A new approach in optimal sensor placement for smart hydraulic monitoring in intermittent water supply (IWS) systems

MSc. Thesis

J.J.G.M. (Joost) Verbart

19-02-2021

TU Delft

world waternet

NAIROBI CITY WATER & SEWERAGE COMPANY (NCWSC)

A new approach in optimal sensor placement for smart hydraulic monitoring in intermittent water supply (IWS) systems

A technical and financial analysis of the use of flow and pressure meters to detect hidden leaks in large cities in sub-Saharan Africa.

by

J.J.G.M. Verbart

in partial fulfilment of the requirements for the degree of

Master of Science

in Civil Engineering, MSc Track: Water Management

at Delft University of Technology
to be defended publicly on the 19^{th} of February, 2021 at 15:00h.

Student number:	4349601	
Project duration:	May 2020 - February 2021	
Thesis committee:	Prof. dr. ir. J.P. van der Hoek,	TU Delft, chair
	Dr. ir. E. Abraham,	TU Delft
	MSc. E. van Andel,	TU Delft
	Ing. R. de Groot,	Waternet

Acknowledgements

For this thesis I would like to express my gratitude to some people without whom this all would never have been possible.

I would like to start with my graduation committee, who guided me through the process. I would like to thank my main supervisor Jan Peter van der Hoek, who has supported me during the initial meetings with WorldWaternet for finding my topic and was always available to give feedback. I would like to thank Edo Abraham, who has taught me a lot with his in-depth knowledge about water supply systems and always knew to give sharp advice on things which I had not incorporated yet. I would like to thank Ellen van Andel, who has helped with viewing my thesis from an entrepreneurial perspective, thereby adding a lot of practical value for myself and the local utilities involved. Finally, I would like to thank Ralf de Groot, who has always been available for questions and help and provided me with a warm introduction within Waternet.

Furthermore, I would like to thank WorldWaternet, especially Alice Fermont, for having trust in me and providing the circumstances in which I could perform my research. I would like to thank my colleagues from Nairobi City Water and Sewerage Company (NCWSC), Sam Robroek, Philip Oruoch, Ephantus Mugo and Geoffrey Kamau, for helping me get a better understanding of the water supply system in Nairobi. I would like to express my gratitude towards Harare Water Department, especially Janneke Moors and Victor Shana, who have been very helpful in providing a helping hand when progressing was difficult and shared data and valuable insights from Harare. At last I would like to thank my colleagues at Waternet, who were always very willing to answer my questions.

On a more personal note I would like to thank my parents. Since this thesis marks the end of my education, I would like to thank you for all your love and support from the beginning until the end.

Finally, I would like to thank the people who have been my personal motivation and inspiration throughout the entire process. It has already been seven years since I lived in Malindi (Kenya) and worked at Jombas Childrens Home and St. Francis Rehabilitation Centre. I would like to thank Rose and Kamotho, for letting me stay in their house and being like an African mother and father to me. Furthermore, I would like to thank father Bapu, the priest who take in orphans and suffering children, started Jombas and has therefore been so helpful to many. At last, I would like to thank all 23 of my brothers and sisters from Jombas Childrens Home. Your humble dreams, love and fighting spirit has taught me more than one could ever learn by books or papers. Above all it has provided me with one of the most valuable thing: a purpose in life. Let this research be for you, that we (as in the picture on the title page) can improve the conditions where you once came from and walk together to a brighter future.

Asante sana

Abstract

This thesis proposes a novel design approach for a monitoring system that can detect hidden leaks in intermittent water supply (IWS) systems. Cities with IWS conditions in their drinking water network, such as Nairobi and Harare, often have a high percentage of non-revenue water (NRW) in their system. Estimations of the amount of NRW in these cities range from 40% to 50%, of which a large part is due to a leaky infrastructure. Intermittency of water supply is usually caused by a shortage of available supply, making it extra poignant to notice that these areas lose significant volumes of water. The leaks are also important locations for contaminant intrusion, which deteriorate the quality of drinking water. Additionally, intermittency of supply results in people using storage to fulfill themselves with their weekly water demand, which provides new challenges when constructing hydraulic models. Hidden leaks, which are leaks that do not appear at the surface, can be noticed in continuously supplied areas through reports of pressure deficiencies or the absence of supply. As these are regular circumstances in IWS areas, these hidden leaks are seldom noticed. Therefore, methods that are applicable in IWS systems need to be developed to detect these hidden leaks.

This thesis proposes a new approach to detect hidden leaks in IWS areas with a smart hydraulic monitoring system. The approach optimizes the design of such a system in a district metered area (DMA) with IWS conditions in sub-Saharan Africa, by balancing information density and investment costs. By using as little equipment as possible, this optimization study aims to be not only scientifically and practically relevant, but also cost-effective.

