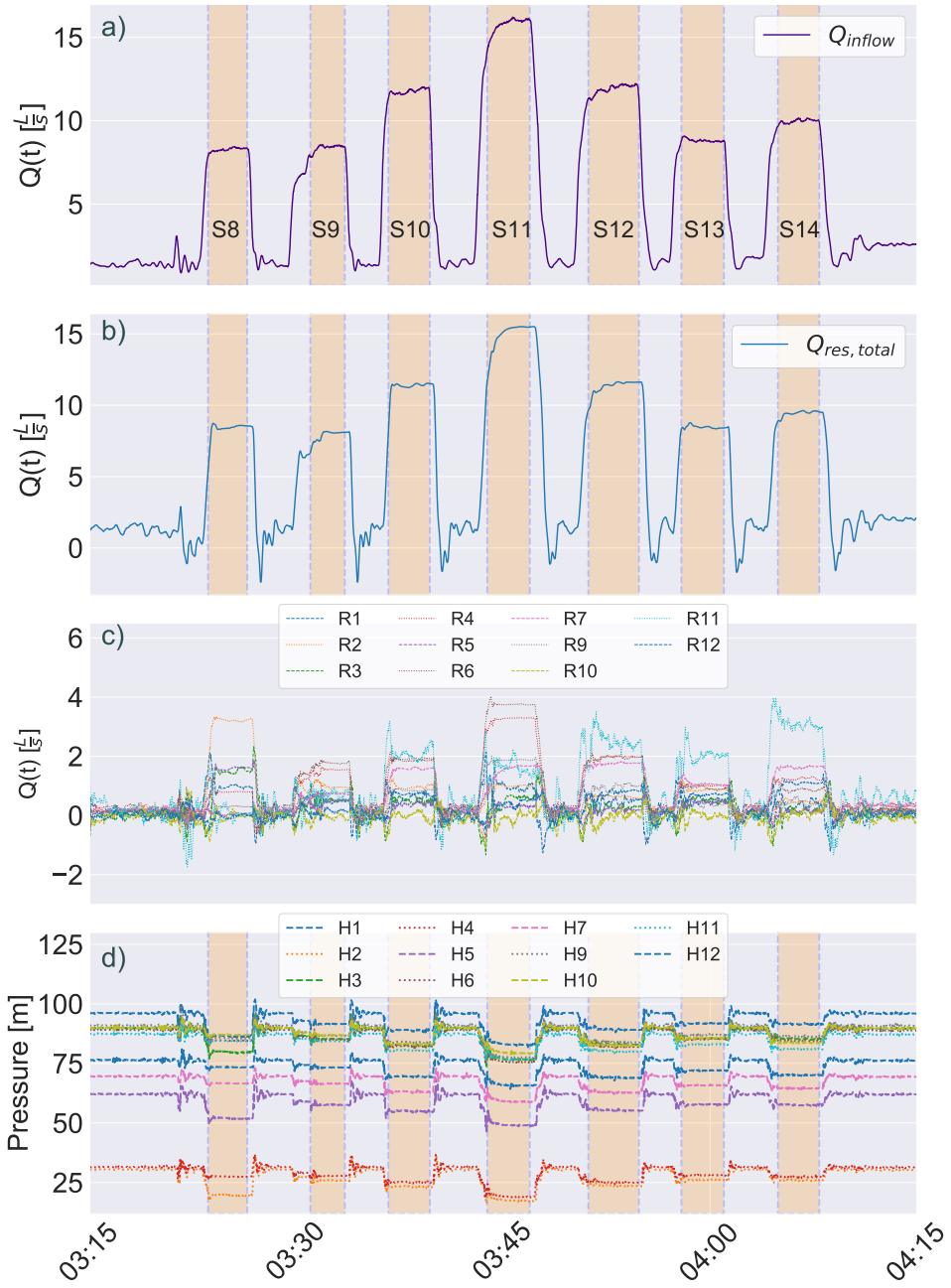


Figure 4: Dual Model leak-free simulation for leak scenario 8 to 14. a) Inflow measured at the system inlet b) Total virtual reservoir flow for all reservoirs simulated with the Dual Model c) Dual Model flow for each virtual reservoir d) Measured pressure.

3.2. Real-world case study results

Table 2 gives the leak localization performance for the real-world case study. The results show that the Dual Model can locate most leaks with decent accuracy, with 21 out of 27 leaks obtaining an FP-fraction below 2 %. The topological distance and the maximum span parameter give short distances for almost all of these leaks, indicating that the Dual Model only suggests leak locations close to the actual location. An exception is leak scenario 10c, which is the only leak obtaining a low FP-fraction (0.9 %), combined with a high topological distance (634 m) and maximum span (657 m). The search area is therefore not necessarily reduced for this leak scenario.

Four out of the six remaining leaks give FP-fractions between 4 % and 10 %, while the last two leaks give even higher FP-fractions. The topological distance and the maximum span parameters further emphasize the weaker performance observed in these scenarios. These six leaks are located at either leak location B (scenario S2b, S3b, S4 and S5a) or leak location F (S12b and S17). All leaks occurring at leak location B are difficult to localize with the Dual Model. It is worth noting that the closest pressure measurement node to leak location B was sensor H8. As previously mentioned, this sensor gave unreliable results and was therefore not included in the simulation of the models. In addition, position B is located at an unmeasured dead-end branch. The six mentioned leaks do not share any other properties than the location, as they range from small leaks (S2b = 1.14 $\frac{L}{s}$) to pipe bursts (S4 = 11.54 $\frac{L}{s}$), and from distinct leak locations to multiple simultaneous leak locations. Therefore, these results indicate that the Dual Model is sensitive to where the leak occurs in the system and where the sensors are placed with respect to the leak.

On a general note, the Dual Model demonstrates an ability to perform well for different leak types at several (but not all) different leak locations in a real-world water distribution network. The model is not significantly affected by the size of the leak, obtaining decent results for both small and large leaks. Furthermore, the Dual Model can localize single leaks in a situation with several leaks co-existing. The aforementioned can be an essential asset of the model, as the high leak percentages observed in European countries indicate that several leaks might be present simultaneously.

The Correlation Model shows varying performance for the different leak scenarios. Approximately a third of the leaks are localized with an FP-fraction below 2 %, while 16 out of 27 leaks obtain an FP-fraction lower than 10 %. Since hydraulic models are meant to reduce the search space before

applying pinpointing techniques, these results are decent. However, many of these leaks, for instance, leaks S3a, S7a, and S12b give high maximum span values even though the FP-fraction is low. Hence, the model suggests pipes distant from each other, indicating that the search area might be more significant than expected. Furthermore, it struggles to localize leak scenarios where several hydrants are opened simultaneously. The model performance is particularly weak for leak scenario 10, where all three sub-scenarios are poorly localized. The method is unable to localize leaks smaller than 3 $\frac{L}{s}$, managing to find only one of the six smallest leaks. Previous findings from Meseguer et al. (2014) showed that the model was able to approximate the leak location when the leak size was around 6 $\frac{L}{s}$, and also showed promising performance for leaks as small as 4 $\frac{L}{s}$. However, smaller leaks were not considered. In general, the Correlation Model performance is weaker than the Dual Model performance, being slightly outperformed for the scenarios most straightforward to locate and strongly outperformed for the more challenging leak scenarios.

Table 2: Leak localization results for all leak scenarios with the best calibrated model (C3) using all 11 sensors.