The methodology that was used to design the monitoring system makes use of a similar concept as the Dynamical Bandwidth Monitor (DBM), which is a smart hydraulic monitoring system that has been applied regularly in networks with continuous supply. The monitoring system consists of sensors that continuously measure flow or pressure and it compares these measurements to a range of expected values, attributing deviations from these expected values to a potential leak. A case study of Ashdown Park, a DMA with IWS conditions in Harare, was used to assess the performance of the design. The flow into this DMA and the pressure at its inlet had been monitored for one year. Two designs of the monitoring system were made, one which mainly consisted of flow sensors and one with mostly pressure sensors, to showcase which type of sensor could best be used in Ashdown Park. A hydraulic model was constructed for the DMA using pressure dependent outflow modelling. Daily demand patterns were constructed from analyzing the inflow measurements and used to calibrate the hydraulic model. The proposed calibration method assumes linear relationships between the demands and inlet pressure on one side and the pressure at a specific node and flow at a specific pipe on the other side. The range of expected flows and pressures within the DMA was calculated by Monte Carlo analyses, during which demand realizations were modelled by using a novel method which made use of a random weighted choice of demand, based on the outflow from a single tap. The ability of the monitoring system to detect leaks during different demand realizations was stored

in a three-dimensional Boolean matrix, which was then used to determine the optimal sensor placement. A social and financial analysis, summarized in a business model canvas, shows more practical challenges and opportunities that could arise from implementing the monitoring system. The lessons learnt from this thesis were used to showcase whether the monitoring system could be applicable for IWS systems around the globe.

Several conclusions can be drawn from the results of this thesis. The daily demand patterns in Ashdown Park showed a different pattern than in continuously supplied systems, showing less strong peaks. This could be due to a constant water demand for filling storage, leaks in the system or different consumer behaviour. The calibration method made it possible to model flows and pressures at the DMA inlet which were comparable to the measurements. The novel method to model demand realizations with a random weighted choice and a single tap capacity, showed promising results since the spread of the modelled inflow was well comparable to the spread of the inflow measurements. This standard tap capacity is especially suitable for IWS areas, since most people in IWS areas usually only have one tap directly connected to the water supply system and water end-use devices are not directly connected to the network. Furthermore, it was found that the water use behaviour of inhabitants of Ashdown Park had been more constant than the supply behaviour of the water utility. This irregular supply behaviour of the utility increased the difficulty of designing a pressure monitoring system. Using a flow monitoring system to detect leaks showed a better performance (leaks could be found on a daily basis in 25% of the pipes in the DMA) than using a monitoring system with pressure sensors (leaks could be found in 1% of the pipes). Making the monitoring system with pressure sensors dependent on the inlet pressure increased its performance (from 1% to 8.3%). Branched parts of the system were more favorable locations to place sensors and sensors at the DMA inlet were crucial for calibrating the hydraulic model. Practical barriers that were identified during this thesis were irregular operational schemes, unknown demand patterns and incomplete GIS data. Furthermore, costs can be saved as soon as leaks are detected, making the financial profitability very dependent on the performance of the system and the occurrence of leaks. The applicability of the monitoring system in IWS areas around the globe is determined by the priorities of a local water utility, its network characteristics and the ability of the local utility to overcome implementation barriers.

The main limitations in this research are due to making some simplified assumptions, such as assuming a constant flow-rate from the tap in all households in Ashdown Park, and due to a lack of understanding of the local situation, since this research was performed in the Netherlands. To validate assumptions and get better understanding of the local situation, it is advised to conduct follow-up research at the location of interest. Especially a pilot project of the proposed monitoring system would likely find more practical barriers and limitations than could be thought off in this thesis and therefore bring more valuable information for the implementation of a smart hydraulic monitoring system.

If prioritized, properly installed and operated, the proposed smart hydraulic monitoring system could generate substantial water savings and provide many social benefits, such as an increased access to clean drinking water and employment opportunities. Above all, it can assist a local utility with fulfilling their responsibility: supplying people with the basic need of drinking water.

Contents

List of Acronyms

DBM Dynamische Bandbreedte Monitor (English: Dynamical Bandwidth Monitor)

DMA District Metered Area

GIS Geographical Information System

H2H House-to-house

IWA International Water Association

IWS Intermittent Water Supply

KPI Key Performance Indicator

MJ Morton Jaffray (Drinking water treatment plant in Harare)

MNF Minimum Night Flow

NCWSC Nairobi City Water and Sewerage Company

NRMSE Normalized Root Mean Squared Error

NRW Non Revenue Water

PDO Pressure Dependant Outflow

RMSE Root Mean Squared Error

TC Tap Capacity

VEI Vitens Evides International

WC Warren Control (Control centre for pumping stations in Harare)

ZWL Zimbabwe dollar

1. Introduction

1.1 Problem description

Every year the World Economic Forum publicises a list of top threats which face our world, ranked by likelihood and impact. In both categories “Water crisis” comes out at the largest societal risk to our world. Among all possible risks, it scores 8^{th} in likelihood and 5^{th} in impact (Edmond, 2020). This highlights the strong need for the sustainable use of our water resources globally and also the impact that can be expected when water shortages are faced. Many world leaders point out the need for re-evaluation of water, which makes it more poignant to see that cities with millions of inhabitants such as Nairobi (Ndegwa, 2016) or Harare, discussed with Shana (2020), loose 40% to 50% of their daily produced drinking water before it ends up at the customer. On top of that, these same cities do not have the production capacity to supply the demand of its entire population, resulting in intermittent water supply (IWS) during which water is supplied to an area at only for a few times a week. This water that does not reach the customer is labelled as Non-Revenue Water (NRW). A proportion of this NRW is lost due to a leaky infrastructure, but financial constraints limit the water utility to make large investments in new infrastructure (KAM, 2020). Apart from the water loss, various research has also shown an impaired water quality in intermittent supply systems (Yassin et al., 2006)(Andey and Kelkar, 2007)(Elala et al., 2011), which is likely caused by contaminant intrusion through leaks (Kumpel and Nelson, 2013).