	Dual Model			Correlation Model		
Leak scenario	TD [m]	MS [m]	FP [%]	TD [m]	MS [m]	FP [%]
S1	2	1	0.3	19	18	0.5
S2a	1	1	0.3	60	57	0.5
S2b	195	411	5.6	308	839	9.1
S3a	60	57	0.5	452	881	2.6
S3b	1050	2145	16.7	1295	1912	30.1
S3c	0	0	0	55	87	1.5
S4	308	427	5.8	241	928	17.3
S5a	964	2145	35.6	729	2352	29.6
S5b	35	54	0.9	15	34	0.8
S6	0	0	0	15	14	0.6
S7a	19	18	0.5	438	916	3.2
S7b	15	14	0.6	438	916	3.2
S8	39	90	1.2	99	194	2.4
S9a	90	26	0.6	177	462	7.3
S9b	68	118	0.5	88	312	0.8
S10a	14	0	0.2	553	765	17.8
S10b	64	151	1.8	611	1666	48.8
S10c	634	657	0.9	620	2282	69.5
S11	21	193	1.5	21	19	0.3
S12a	43	72	1.4	68	292	0.6
S12b	300	931	4.7	91	514	4.3
S13a	62	137	1.8	177	724	14.0
S13b	14	19	0.6	460	2282	32.5
S14	290	299	1.1	167	668	12.5
S15	19	17	0.3	19	18	0.5
S16	68	113	1.5	734	1554	29.5
S17	789	1531	8.2	1488	2352	76.3

3.3. Model performance with differently calibrated models

The Dual Model and the Correlation Model were simulated with three different model calibrations to analyse their ability to handle uncertain input

parameters. It was found that the Dual Model did not need a well-calibrated model to give good results for large ($> 7\ \frac{L}{s}$) single leak scenarios, as the FP-fraction was almost identical with the three different calibrations (Figure 5a). Furthermore, the Dual Model gave a lower average topological distance for the least calibrated situation compared to the calibrated and the partly calibrated network for the large single leak scenarios, which is illustrated in Figure 5c. Given that large leaks are capable of causing significant damage, this is a promising result for water utilities dealing with uncertainty, as it reduces the need for model calibration. Model calibration requires knowledge and expertise, which can be difficult (or expensive) for municipalities to obtain.

The performance with the least calibrated and the partly calibrated hydraulic model worsens when all leak scenarios are included (Figure 5b and 5d). The reduction in performance is especially significant for the minor leaks. Leak scenarios S15, S16, and S17, respectively, have an increase in the FP-fraction of 2.7 %, 18.4 %, and 8.2 % from C3 to C2, and even larger from C2 to C1. Hence, the Dual Model relies more on a well-calibrated model for the minor leak scenarios. The leak scenarios with multiple hydrants opened simultaneously are not as affected by excluding the roughness calibration as long as the nodal elevations are adjusted, maintaining almost the same performance from C3 to C2 (see appendix Table B.6 for the results with the model calibrated for elevations). However, when the elevations are not corrected, the multi-leak scenarios are difficult to localize (See Appendix Table B.7 for the results with the least calibrated model). To summarise, the Dual Model can localize single large leaks without a proper calibration but struggles with leak outflows around 1-1.5 $\frac{L}{s}$ and is reliant on adjusting the elevations when the complexity of the leak scenario increases. The Dual Model's sensitivity to the node elevation is further discussed in Subsection 3.5, along with other limitations of the Dual Model.

The Correlation Model relies on a well-calibrated model to give good results for all types of leaks, even the large pipe bursts. In the best scenario when the hydraulic model is well-calibrated, the results are decent but not as good as the results obtained with the Dual Model. However, the performance rapidly deteriorates when the model deviates more from the real network. In contrast to the Dual Model, the Correlation Model is unaffected by elevation adjustments, showing similar performance for both the least calibrated and the partly calibrated network. However, a minority of the leaks are found close to the actual leak location with these calibrations. In addition, the per-

formance is equally poor for single and multi-leak scenarios (Appendix Table B.7). Furthermore, the topological distance and maximum span values are much higher compared to the well-calibrated model. Hence, the Correlation Model cannot tolerate uncertain model input parameters. Similar behaviour has been observed in the past; Meseguer et al. (2014) concluded that the Correlation Model relies on an accurate hydraulic model both in the parameters (e.g. roughness) and in the topological structure to obtain good leak localization results. Therefore, these results imply that the Dual Model's ability to amplify the leak signal makes the leak localization more robust, which corresponds well with previous findings about the model (Steffelbauer et al., 2022).

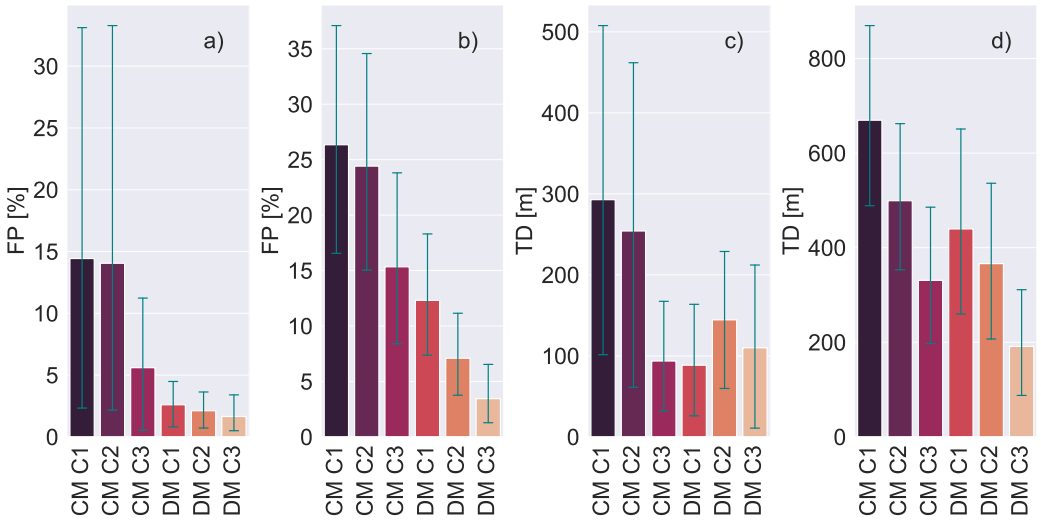


Figure 5: Average model performance with different calibrations. C1 is the least calibrated network, C2 is calibrated for elevation, and C3 is calibrated for roughness, minor-loss, and elevation. a) Average false positive fraction for the large single leak scenarios, b) Average false positive fraction for all leak scenarios, c) Average topological distance for the large single leak scenarios, d) Average topological distance for all leak scenarios. CM - Correlation Model, DM - Dual Model. 95 % confidence interval marked with a thin teal-coloured line (common for all figures with a confidence interval).

3.4. Model performance with 11, 5 and 3 sensors

It was found that the Dual Model was able to maintain model performance for most of the leaks with 5 and 3 sensors combined with a well-calibrated model (Figure 6). The results for the large single leak scenarios were completely unaffected by reducing the number of sensors, (see Appendix Table

B.8 for a full list of the results with 5 sensors and Appendix Table B.9 for the results with 3 sensors.) maintaining or even obtaining even better results than with 11 sensors. The multi-leak scenarios had a slight increase in the FP-fraction for most of the leaks, but several leak scenarios also had better results with 5 sensors than 11.

The main reason the average results are poorer with fewer sensors is that some of the small multi-leak scenarios are significantly worse localized. For instance, leak scenario S9a has an FP-fraction with 11 sensors of 0.6 %, compared to 16.7 % for 5 sensors and 17.6 % for 3 sensors. Leak scenario S10a shows almost a linear increase in the FP-fraction, obtaining an FP-fraction of only 0.2 % with 11 sensors, which increases to 18.8 % with 5 sensors and further increases to 33.1 % with 3 sensors. Common for these two leaks is that they are very small, with respectively a leak outflow of 1.81 (S9a) and 1.27 $\frac{L}{s}$ (S10a). In addition, both leak scenarios have several hydrants opened simultaneously. On the bright side, the single leak scenarios with small outflows do not share these problems. Surprisingly, with three sensors, the results are better for leak scenario S15-S17 than with 11 sensors. Hence, reducing the number of sensors makes the Dual Model performance weaker only for the small leak scenarios with multiple hydrants opened. In summary, the Dual Model show good performance with few sensors, but struggles with the most challenging leak scenarios with 5 and 3 sensors.

The Correlation Model shows the same development as the Dual Model, with a gradual decrease in the model performance with fewer sensors. However, the Correlation Model initially shows worse performance, as DM3 outperforms CM11. The small leak outflows strongly influence the difference between the Correlation Model and the Dual Model, as the Correlation Model is unable to localize leaks smaller than 3 $\frac{L}{s}$ no matter the number of sensors. For instance, leak scenario S17 has an FP-fraction of 76.3 % with 11 sensors, 48.5 % with 5 sensors, and 60.2 % with 3 sensors. These numbers indicate that increasing the number of sensors probably would not help this model localize leaks with small magnitudes. The multi-leak scenarios also showed a gradual decrease in model performance when reducing the number of sensors. On the positive side, the localization of the large single leak scenarios showed no apparent difference between 3, 5 and 11 sensors. These results comply with previous findings in Perez et al. (2014), where it was concluded that the Correlation Model performance showed no significant improvement when increasing the number of sensors from 5 to 7 when a single leak with size 6 $\frac{L}{s}$ were considered.