Let us compare the struggles that are involved with improving the performances of drinking water system by comparing the situation at hand to the growth of the mobile network in Africa. The last years Africa has had a gigantic growth of its mobile network, thereby connecting a lot of people to internet and the entrepreneurial opportunities that come along with it. For example, Nyirenda-Jere and Biru (2015) showed that Africa’s international bandwidth increased 20-fold and its terrestrial network more than doubled over the period from 2009 to 2014. These steep growth rates were partially possible because there was not yet an extensive infrastructure in place to deal with the service provided. Also, technologies that did not have to be placed at the location, such as satellites, made it possible for the mobile network to grow at the rate it did. The drinking water supply network however, can not be newly built as it evolves as the city evolves. Nairobi, for example, has grown with a factor 34 in the last 70 years (World Bank, 2020). With limited funds it has always been a challenge to connect everyone to its drinking water system, let alone make detailed descriptions of the locations and characteristics of the installed network. Amsterdam, which has only grown by a factor 1.4, has information about its drinking water network that is up-to-date. This has allowed its utility to use this data to build models and perform difficult calculations. In Nairobi however, there is not such a basis, mainly due to the large growth challenges that were faced. The transition to a water system that connects everyone and has a low NRW is a lot harder than was the case for improving the mobile network, as the former continues on previously built infrastructure and registration. This research therefore aims on taking the current situation of the drinking water system (with IWS) as the basis and aims on improving this system, still functioning with IWS. By accepting IWS as the existing situation and the situation to be, it opposes other research which focus on the transition from IWS to continuous supply systems (Mohapatra et al., 2014)(Ilaya-Ayza et al., 2016).

Monitoring the performance of the system of today can give valuable information for predicting the performance in the future, when water demand and supply change and infrastructure deteriorates. Furthermore, a well-working monitoring system is a key instrument in efforts to reduce NRW (Jang and Choi, 2017) and could as well be used to identify intrusion hot spots (Sakomoto et al., 2020). It can be used to construct water balances in isolated and metered sections of the drinking water network, called District Metered Areas (DMA). By isolating and monitoring these different parts, a utility is able to find which city section looses most water and is therefore a weak point of the network (Jang and Choi, 2017). Investments in a monitoring system can therefore not only help to set up management strategies to provide enough water access and quality in the future, they can also lead to higher revenues and more efficient water use for the water utility.

An elaborate interview with the local NRW-manager of the local water utility of Nairobi, engineer Mugo, showed which data is already monitored in Nairobi and which data is still lacking.[1]. The main issue for Nairobi's utility comes with the assessment of real losses, that are the volumes of leaks in the system. Whereas visible leaks are reported to the utility, hidden leaks are a nightmare. These hidden leaks are leaks that do not appear at the surface, but still leak away through perforated pipes. These leaks increase the amount of NRW and are potential locations for contaminant intrusion. They are difficult to locate, since they are not visible at the surface. In continuous systems, they can be notified due to customer complaints about a decrease in pressure or absence of supply. Since these are regular circumstances in an intermittent supply, these customer notifications are less likely to occur and different methods to become aware of their presence should be explored.

This research uses a new approach in order to optimize the monitoring system of a DMA with IWS conditions in sub-Saharan Africa, by balancing information density and investment costs. It has a twofold objective. Firstly, it aims at designing a monitoring system with flow and pressure meters in a DMA with IWS that can create awareness of hidden leaks in a the drinking water supply network and give direction to their location, using as little measurement devices as possible. Secondly, it aims at exploring new sustainable business cases for the implementation of this monitoring system. By using as little equipment as possible, this optimization study aims to be not only scientifically practically relevant, but also cost-effective. Finally, by providing suggestions about the possible business cases that come around with the new system, this research aims to provide a full picture which also shows required financial and social relevance of the system.

[1]Notes of the entire interview can be read at appendix A.1

1.2 Organizational framework

This thesis is conducted as part of the WaterWorX program, a program of 10 Dutch water utilities together with the Dutch ministry of foreign affairs that has set the goal to increase access to sustainable water services for 10 million people between 2017 – 2030. The Dutch water utility overseeing this thesis is Waternet. Their international branch, WorldWaternet, has the goal to support foreign public water utilities by offering sustainable, integral solutions for water challenges. They have constructed a Water Operational Partnership (WOP) with the Nairobi City Water and Sewerage Company (NCWSC). Later in this thesis, data from the Harare Water Department was used as a case study. This data was retrieved through Vitens Evides International (VEI), which is also an international branch of Dutch water utilities (Vitens and Evides) operating as part of the WaterWorX program (see Figure 1.1).

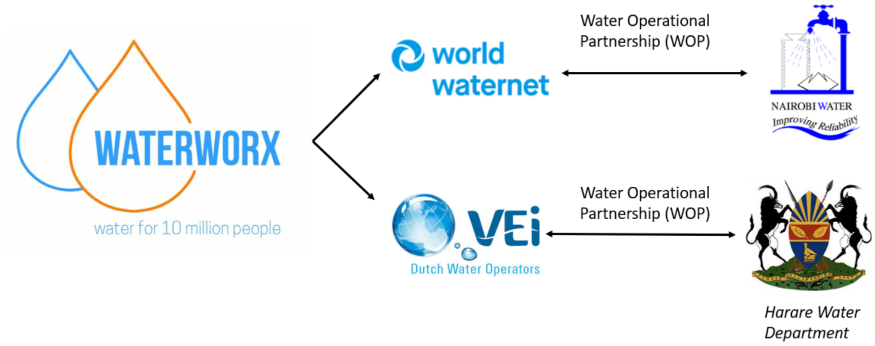

Figure 1.1: Organizational framework of this thesis.

2. Literature review

This literature review chapter describes the literature that was consulted during this thesis. It provides background information about the international NRW-framework and methods to measure NRW. Furthermore, it shows the challenges that arise when modelling IWS systems. Lastly, it provides more information that was used for designing the leak localization system.

2.1 NRW background information

The literature about NRW is mainly used to explore the international guidelines and terminology, to find common ground for discussing the challenges in NRW for NCWSC.

The IWA water balance

The most important strategies that are currently adopted by WaterWorX and other international standards to determine the amount of NRW in a network are the bottom-up approach and the top-down approach (Ziegler et al., 2011). Both strategies divide the water balance over a network in a part of the water that is not billed to the customers (NRW) and a part that is billed to the customer. The water balance that is mostly used is the one adopted by the International Water Association (IWA), which can be seen in Table 2.1 (Lambert and Hirner, 2000)(Ziegler et al., 2011). It has five different columns, splitting the water entering the system in different segments. The original version does not include any headings in the columns, but in this thesis different levels are added in the headings to facilitate referencing and clarify the use of this water balance. As a next step, the differences between the top-down approach and bottom-up approach will be regarded.

Level 0	Level 1	Level 2	Level 3	Type of water
System input volume Q_{in}	Authorised consumption Q_A	Billed authorised consumption Q_{BA}	Billed water exported	Revenue water
			Billed metered consumption	
			Billed unmetered consumption	
		Unbilled authorised consumption Q_{UA}	Unbilled metered consumption	Non-revenue water
			Unbilled unmetered consumption	
	Water losses Q_L	Apparent losses Q_{AL}	Unauthorised consumption	
			Customer meter inaccuracies and data handling errors	
		Real losses Q_{RL}	Leakage on transmission and distribution mains	
			Leakage and overflow at storage tanks	
			Leakage on service connections up to point of customer meter	

Table 2.1: IWA water balance with column headings

Top-down vs. Bottom-up

As a first step in the top-down approach, the system input volume (level 0) is determined. Afterwards it determines volumes of water on level 2 from the top of the balance downwards. So it first determines the billed authorised consumption (Q_{BA}), then the unbilled authorised consumption (Q_{UA}) and then the apparent losses (Q_{RL}). These volumes are estimated by first estimating their subvolumes (level 3) and summing them to get the cohering level 2 volumes. The remaining real losses (Q_{RL}) are determined by subtracting the preceding volumes from the system input volume, so not by actual measurements. Advantages of using this method is that it is quick and therefore cheap, since most volumes can be estimated from the consumers data that are present in the database of the utility. Nevertheless, the estimations in this method are prone to inaccuracies. These inaccuracies aggregate when water balances of larger time periods are constructed. [1]

The bottom-up approach starts with measuring the real losses in the system. Since the losses are measured and calculated, they can be determined more accurately than with the top-down approach. The other losses in level 2 (apparent losses, unbilled authorised consumption and billed authorised consumption) are estimated the same way as in the top-down approach. However, estimation errors are more easily noticed since the sum of the aforementioned volumes should equal the system input volume. An advantage of this approach is that the results are more reliable and substantiated, whereas disadvantages are that measuring real losses is a challenging practice and that the approach is time consuming and expensive.

In order to better understand both approaches, a closer look is taken to the definition of the volumes of the IWA water balance and how they can be measured.

System input volume (level 0)

The success of the water balance is highly dependent on the right determination of the incoming volume of water into an area. By measuring the incoming volume and outgoing volumes in the area, you consequently create a District Metered Area (DMA). A DMA is originally defined by Morrison et al. (2007) as: “A discrete area of a distribution system usually created by the closure of valves or complete disconnection of pipe work in which the quantities of water entering and leaving the area are metered.” Since you measure incoming volumes, one needs flow meters in order to determine a system input volume.