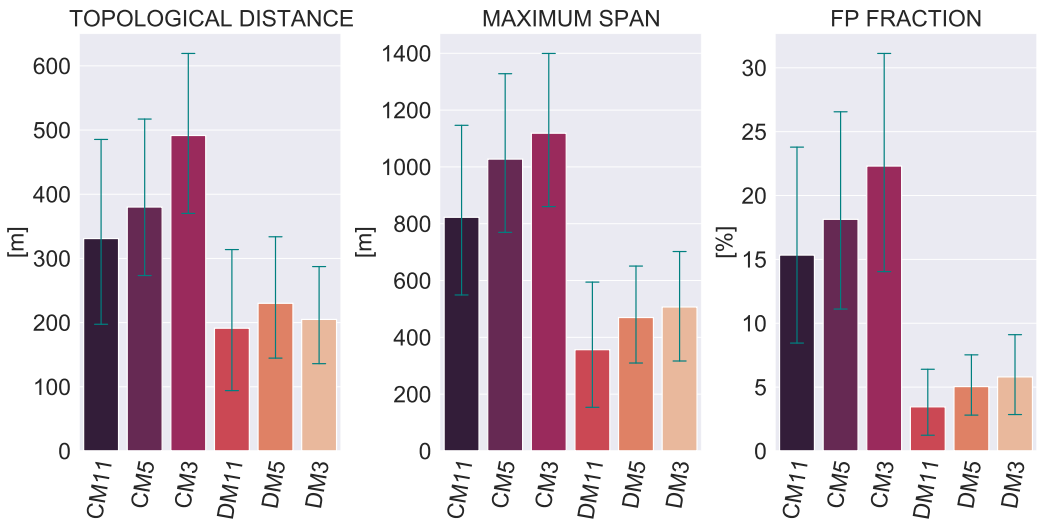


Figure 6: Average model performance with 11, 5 and 3 sensors for the Correlation Model (CM) and the Dual Model (DM) using network calibration C3.

3.5. Limitations and possibilities with the Dual Model

Eleven simulations were performed for each leak scenario with a random parameter included in the elevation to test the Dual Model's dependency of node elevation. The random parameter was drawn from a normal distribution with a standard deviation of 1. The average of these simulations is depicted in Figure 7, showing that the model performance remarkably worsened for all single leak scenarios. The small leak scenarios had the most noticeable difference between correct and random elevations, with FP-fractions increasing on average between 40 and 50 % (detailed results can be found in Appendix Table B.10).

The random elevation parameter affects the virtual reservoir flow, creating virtual reservoir flows in a leak-free situation. These flows serve as a noise parameter in the model, which negatively influences the model performance. The small leaks are often smaller than the flows caused by this random elevation parameter, making the leaks near impossible to localize. The model maintains acceptable performance for most large leak scenarios, although these also show decreasing leak localization precision. In addition, Figure 7 only includes the single leak scenarios. The model performance would probably deteriorate even stronger for the leak scenarios with several open hydrants. Therefore, the Dual Model's main limitation is the need to adjust the elevation coordinates for the reservoir nodes. The aforementioned

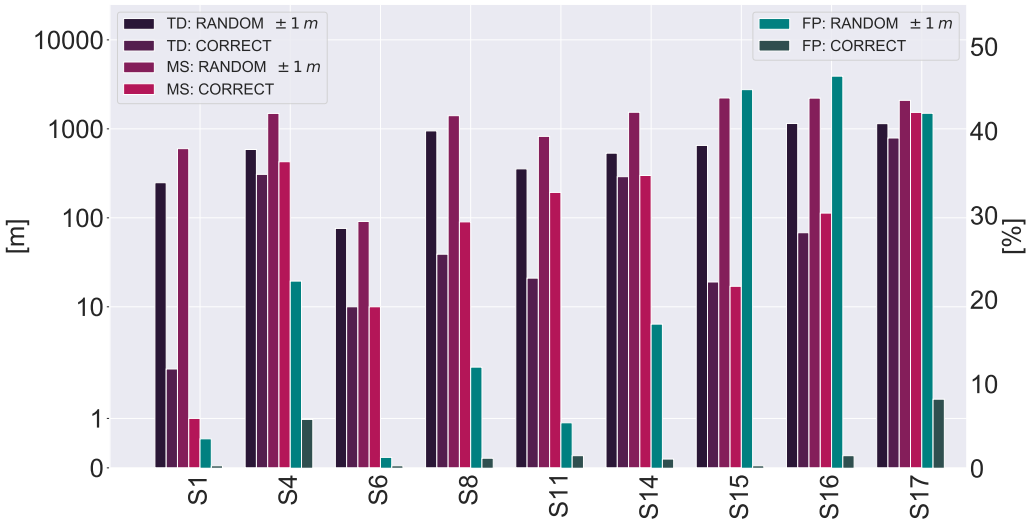


Figure 7: Dual Model performance with a random error elevation parameter included in the reservoir elevation. The random parameter is drawn from a normal distribution with $\sigma = 1$.

adjustments should be made even if these differ from the actual elevation of the nodes. The elevation must be changed to produce zero flow in a leak-free situation to enable the localization of small leaks.

Other limitations of the Dual Model include the dependency of a well-calibrated model if the leak is small (1-2 $\frac{L}{s}$) and the reduced precision seen at specific leak locations. Given the short measurement interval (< 5 minutes) used to localize the leaks, these results might improve with more extended measurement periods. Another possible solution would be to conduct a simple calibration of the pipe roughnesses, which should not be too time-consuming, considering that roughness intervals already exist for different types of pipe materials. The Dual Model has, as mentioned, only been tested for single-leak scenarios, even when several leaks co-existed. These leak scenarios obtained slightly poorer results than those containing only one leak. However, it is probable to believe that the Dual Model performance would improve in a multi-leak scenario if the leaks added as extra demand instead were modelled with an emitter coefficient.

The Dual Model was not tested with a poorly calibrated model in combination with few sensors. Given that more sensors are needed with increased uncertainty (Steffelbauer and Fuchs-Hanusch, 2016), this might be a limitation of the model. However, the limitations of the Dual Model already

show the significant possibilities of this model, as; i) elevation adjustments is relatively straightforward and rapid to rectify, ii) simple calibrations are likely to solve the issues with localizing small leaks in a multi-leak scenario, and iii) leak location dependencies might be solved with increased knowledge about optimal sensor placement.

The advantages of the Dual Model's ability to perform under uncertain conditions should not be underestimated. The Dual Model showed robust behaviour even with a deficient calibration, indicating that the model is able to operate with erroneous data. In addition, it maintained model performance with few sensors, which demonstrates that the model can function with small amounts of available data. Many water utilities around the world struggle with data scarcity and lack of knowledge (Scheidegger et al., 2015), emphasizing the importance of a model being able to operate with uncertainty. Furthermore, the significant leak percentages seen in Europe indicate that several leaks co-exist in water distribution systems. Many methods, including the Correlation Model, cannot search for several leaks simultaneously. In contrast, the Dual Model is formulated in a way that allows searching for several distinct leaks occurring at the same time. Although this has not yet been tested, it is, at the minimum, a possible advantage of this model.

Previous research on the Dual Model questioned how the model would perform with limited demand information and few pressure sensors (Steffelbauer et al., 2022). The findings of this paper suggest that the Dual Model is (almost) unaffected by detailed information about water consumption (Appendix: Figure B.11 and Table B.11). In addition, the model performance is not significantly affected by reducing the sensor numbers for the large leak scenarios. However, a negative influence is seen for the most complex leak scenarios that combine small outflows with several simultaneous leak locations.

4. Conclusions and Recommendations

Localizing leaks is essential for water utilities struggling with significant water losses. This paper aimed to test the Dual Model's ability to localize leaks in a real-world case study for the first time. It was found that the Dual Model could localize actual leaks with different locations and magnitudes, with 19 out of 27 leaks being located in close proximity (< 100 m) to the actual leak location. In addition, the Dual Model was able to localize both small leaks ($1\ \frac{L}{s}$) and single leaks in a situation with several leaks present at the same time. The Correlation Model did not show the same performance and struggled with localizing all leaks smaller than $3\ \frac{L}{s}$.