Billed and unbilled authorised consumption (level 2)

The billed authorised consumption consists of: 1) billed water exported, 2) billed metered consumption and 3) billed unmetered consumption. Billed water exported takes account for the volume of water that is produced by NCWSC but sold to another utility. Billed metered consumption represents the volume of customers that are billed according to the volume of water that they consumed. Billed unmetered consumption is the volume of water that is billed, but not based on an actual measurement. Usually this is the

[1] This happens for instance when estimations of monthly billed volumes are aggregated for a yearly assessment of the percentage of NRW in the supplied area.

case when the customer meter is broken, or the meter reader did not acquire the right data.

The unbilled authorised consumption is the volume of water that is used, but not billed for a reason. This can be for instance the volume of water that is used for firefighting, which is not billed in most countries or for flushing the drinking water system. If this volume is measured, it is labelled “unbilled metered consumption”. If this volume is not measured, it is labelled “unbilled unmetered consumption”.

The methods to determine the volume of billed and unbilled authorized consumption are per definition dependent on rules and regulations of the authority. Metered consumption can be measured by using an automated software or by meter reading personnel. Types of water that are not billed, but authorized, can be found within the regulations of the utility.

Apparent losses (level 2)

Apparent losses account for the drinking water that has been produced, transported and has reached the customer, but it has not been registered. Sometimes apparent losses are referred to as “commercial losses”, both terms have similar meanings (VEI, 2020). The apparent losses can have three causes: 1) Meter inaccuracies, 2) Data handling errors and 3) Unauthorised consumption.

Meter inaccuracies can happen at every point where a meter is placed. When determining meter inaccuracies, usually only meter inaccuracies in customer meters are taken into account since these represent the largest volumes (Ziegler et al., 2011). The data handling errors are human errors that are often linked to the database software that is used to register the data. Unauthorised consumption comes from people that consume water illegally and can also be described as “water theft”. Forms of unauthorised consumption are: Meter by-pass, illegal connections, illegal re-connection, fetching water before the meter, meter reversal and meter tampering (UN-Habitat, 2012).

The volume lost due to meter inaccuracies can be estimated by making a selection of in field meters from a similar brand, size and age group and testing them on a bench. The inaccuracies that can be concluded from these test can be used to re-estimate the measurements from the whole group (scale them up or down, depending on the inaccuracy) (Ziegler et al., 2011). Another way to estimate the inaccuracy of meters is to use a bucket and calculate the flow into the bucket. It is advised to calibrate DMA meters and large consumer meters every 6-12 months. The frequency of calibration of consumer meters should be guided by the performance history, according to VEI (2020). Data handling errors are human errors that can be reduced by training personnel and creating a better understanding of the database (Ziegler et al., 2011). The amount of illegal water use can be best determined by conducting surveys in a pilot area or giving incentives to the local inhabitants to notify illegal connections at the local utility (UN-Habitat, 2012).

Real losses (level 2)

Real losses can be defined based on their location and their visibility (Ziegler et al., 2011). The different locations where leaks can occur are at: 1) Transmission and distribution mains, 2) service connections and 3) storage tanks. Leaks at transmission and distribution mains can occur at pipes, joints and valves and usually have medium to high flow rates and short to medium run-times. Leaks at service connections can occur at joints and fittings, which often low flow rates and long run-times. In terms of visibility, leaks can be classified as “visible”, “hidden” and “background leakage”. Visible leaks will quickly appear at the surface and can therefore be identified and repaired quickly. Hidden leaks will not appear at the surface, but can be detected by leak detection methods due to its significant size. Farley (2001) assumes that background leaks are almost never detected and they have a large share in the amount of real water losses and a long run-time. They can also be referred to as “Unavoidable Real Losses (UARL)”.

Let us first consider the way leaks in the transmission and distribution mains are measured. Visible leaks are usually reported by local inhabitants. A repair team measures the leak and repairs it. The volume lost is given by the flow rate through the leak and the response time of the repair, which is often assumed equal to the run-time of the leak. Hidden leaks have to actively be identified. The process of leak identification, location and repair is known as Active Leakage Control (ALC) (Morrison et al., 2007). First, the utility needs to become aware of the hidden leak. This can be done by applying continuous monitoring and flow analysis. Structural changes of the flow regime within pipes can be due to the formation of hidden leaks. Often continuous measurements are not available, or difficult to analyse since flow patterns differ too much over the day. In those cases a Minimum Night Flow (MNF) analysis can be used to determine the amount of hidden leaks. In an MNF-analysis a minimum night consumption is determined, as well as the background leakage. The difference between the sum of the two and the flow that is actually flowing in the tube is regarded as the volume of the hidden leaks. After being aware of a leak, its location should be narrowed down to a range within 300m. This can be done by using step-testing in DMA’s. During step-testing, parts of the DMA are closed to form different sections and water is allowed to flow within these section. At every section it is measured if water is lost. Also leak noise loggers and sounding surveys can be used to detect the location within a 300m-reach (Kober and Gangl, 2009). Finally, the location range should be narrowed down to the exact location of the leak (within a range of +/- 0.3m). Applied methods to achieve this are a combination of listening sticks and ground microphones, leak noise correlation and non-acoustic methods such as tracing gas or ground penetrating radar (Ziegler et al., 2011). Due to the small magnitude, background leakage can not be detected by active leak detection. There is an empirical formula constructed by the Lambert et al. (1999) on behalf of the IWA for the background leakage that takes into account the network length, the number of service connections, the the length of private pipes after the property line up to the customer meter and the average operating pressure[2].