The most important finding of the Dual Model is the model's ability to handle uncertain input parameters, maintaining acceptable performance with a weaker calibrated model. The model managed to localize most leaks without roughness calibration, even obtaining slightly better results for the large single leak scenarios with all pipes given the same roughness coefficient. However, the model struggled with localizing minor leaks without calibration. The performance was especially poor without adjusting the nodal elevations. In contrast to the Dual Model, the Correlation Model could not tolerate uncertain input parameters, showing rapidly deteriorating model performance without a well-calibrated model.

The Dual Model performance was almost unaffected by reducing the number of sensors for the large single leak scenarios. In addition, the small single leak scenarios also maintained acceptable performance at most leak locations. However, the most complex leak scenarios were difficult to localize with few sensors, indicating that more sensors is needed when the complexity of the leak increases. The Correlation Model was utterly unaffected by increasing the number of sensors for the large single leak scenarios but very reliant on the sensor count for the more complex leak scenarios. In general, the Dual Model significantly outperformed the Correlation Model with few sensors, obtaining better results with only 3 sensors than the Correlation Model did with 11 sensors.

The Dual Model's capabilities of handling uncertainty is a crucial asset making it a promising tool for water utilities. Given the many uncertainties water utilities struggle with (e.g. lack of pipe replacement history, incomplete register for pipe characteristics, and limited information about water consumption), this benefit is highly needed. In addition, the model seems to be cost-efficient, as calibration can be time-consuming and pressure sensors

are relatively expensive.

The Dual Model's main limitation is the need to change the elevation of the pressure measurement nodes to produce zero virtual reservoir flow in a leak-free situation. It was found that the model did not manage to locate small leaks without these adjustments. The small leaks were also notably more challenging to localize with an uncalibrated model, indicating that the Dual Model might need a simple pipe roughness calibration to locate leaks smaller than 1.5 $\frac{L}{s}$. Finally, the model was sensitive to where the leak was located in the system, obtaining poor performance for all leak scenarios at one leak location (position B).

The future research needed to improve the Dual Model could centre around understanding why the Dual Model is sensitive to the leak's location and elevation. If these challenges can be solved, the Dual Model provides a robust and relatively straightforward tool for leak detection and localization. Further work should also focus on optimal sensor placement and applying the Dual Model to other real systems with different network characteristics. An in-depth sensitivity analysis comparing the different input parameters (e.g. elevation, roughness, demand, number of sensors) would also increase knowledge and understanding of the Dual Model's behaviour, which is necessary if the Dual Model is to be implemented as a valuable tool for water utilities in the future.

Bibliography

Adedeji, K., Hamam, Y., Abe, B., Abu-Mahfouz, A., 2017. Towards achieving a reliable leakage detection and localization algorithm for application in water piping networks: An overview. IEEE Access 5, 20272–20285. doi:10.1109/ACCESS.2017.2752802.

Bendz, A., Boholm, Å., 2020. Indispensable, yet invisible: Drinking water management as a local political issue in swedish municipalities. Local Government Studies 46, 800–819. doi:10.1080/03003930.2019.1682557.

Blocher, C., Pecci, F., Stoianov, I., 2020. Localizing leakage hotspots in water distribution networks via the regularization of an inverse problem. Journal of Hydraulic Engineering 146, 04020025. doi:10.1061/(asce)hy.1943-7900.0001721.

Boulos, P., Aboujaoude, A., 2011. Managing leaks using flow step-testing, network modeling, and field measurement. Journal (American Water Works Association) 103, 90–97.

Casillas, M., Garza-Castañón, L., Puig, V., Vargas-Martinez, A., 2015. Leak signature space: An original representation for robust leak location in water distribution networks. Water (Switzerland) 7, 1129–1148. doi:10.3390/w7031129.

Colombo, A., Karney, B., Asce, M., 2002. Energy and costs of leaky pipes: Toward comprehensive picture. Journal of Water Resources and Planning Management 128, 441–450. doi:10.1061/ASCE0733-94962002128:6441.

Dijkstra, E., 1959. A note on two problems in connexion with graphs. Numerische mathematik 1, 269–271.

EurEau, 2017. Europe's water in figures: An overview of the european drinking water and waste water sectors. 2017 edition. Brussels, Belgium: EurEau The European Federation of National Associations of Water Services, p. 22.

Farah, E., Shahrour, I., 2017. Leakage detection using smart water system: Combination of water balance and automated minimum night flow. Water Resources Management 31, 4821–4833. doi:10.1007/s11269-017-1780-9.

Farley, M., Trow, S., 2003. Losses in water distribution networks: a practitioner's guide to assessment, monitoring and control. 1. ed., IWA.

Ferrandez-Gamot, L., Busson, P., Blesa, J., Tornil-Sin, S., Puig, V., Duviella, E., Soldevila, A., 2015. Leak localization in water distribution networks using pressure residuals and classifiers. IFAC-PapersOnLine 28, 220–225. doi:10.1016/j.ifacol.2015.09.531.

Garcia, D., Gonzalez, D., Quevedo, J., Puig, V., Saludes, J., 2015. Water demand estimation and outlier detection from smart meter data using classification and big data methods leak detection in pipelines view project fault tolerant control view project. In Proc., 2nd New Developments in IT & Water Conf., 1-8. London: IWA Publishing.

Gibson, J., Karney, B., Guo, Y., 2019. Predicting health risks from intrusion into drinking water pipes over time. Journal of Water Resources Planning and Management 145, 04019001. doi:10.1061/(asce)wr.1943-5452.0001039.

Hu, Z., Chen, B., Chen, W., Tan, D., Shen, D., 2021. Review of model-based and data-driven approaches for leak detection and location in water distribution systems. Water Supply 21, 3282–3306. doi:https://doi.org/10.2166/ws.2021.101.

Hutton, C., Kapelan, Z., Vamvakeridou-Lyroudia, L., Savić, D., 2014. Dealing with uncertainty in water distribution system models: A framework for real-time modeling and data assimilation. Journal of Water Resources Planning and Management 140, 169–183. doi:10.1061/(asce)wr.1943-5452.0000325.

Jun, S., Arbesser-Rastburg, G., Fuchs-Hanusch, D., Lansey, K., 2022. Response surfaces for water distribution system pipe roughness calibration. Journal of Water Resources Planning and Management 148. doi:10.1061/(asce)wr.1943-5452.0001518.

Kafle, M., Fong, S., Narasimhan, S., 2022. Active acoustic leak detection and localization in a plastic pipe using time delay estimation. Applied Acoustics 187, 108482. doi:10.1016/j.apacoust.2021.108482.

Kang, J., Park, Y., Lee, J., Wang, S., Eom, D., 2018. Novel leakage detection by ensemble cnn-svm and graph-based localization in water distribu-

tion systems. IEEE Transactions on Industrial Electronics 65, 4279–4289. doi:10.1109/TIE.2017.2764861.

Kleidorfer, M., Möderl, M., Tscheikner-Gratl, F., Hammerer, M., Kinzel, H., Rauch, W., 2013. Integrated planning of rehabilitation strategies for sewers. Water Science and Technology 68, 176–183. doi:10.2166/wst.2013.223.

Li, R., Huang, H., Xin, K., Tao, T., 2015. A review of methods for burst/leakage detection and location in water distribution systems. Water Science and Technology: Water Supply 15, 429–441. doi:10.2166/ws.2014.131.

Lippacher, J., 2018. Methoden der kalibrierung von trinkwasserverteilnetzen und deren einfluss auf die modellbasierte leckagelokalisierung. Graz, Austria: Master Thesis, Graz University of Technology.

Meseguer, J., Mirats-Tur, J., Cembrano, G., Puig, V., Quevedo, J., Pérez, R., Sanz, G., Ibarra, D., 2014. A decision support system for on-line leakage localization. Environmental Modelling and Software 60, 331–345. doi:10.1016/j.envsoft.2014.06.025.

Moser, G., Paal, S., Jlelaty, D., Smith, I., 2016. An electrical network for evaluating monitoring strategies intended for hydraulic pressurized networks. Advanced Engineering Informatics 30, 672–686. doi:10.1016/j.aei.2016.09.003.

Moser, G., Paal, S., Smith, I., 2017. Measurement system design for leak detection in hydraulic pressurized networks. Structure and Infrastructure Engineering 13, 918–928. doi:10.1080/15732479.2016.1225312.

Nygård, K., Wahl, E., Krogh, T., Tveit, O., Bøhleng, E., Tverdal, A., Aavitsland, P., 2007. Breaks and maintenance work in the water distribution systems and gastrointestinal illness: A cohort study. International Journal of Epidemiology 36, 873–880. doi:10.1093/ije/dym029.

Pérez, R., Cugueró, M., Cugueró, J.and Sanz, G., 2014. Accuracy assessment of leak localisation method depending on available measurements. Procedia Engineering 70, 1304–1313. doi:10.1016/j.proeng.2014.02.144.

Perez, R., Sanz, G., Puig, V., Quevedo, J., Escofet, M., Nejjari, F., Meseguer, J., Cembrano, G., Mirats Tur, J., Sarrate, R., 2014. Leak localization

in water networks: A model-based methodology using pressure sensors applied to a real network in barcelona [applications of control]. IEEE Control Systems 34, 24–36. doi:10.1109/MCS.2014.2320336.

Porto, R., 2006. Hidráulica básica. 4. ed. Edição Projeto REENGE, EESC/USP.

Pudar, R., Liggett, J., 1992. Leaks in pipe networks. Journal of Hydraulic Engineering 118, 1031–1046. doi:10.1061/(ASCE)0733-9429(1992)118:7(1031).

Puust, R., Kapelan, Z., Savic, D., Koppel, T., 2010. A review of methods for leakage management in pipe networks. Urban Water Journal 7, 25–45. doi:10.1080/15730621003610878.

Pérez, R., Puig, V., Pascual, J., Quevedo, J., Landeros, E., Peralta, A., 2011a. Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks. Control Engineering Practice 19, 1157–1167. doi:https://doi.org/10.1016/j.conengprac.2011.06.004.

Pérez, R., Quevedo, J., Puig, V., Nejjari, F., Cugueró, M., Sanz, G., Mirats, J., 2011b. Leakage isolation in water distribution networks: A comparative study of two methodologies on a real case study. 2011 19th Mediterranean Conference on Control and Automation, MED 2011 19, 138–143. doi:10.1109/MED.2011.5982979.

Righetti, M., Bort, C., Bottazzi, M., Menapace, A., Zanfei, A., 2019. Optimal selection and monitoring of nodes aimed at supporting leakages identification in wds. Water (Switzerland) 11, 126. doi:10.3390/w11030629.

Rossman, L., 1994. EPANET Users Manual : Project Summary. Cincinnati, OH :U.S. Environmental Protection Agency, Risk Reduction Engineering Laboratory.

Savic, D., Kapelan, Z., Jonkergouw, P., 2009. Quo vadis water distribution model calibration? Urban Water Journal 6, 3–22. doi:10.1080/15730620802613380.

Scheidegger, A., Leitão, J., Scholten, L., 2015. Statistical failure models for water distribution pipes - a review from a unified perspective. Water Research 83, 237–247. doi:10.1016/j.watres.2015.06.027.

Scheidegger, A., Scholten, L., Maurer, M., Reichert, P., 2013. Extension of pipe failure models to consider the absence ofdata from replaced pipes. Water Research 47, 3696–3705. doi:10.1016/j.watres.2013.04.017.

Soldevila, A., Blesa, J., Tornil-Sin, S., Duviella, E., Fernandez-Canti, R., Puig, V., 2016. Leak localization in water distribution networks using a mixed model-based/data-driven approach. Control Engineering Practice 55, 162–173. doi:10.1016/j.conengprac.2016.07.006.

Steffelbauer, D., Blokker, E., Buchberger, S., Asce, M., Knobbe, A., Abraham, E., 2021. Dynamic time warping clustering to discover socioeconomic characteristics in smart water meter data. Journal of Water Resources and Planning Management 147. doi:10.1061/(ASCE)WR.1943.

Steffelbauer, D., Deuerlein, J., Gilbert, D., Abraham, E., Piller, O., 2022. Pressure-leak duality for leak detection and localization in water distribution systems. Journal of Water Resources Planning and Management 148. doi:10.1061/(ASCE)WR.1943-5452.0001515.

Steffelbauer, D., Deuerlein, J., Gilbert, D., Piller, O., Abraham, E., 2020. Dual model for leak detection and localization. doi:10.5281/zenodo.3923907ï. battLeDIM 2020 -Battle of the Leakage Detection and Isolation Methods Team: Under Pressure. CCWI / WDSA 2020, pre-conference workshop BattleDIM, Beijing, China.

Steffelbauer, D., Fuchs-Hanusch, D., 2016. Efficient sensor placement for leak localization considering uncertainties. Water Resources Management 30, 5517–5533. doi:10.1007/s11269-016-1504-6.

Steffelbauer, D.B., 2018. Model-Based Leak Localization in Water Distribution Systems. Ph.D. thesis. Graz University of Technology.

Van Rossum, G., Drake Jr, F.L., 1995. Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam.

Virtanen, P., Gommers, R., Oliphant, T., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S., Brett, M., Wilson, J., Millman, K., Mayorov, N., Nelson, A., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen,

I., Quintero, E., Harris, C., Archibald, A., Ribeiro, A., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors, 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods 17, 261–272. doi:10.1038/s41592-019-0686-2.

Vrachimis, S., Eliades, D., Taormina, R., Ostfeld, A., Kapelan, Z., Liu, S., Kyriakou, M., Pavlou, P., Qiu, M., Polycarpou, M., 2020. "Dataset of BattLeDIM: Battle of the leakage detection and isolationmethods."InProc., 2nd Int CCWI/WDSA Joint Conf. Kingston, ON,Canada: Queen's Univ.

Walski, T., 1983. Technique for calibrating network models. Journal of Water Resources Planning and Management 109, 360–372. doi:10.1061/(asce)0733-9496(1983)109:4(360).

Zaman, D., Tiwari, M., Gupta, A., Sen, D., 2020. A review of leakage detection strategies for pressurised pipeline in steady-state. Engineering Failure Analysis 109, 104264. doi:10.1016/j.engfailanal.2019.104264.

Appendix A. Leak scenarios

Table A.3: Leak scenario 1 to 14 (the large leak scenarios). A, B, C, D, E and F marks the location of the leak according to Figure 2.

Leak scenario	Leak start time [H:M:S]	Leak end time [H:M:S]	A $\frac{L}{s}$	B $\frac{L}{s}$	C $\frac{L}{s}$
S1	01:37:45	01:40:30	15.03	–	–
S2	01:47:30	01:49:45	8.06	4.11	–
S3	01:58:00	01:59:30	4.97	1.14	5.10
S4	02:04:00	02:06:00	–	11.54	–
S5	02:09:15	02:11:45	–	2.62	7.05
S6	02:15:45	02:18:15	–	–	12.03
S7	02:22:15	02:25:15	6.93	–	7.06
Leak scenario	Leak start time [H:M:S]	Leak end time [H:M:S]	D $\frac{L}{s}$	E $\frac{L}{s}$	F $\frac{L}{s}$
S8	03:23:45	03:26:30	7.60	–	–
S9	03:31:30	03:33:30	1.81	6.06	–
S10	03:36:45	03:39:30	1.27	5.08	5.05
S11	03:44:00	03:45:45	–	15.97	–
S12	03:51:45	03:54:45	–	6.00	6.00
S13	03:58:30	04:01:00	2.20	–	6.02
S14	04:04:30	04:07:45	–	–	9.07

Table A.4: Information about leak scenario 15 to 17 (the small leak scenarios).

Leak scenario	Leak time	Leak position	Leak size $\frac{L}{s}$
S15	02:30:00 - 02:45:00	A	1.18
S16	02:57:00 - 03:10:00	E	1.14
S17	04:12:00 - 04:23:00	F	1.34

Table A.5: Leak scenarios with only one leak.