Leaks up to the point of the customer meter can be approached differently than leakage in mains, since in the location of the leakage is already known. The volume of these leakages can be estimated by doing house-to-house visits and surveys. Eventually, a very practical strategy to decrease this volume is to replace all leaking household connections by high

[2]This formula however is based on data from developed countries with a continuous water supply, so one should be careful in applying this formula in IWS systems. Another method has been constructed by DeSilva et al. (2005) that takes into account the pipe material, number of joints and a pressure factor. This method has been applied in South Australia and resulted in a background leakage that was only 67% of the IWA-value. This formula has less international recognition.

quality connections provided by the utility as suggested by Preston and Sturm (2002). Leakage and overflow at storage tanks are usually visible, since most tanks are placed above ground. The volume of small leaks can be determined by using a bucket to measure the flow. Hidden leaks at the storage tank can be identified by continuous monitoring, when flow meters are installed at the inflow and outflow point of the tank.

2.2 Modelling IWS systems

The behaviour of water flows and pressures in intermittent supply systems is often very difficult to predict and therefore challenging to model. Main reasons for this difficulty are usually differences in demands, high leakage levels and the occurrence of low pressures in the system compared to continuous supply systems (De Marchis et al., 2010).

Demands in IWS systems

The differences in demand pattern, compared to pressurized systems, is mainly due to the need of people to store water during days with supply to ensure themselves with water during the days without supply. This results in a different pattern and an increased water demand during the days with supply (De Marchis et al., 2010). Next to demand patterns, per capita water demand can differ as well. A relatively high per capita demand is mentioned in a modelling report of the Republic of Zimbabwe (2017). Here, they refer to a commercial report of 2016 which indicated an average consumption of 497 L/p/day for high density areas and 412 L/p/day for low density areas. The per capita water use was calculated by dividing the inflow into an area by its number of inhabitants. Underlying reasons for this high base demand could therefore be meter inaccuracies, high leakage levels, unregistered bulk consumers and a higher personal water use than expected. [3] Lastly, it is a challenge to model the spatial and temporal variation of demand within the network. A stochastic model that can be used to model this spatial variation in continuous supply systems is SIMDEUM. This model uses different parameters for household's occupancy (e.g. household size, ages), household appliances (typical flows and volumes) and consumers' water using behaviour (e.g. number of toilet flushes, duration of shower) (Blokker et al., 2017). However, no model was found that had similar water use statistics for IWS areas, stressing the need for more survey data on consumption behaviour.

Leakages in IWS systems

High leakage levels are often due to the fast deterioration of an IWS network (Al-Washali et al., 2019). These leakages can be modelled by adding a certain demand to a model (Casillas Ponce et al., 2013) or by relating the leak volume to the pressure in the system (Crowl and Louvar, 2011).

[3] Whilst having measured this high use, the same report considers these numbers unreliable, stating that a large proportion of the consumer meters would not be working or read and consumption would largely be based on estimates. Therefore, the report uses unit demands of South Africa for its hydraulic model later on.

Low pressures and the nominal pressure

At last, intermittent supply systems regularly experience low pressures in the system. If the system's pressure drops below a certain minimum, only part the system can be supplied with water. This results for some systems in a water demand that is highly dependent on the system's pressure. In order to model such as situation, pressure dependant outflow (PDO) modelling should be applied instead of demand based modelling, according to Vairavamoorthy et al. (2001). In these models, the demand decreases when the pressure drops below a certain limit. This limit is called the "nominal pressure", represented by H_s in Figure 2.1 (Wagner et al., 1988). The pressure at which no demand occurs is called the "minimum pressure" (represented by H_m in Figure 2.1). This Figure assumes a non-linear relation and is used in WNTR, which is a Python-package that can be used to run pressure dependant outflow models (NTESS, 2019).