Leak scenario	Leak start time [H:M:S]	Leak end time [H:M:S]	Leak position see figure 2	Leak size $[\frac{L}{s}]$
S1	01:37:45	01:40:30	A	15.03
S4	02:04:00	02:06:00	B	11.54
S6	02:15:45	02:18:15	C	12.03
S8	03:23:45	03:26:30	D	7.60
S11	03:44:00	03:45:45	E	15.97
S14	04:04:30	04:07:45	F	9.07
S15	02:30:00	02:45:00	A	1.18
S16	02:57:00	03:10:00	E	1.14
S17	04:12:00	04:23:00	F	1.34

Appendix B. Results

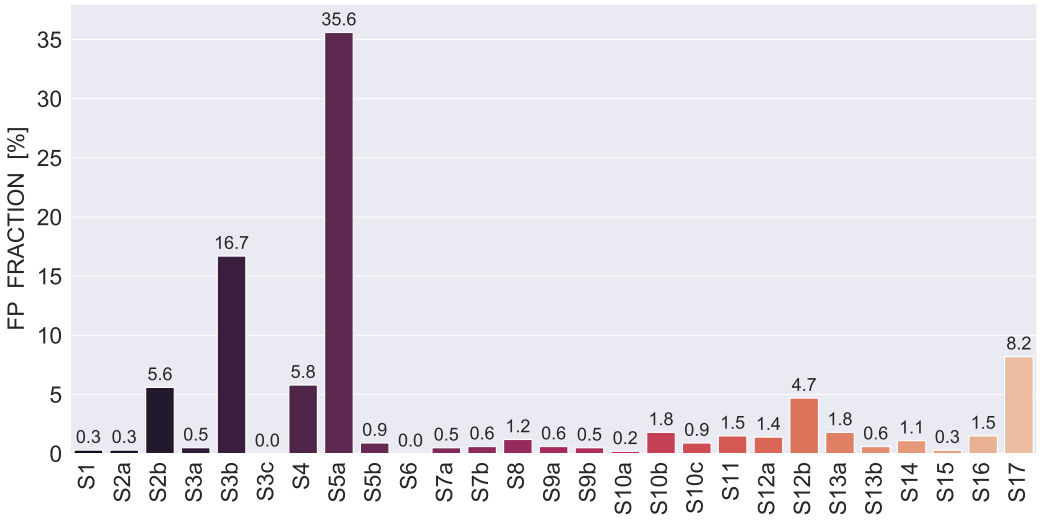


Figure B.8: Dual Model fraction of false positive pipes for all leak scenarios with a well-calibrated model (C3). These results are also depicted in Table 2 on page number 20 in the paper.

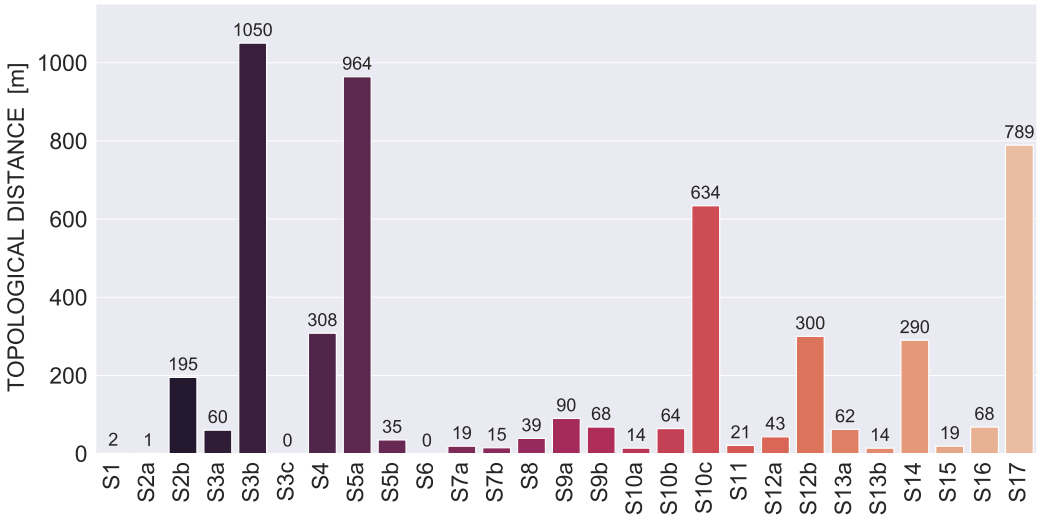


Figure B.9: Dual Model topological distance performance for all leak scenarios with a well-calibrated model (C3). These results are also depicted in Table 2 on page number 20 in the paper.

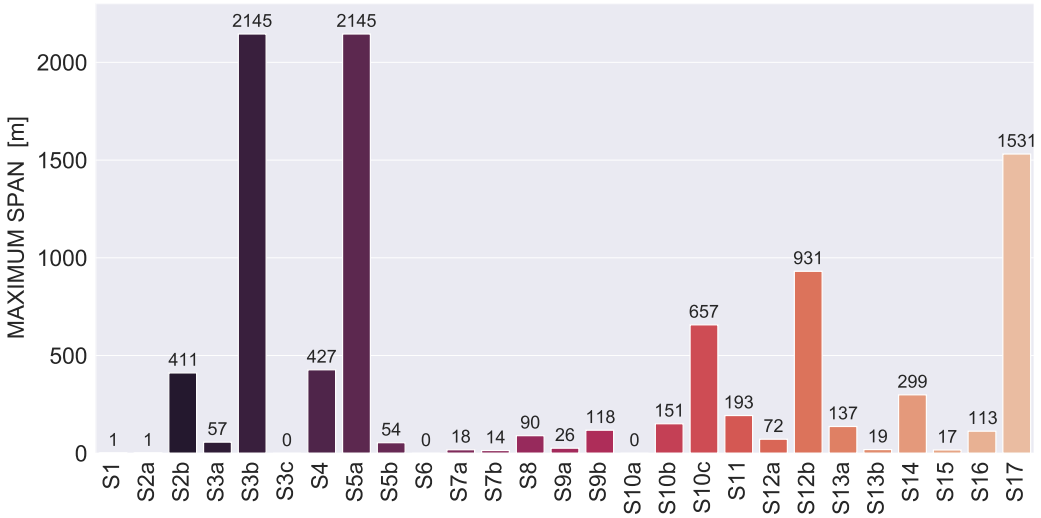


Figure B.10: Dual Model maximum span performance for all leak scenarios with a well-calibrated model (C3). These results are also depicted in Table 2 on page number 20 in the paper.

Table B.6: Leak localization results with demand from billing information, model calibration C2 (only elevation calibrated), and 11 sensors used.

	Dual Model			Correlation Model		
Leak scenario	TD [m]	MS [m]	FP [%]	TD [m]	MS [m]	FP [%]
S1	257	257	0.5	648	854	10.5
S2a	79	78	0.6	682	947	13.8
S2b	1291	1458	11.2	1458	2282	51.4
S3a	23	0	0.2	354	610	15.0
S3b	1003	2352	30.7	1122	2282	71.6
S3c	0	0	0	73	72	1.4
S4	136	353	4.3	551	1703	58.4
S5a	1287	2352	43.5	1298	2352	56.2
S5b	35	66	1.1	35	72	1.4
S6	15	7	0.3	35	54	0.9
S7a	19	18	0.5	354	634	13.4
S7b	15	34	0.8	73	72	1.4
S8	156	603	5.3	36	35	0.3
S9a	296	764	8.2	145	265	3.8
S9b	38	104	2.1	347	1984	6.1
S10a	379	830	7.0	516	651	16.9
S10b	662	973	7.6	761	1703	47.0
S10c	310	368	1.2	520	2282	72.3
S11	13	27	1.1	21	66	0.5
S12a	539	809	9.7	21	197	2.0
S12b	155	873	7.6	241	618	8.7
S13a	249	753	6.5	516	765	17.8
S13b	14	51	1.1	520	2282	59.0
S14	290	316	1.2	234	715	13.7
S15	244	221	3.0	682	879	11.4
S16	797	2131	19.9	797	1415	24.8
S17	1588	2119	16.4	1441	2352	79.6

Table B.7: Leak localization results with demand from billing information, model calibration C1 (uncalibrated), and 11 sensors used.

	Dual Model			Correlation Model		
Leak scenario	TD [m]	MS [m]	FP [%]	TD [m]	MS [m]	FP [%]
S1	2	1	0.3	648	879	10.6
S2a	19	57	0.5	707	978	14.1
S2b	1158	1458	12.2	1459	2282	51.2
S3a	493	762	5.5	470	641	15.3
S3b	956	2272	29.6	1122	2282	71.6
S3c	8	7	0.3	73	72	1.4
S4	118	286	2.9	551	1771	59.0
S5a	353	2188	19.6	1459	2352	54.7
S5b	15	14	0.6	15	14	0.6
S6	0	0	0.0	15	34	0.8
S7a	426	634	3.2	372	662	14.9
S7b	8	0	0.2	73	72	1.4
S8	249	603	5.2	36	73	0.8
S9a	404	2145	14.6	246	254	2.3
S9b	761	1399	6.5	304	2009	11.4
S10a	1978	2101	20.5	533	647	16.4
S10b	761	1071	10.2	744	1703	46.5
S10c	126	1916	13.7	1613	2297	75.2
S11	38	53	1.4	68	66	0.5
S12a	142	1097	9.3	923	2297	8.8
S12b	113	1811	16.9	602	816	14.7
S13a	341	764	7.8	468	931	17.9
S13b	99	1996	12.6	1654	2297	72.2
S14	124	486	5.8	439	816	14.9
S15	1402	2297	75.7	925	1554	28.3
S16	1337	2207	37.5	902	1481	20.5
S17	439	2197	19.9	1654	2297	85.3

Table B.8: Leak localization results using only 5 sensors, network calibration (C3), and demand from billing information.

	Dual Model			Correlation Model		
Leak scenario	TD [m]	MS [m]	FP [%]	TD [m]	MS [m]	FP [%]
S1	19	17	0.3	57	57	0.5
S2a	1	0	0.2	60	1082	0.8
S2b	168	215	2.7	253	351	5.0
S3a	23	0	0.2	97	878	5.5
S3b	241	1567	11.9	178	1912	28.1
S3c	209	236	3.5	301	422	2.1
S4	308	384	4.0	578	832	16.1
S5a	485	1543	21.1	905	2145	52.6
S5b	209	264	1.4	238	230	4.3
S6	130	236	1.5	234	310	3.5
S7a	1	0	0.2	438	916	5.5
S7b	196	202	2.9	301	373	2.0
S8	1	0	0.2	258	931	10.3
S9a	591	724	16.7	591	651	15.3
S9b	244	243	1.2	108	895	4.0
S10a	643	943	18.8	516	1400	23.6
S10b	173	749	4.0	799	2282	62.5
S10c	460	606	2.4	406	2282	74.8
S11	21	189	1.4	180	116	1.2
S12a	180	318	3.8	561	1363	19.3
S12b	81	557	6.4	36	325	3.6
S13a	86	520	2.6	442	830	18.1
S13b	65	64	0.3	460	2282	33.0
S14	46	627	1.5	431	870	15.0
S15	0	0	0	60	95	0.6
S16	493	750	10.9	312	1559	33.9
S17	1138	1729	16.0	1461	2352	48.5

Table B.9: Leak localization results using only 3 sensors, network calibration (C3), and demand from billing information.

	Dual Model			Correlation Model		
Leak scenario	TD [m]	MS [m]	FP [%]	TD [m]	MS [m]	FP [%]
S1	19	57	0.5	60	57	0.5
S2a	23	22	0.3	97	1156	2.7
S2b	285	695	15.5	430	771	15.5
S3a	23	0	0.2	452	916	5.6
S3b	509	1668	22.2	857	1912	27.1
S3c	208	236	3.2	301	480	2.4
S4	87	384	2.6	875	1243	26.3
S5a	485	1397	17.8	327	2145	65.3
S5b	174	215	3.2	238	401	7.0
S6	15	7	0.3	408	480	6.2
S7a	23	22	0.3	438	916	5.5
S7b	208	236	3.2	301	480	2.4
S8	1	0	0.3	643	949	12.0
S9a	643	830	17.6	439	685	16.0
S9b	245	283	1.8	108	1274	15.7
S10a	643	1400	33.1	516	765	17.8
S10b	97	1350	5.5	761	2282	62.3
S10c	460	577	2.7	1461	2282	74.8
S11	172	189	1.4	204	665	1.7
S12a	172	466	4.7	589	1778	39.5
S12b	111	557	5.8	290	432	5.5
S13a	90	520	2.7	382	830	18.1
S13b	65	195	1.4	460	2282	42.9
S14	46	654	1.8	602	870	17.5
S15	0	0	0	60	95	0.6
S16	493	1486	5.0	600	1703	51.2
S17	237	240	3.5	1461	2352	60.2

Table B.10: Leak localization results with random node elevation errors drawn from a normal distribution with standard deviation of 1. For each leak scenario, 11 simulations were performed, and the average value computed. The average value and the 95 % confidence interval is visualized. Only the leak scenarios containing only one leak were tested, as the simulations were quite time consuming.

	Dual Model			Correlation Model		
Leak scenario	TD [m]	MS [m]	FP [%]	TD [m]	MS [m]	FP [%]
S1	248	599	3.5	35	51	0.5
S4	586	1491	22.2	526	970	22.8
S6	76	91	1.3	25	28	0.6
S8	948	1406	12.0	179	399	5.8
S11	355	823	5.4	307	463	6.6
S14	533	1536	17.1	455	884	11.4
S15	651	2223	44.9	1127	1974	47.4
S16	1148	2215	46.5	1227	2034	60.8
S17	1144	2091	42.1	1067	1913	47.9

Table B.11: Leak localization results with demand distributed equally to all of the system's nodes, model calibration (C3) and 11 sensors used.

	Dual Model			Correlation Model		
Leak scenario	TD [m]	MS [m]	FP [%]	TD [m]	MS [m]	FP [%]
S1	19	18	0.5	19	17	0.3
S2a	1	0	0.2	60	57	0.5
S2b	195	411	5.5	325	896	9.9
S3a	60	57	0.5	60	574	1.1
S3b	1050	2145	17.2	1050	1912	29.5
S3c	0	0	0	73	87	1.7
S4	308	419	5.8	241	928	19.6
S5a	964	2145	37.7	729	2352	31.0
S5b	35	54	0.9	15	34	0.8
S6	15	0	0.2	15	34	0.8
S7a	19	18	0.5	438	830	1.5
S7b	15	14	0.6	35	87	1.5
S8	55	98	1.4	112	207	2.7
S9a	99	13	0.5	90	187	2.4
S9b	68	268	0.6	88	312	0.8
S10a	39	47	0.6	553	765	17.8
S10b	64	123	1.7	611	1703	52.0
S10c	634	691	1.4	520	2282	70.7
S11	21	189	1.4	21	19	0.3
S12a	43	72	1.4	68	268	2.0
S12b	300	931	4.1	147	514	3.8
S13a	73	144	2.0	229	765	16.6
S13b	14	19	0.6	460	2282	37.1
S14	290	299	1.2	167	681	12.8
S15	0	0	0.0	19	17	0.3
S16	68	369	2.6	734	1536	28.4
S17	855	1806	9.1	1441	2352	80.4

Table B.12: Dual Model leak localization results for five different calibrations of the Graz-Ragnitz network. i) CAL1 is completely uncalibrated, ii) CAL2 includes calibration of the partially-closed valve, iii) CAL3 includes valve calibration in addition to calibration of the roughness parameter of the pipes using Scipy's Differential Evolution Algorithm by Virtanen et al. (2020), iv) CAL4 is the best calibration found in Steffelbauer (2018) and v) CAL5 is CAL4 plus correction of the nodal elevations. Only leak scenarios containing one leak location and scenarios with leak outflows larger than 2 $\frac{L}{s}$ were considered. Note that the Dual Model shows similar performance for all of these calibrations.

Dual Model Results: Topological Distance					
Scenario	CAL1 [m]	CAL 2 [m]	CAL3 [m]	CAL4 [m]	CAL5 [m]
S1	2	19	0	19	2
S4	118	128	286	286	308
S6	0	0	52	0	0
S8	249	73	73	112	39
S11	38	59	103	21	21
S14	124	124	216	124	290
Dual Model Results: Maximum Span					
Scenario	CAL1 [m]	CAL2 [m]	CAL3 [m]	CAL4 [m]	CAL5 [m]
S1	1	18	0	17	1
S4	286	342	307	353	427
S6	0	0	44	0	0
S8	603	98	98	176	90
S11	53	79	181	189	193
S14	486	1004	371	422	299
Dual Model Results: FP Fraction					
Scenario	CAL1 [%]	CAL2 [%]	CAL3 [%]	CAL4 [%]	CAL5 [%]
S1	0.3	0.5	0.0	0.3	0.3
S4	2.9	4.0	2.0	2.7	5.8
S6	0.0	0.0	0.5	0.0	0.0
S8	5.2	1.4	1.4	2.3	1.2
S11	1.4	1.8	1.5	1.4	1.5
S14	5.8	7.6	4.1	5.4	1.1

Table B.13: Leak localization results for five different calibrations of the Graz-Ragnitz network for the Correlation Model. See explanation in Table B.12. Note that the Correlation Model is reliant on a well-calibrated network.