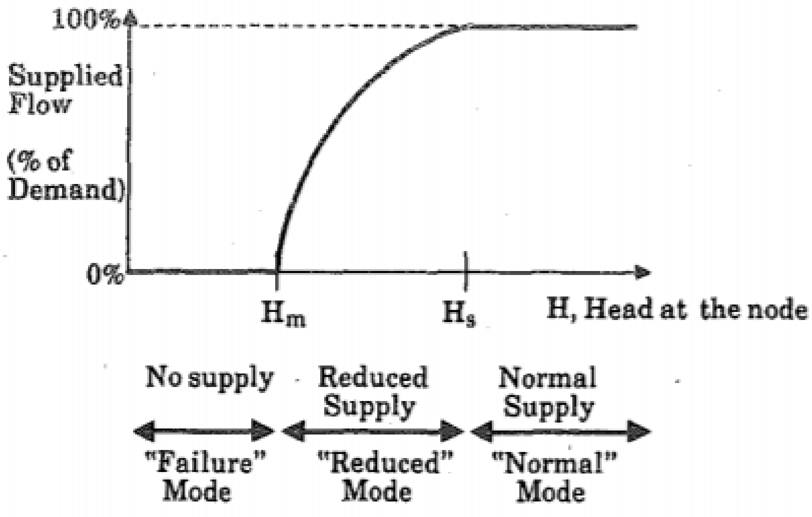

Figure 2.1: Demand decreases when insufficient pressure is available

2.3 Design of a leak localization and detection system

In order to optimally allocate sensors for leak detection purposes, first a leak detection method was chosen. More research was done into allocating demand and implementing leaks in the hydraulic model, as these fields provided some challenges. Finally, some extra literature was studied for solving the optimization problem.

Leak detection method

In order detect hidden leaks, the same leak detection principle is used as for the "Dynamische Bandbreedte Monitor" (DBM)[4], a leak detection method design by the Dutch utility Vitens. It uses historical measurement data to construct a bandwidth of expected hydraulic values in the water supply system, so either expected pressures or expected flows. This bandwidth is bounded by so-called "alarm values". An alarm is raised when measured values exceed this bandwidth, as this exceedance might be caused by a leak

[4] In English: Dynamic Bandwidth Monitor

(Van Vossen-van den Berg, 2017). To distinguish between false alarms and an actual leak, it is advised to take a certain detection time frame into account during which the measurement values need to exceed the alarm values before an alarm is raised (Van Steen, 2020). Furthermore, it should be noted that most research in sensor allocation for leak detection in water distribution networks focuses around the placement of pressure sensors (Perez et al., 2009)(Khorshidi et al., 2020), whilst in cities such as Nairobi and Harare flow sensors are used with higher frequency than pressure sensors, as told by Shana (2020) and Mugo (2020).

Modelling demand allocation

The key to a successful DBM is the construction of the boundaries for the expected range of values, the so called “alarm values”. The boundaries represent extreme scenarios which can still be expected and occur usually when demand concentrates locally. Therefore the method of spatially allocating demand is important for constructing of these alarm values. Spatially allocating demand has been a challenge for many years in continuous supply areas (Kanakoudis and Gonelas, 2014) as well as in IWS areas (De Marchis et al., 2010). For this spatial allocation, researchers and operators constantly have to trade off the ability to accurately model the human water consumption behaviour against the simplicity of the model to allow for fast hydraulic calculations. One of the simplifications that induce errors is the representation of demand in only two terminal nodes connected to a pipeline, whilst actual withdrawals might take place at several locations along the pipeline (Kanakoudis and Gonelas, 2014). Furthermore, the random component within human behaviour makes it impossible to model water demand allocation completely accurate. Other struggles arise in assessing the quantity of demand. This quantity can be estimated by using bottom-up approaches, usually based on clustered water meter data, and top-down approaches, based on a water balance constructed from the water volume that enters the network. According to Kanakoudis and Gonelas (2014), most proposed methods are a combination of these approaches .

Later in this thesis, a DMA in Harare (Zimbabwe) called “Ashdown Park” is used as a case study for which a hydraulic model is constructed to model IWS conditions. VEI had shared an EPANET model of Ashdown Park, which formed the basis of the hydraulic model which will later be constructed. The spatial demands in this EPANET model have been allocated by employees of VEI using service polygons within a GIS network. Each polygon contained a certain number of households, each with the same average household consumption. GIS can serve as a logical and helpful demand allocation tool, due to its spatial analysis capabilities (Haestad Methods et al., 2003). EPANET, which is often used in hydraulic modelling, can be used to show the average flows in the network at specific times, but it does not include a multitude of possible flow scenarios at the same time within the network’s pipes (Rossman, 2000). For this additional software such as SIMDEUM can be used, which uses random draws from probability density functions that represent human behaviour and parameters with physical meaning, to predict which scenarios might occur in the area (Blokker et al., 2017). As previously mentioned, this type of software is unfortunately not yet available for IWS systems.

Modelling leaks

Determining the leak size, which should be used in the leak detection model, is a problem that many researches have faced. The water utilities of Harare and Nairobi do register their leaks, but unfortunately not their sizes. Therefore, a similar leak size as used in other literature should be taken into account. A range of leak sizes that were used in comparable literature is listed below:

- **7.5 - 15 m3/h**. These leak sizes were artificial and used in a modelling study to test a leak localization method in Leimuiden (the Netherlands) (Moors, 2016).
- **2.5 - 7.5 m3/h**. These artificial leak sizes were used in a modelling study to determine a suitable leak detection time (Van Steen, 2020).
- **20.6 m3/h**. This leak was created by opening a fire hydrant in Barcelona and used to test a new data driven leak localization method (Soldevila et al., 2019).
- **8.1 m3/h ; 0.72 - 13.0 m3/h**. A real DMA in the UK with approximately 1000 properties showed a leak flow rate of 8.1 m3/h, although it could not be made certain whether this was one leak or multiple leaks. In a model study with the same DMA, the algorithm was tested with single leaks ranging from 0.72 to 13.0 m3/h (Sophocleous et al., 2019).
- **2.52 - 10.8 m3/h ; 2.52 - 22.7 m3/h**. This paper describes a new approach for model-based leak detection and localization in the drinking water network in Barcelona. In this paper, the water utility in Barcelona states that the leaks to be detected range from 2.52 m3/h to 10.8 m3/h. The modelling study aims to isolate leaks that range from 2.52 m3/h to 22.7 m3/h (Casillas Ponce et al., 2013).