Correlation Model Results: Topological Distance					
Scenario	CAL1 [m]	CAL2 [m]	CAL3 [m]	CAL4 [m]	CAL5 [m]
S1	648	19	19	19	19
S4	551	419	336	336	241
S6	15	15	15	6	15
S8	36	276	216	55	99
S11	68	21	2	21	21
S14	439	439	431	439	167
Correlation Model Results: Maximum Span					
Scenario	CAL1 [m]	CAL 2 [m]	CAL3 [m]	CAL4 [m]	CAL5 [m]
S1	879	17	17	18	18
S4	1771	1673	1673	958	928
S6	34	14	7	4	14
S8	73	276	215	98	194
S11	66	66	0	66	19
S14	816	816	738	816	668
Correlation Model Results: FP Fraction					
Scenario	CAL1 [%]	CAL 2 [%]	CAL3 [%]	CAL4 [%]	CAL5 [%]
S1	10.6	0.3	0.3	0.5	0.5
S4	59.0	49.1	40.1	17.8	17.3
S6	0.8	0.6	0.3	0.5	0.6
S8	0.8	3.3	3.3	1.4	2.4
S11	0.5	0.5	0.2	0.5	0.3
S14	14.9	14.0	13.1	13.8	12.5

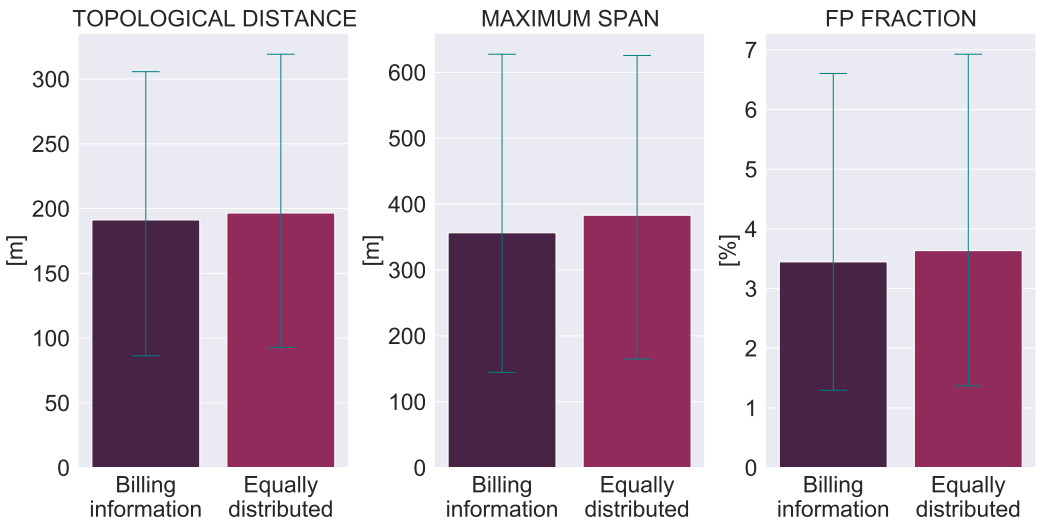


Figure B.11: Average Dual Model performance with the demand equally distributed to all of the system's nodes, and the demand retrieved from the billing information.

Appendix C. Calibration

Pipe Group	Roughness coefficient [mm]
G1	0.0148
G2	0.0061
G3	0.2594
G4	0.9101
G5	1.9980

Table C.14: Roughness coefficients found in Lippacher (2018) for each pipe group.

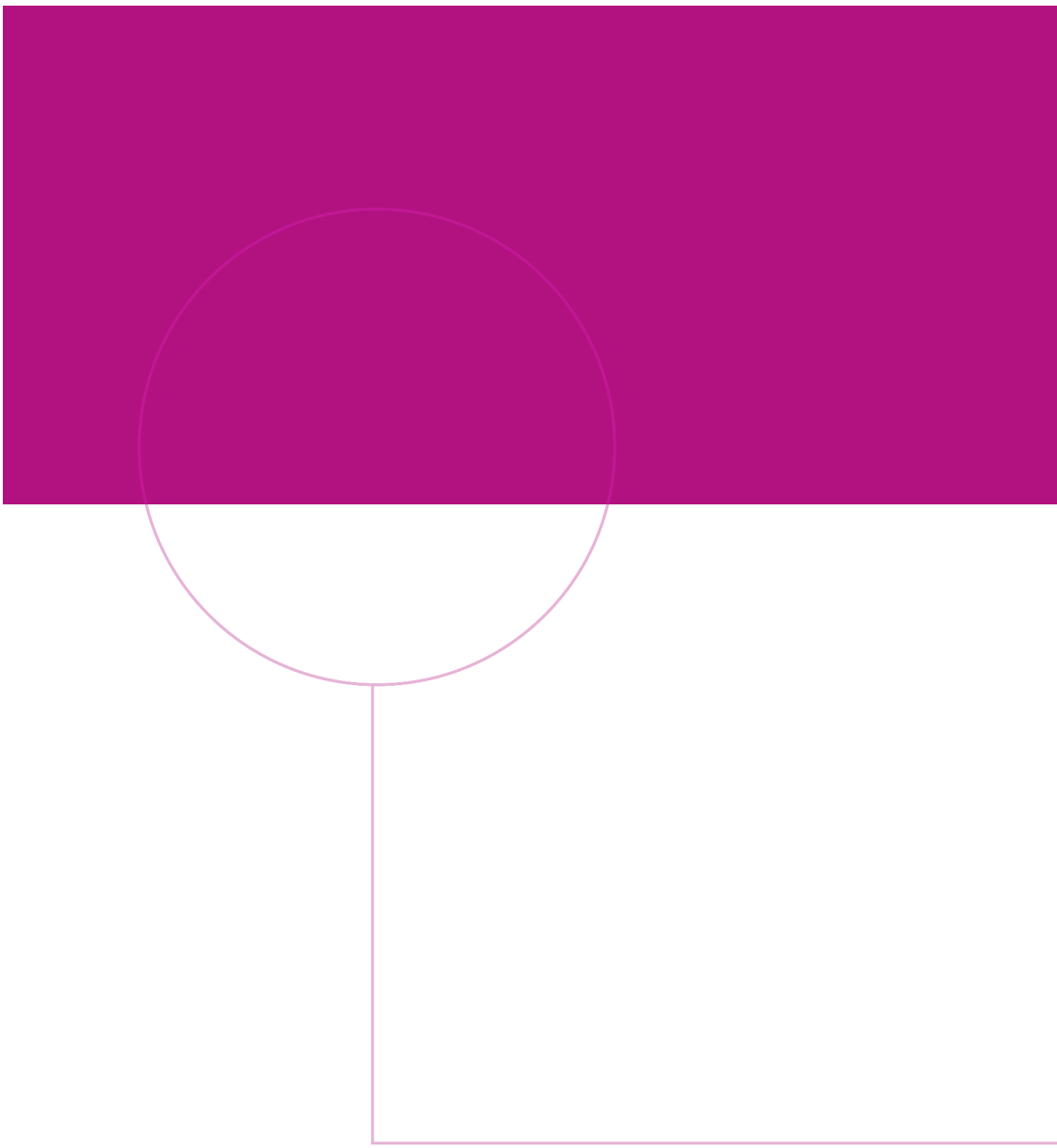
Appendix D. Project thesis

The project thesis was performed on a simple virtual water distribution network. The purpose of the project thesis was to develop an understanding of the two models. In addition, the project thesis was a tool to improve the author's programming skills. The thesis formed the idea of testing the Dual Model on weaker calibrated models in a real-world scenario, as the results were promising on the virtual network. However, the thesis is not directly relevant to the master thesis and is only attached to avoid plagiarism because

of a few sentences being quite similar in both theses. The project thesis can be found attached together with the Python scripts in a zip-file in Inspera.

Appendix E. Python scripts

Attached as a zipped folder in Inspera.



NTNU

Kunnskap for en bedre verden