The leak size range of 2.52 - 10.8 m3/h was defined by the experiences of a local water utility in the research of Casillas Ponce et al. (2013), which gives this range more practical weight than most other leak sizes that were found in literature. Furthermore, the leak size that is implemented in the hydraulic model should be in reasonable proportion to the flow into the DMA (Moors, 2016).

Notably, most literature about leak localization and detection methods defines a leak size in their model in terms of a given flow rate (Moors, 2016)(Van Steen, 2020)(Soldevila et al., 2019)(Sophocleous et al., 2019)(Casillas Ponce et al., 2013). The main reason for this is the usage of demand-driven models, such as EPANET, that require to insert the leak into the model as a demand. However, other research has shown that leak size might better be defined by its surface area, as its flow rate depends on the occurring pressure (Marchis and Milici, 2019). In pressure-dependant modelling, which is performed in the models in this thesis, the flow-rate that is lost through the leak (d_{leak}) can be written as a function of the pressure in the leaky pipe (p), a unitless discharge coefficient (Cd), the area of the hole (A), a unitless exponent related to characteristics of the leak (α) and the density of the fluid (ρ) (Crowl and Louvar, 2011). The leak volume in WNTR is therefore defined as in equation 2.1 (NTESS, 2019).

$$d_{leak} = C_d A p^{\alpha} \sqrt{\frac{2}{\rho}} \tag{2.1}$$

Intermittency in supply systems can also reduce the total leakage in the network through equation 2.1 as well, since there is no volume lost in areas without pressure. This absence of pressure occurs on days without supply.

Boolean leak detection

In similar methods to design a configuration of sensors for leak detection purposes, Boolean classifiers are used to describe whether a leak is found at a certain location or not (Perez et al., 2009)(Khorshidi et al., 2020). These methods store information whether a leak can be detected by a sensor in another node as 1's in a Boolean matrix with two dimensions (one dimension for the leaks and one dimension for the sensors). A shortcoming of these methods is that it only takes a single demand realization into account, thereby ignoring other flow scenarios which might occur during different demand realizations.

Optimization

The optimization solver Gurobi can solve certain classes of optimization problems (Gurobi Optimization LLC, 2020). It can optimize quadratic problems, which is useful to find the least squares solution between two vectors (Boyd and Vandenberghe, 2018), and solve mixed-integer problems, such as the Boolean case which solely consists of the integers 0 and 1. It therefore allows for mixed-integer quadratic-constrained programming (MIQCP).

3. Objective and research questions

3.1 Objective

As previously stated, the main objective of this research will be:

> *"Design an optimized configuration of flow or pressure meters in a monitoring system in a DMA with IWS that can locate hidden leaks with as little equipment as possible and explore the business opportunities of such a system."*

3.2 Research questions

The research is divided into four parts: A theoretical study (1), a modelling study (2), constructing a business model (3) and a study into the world-wide applicability of this study (4). The research questions divided over the different parts will be as follows:

Part 1: Theoretical background study

Q1: Which type of data is important to monitor in Nairobi?[1]

Part 2: Modelling case study Harare: Optimally design a smart hydraulic monitoring system for a DMA with IWS

Q2: How can intermittency be included in a hydraulic model?

Q3: How can flow and pressure meters be used to detect hidden leaks?

Q4: How can the placement of the meters in the monitoring system be optimized?

Part 3: Business model

Q5: What are the financial and social benefits of implementing a monitoring system?

Part 4: General applicability for IWS systems globally

Q6: To what extent can the outcomes of this research be used to construct monitoring systems for Water supply networks with IWS conditions around the globe?

[1] This question was answered by conducting an elaborate interview with the NRW-manager of NCWSC (appendix A.1). The conclusions for this interview were used to define the objective of this thesis. Therefore, the question was answered before the objective was formulated.

4. Methodology

This chapter shows the main structure that was used during this research and its documentation. Furthermore, it describes the methodologies that were applied for answering the research questions.

4.1 Thesis structure

The structure of this thesis can be seen in Figure 4.1. This thesis started with a structured interview with the local NRW-manager of NCWSC, identifying which information was lacking the most for their existing NRW strategies. This interview was used for formulating the research objective. Afterwards, a case study of an IWS area in Harare was used for a modelling study to design a monitoring system that can detect hidden leaks in an IWS system. Reason for the switch of focus location from Nairobi to Harare will be described below. The results of the case study are shown in chapter 5, 6 and 7. At last, a business plan for the monitoring was constructed (chapter 8.3) and conclusions were drawn to which extent the monitoring system could be useful for IWS systems around the globe (chapter 9).

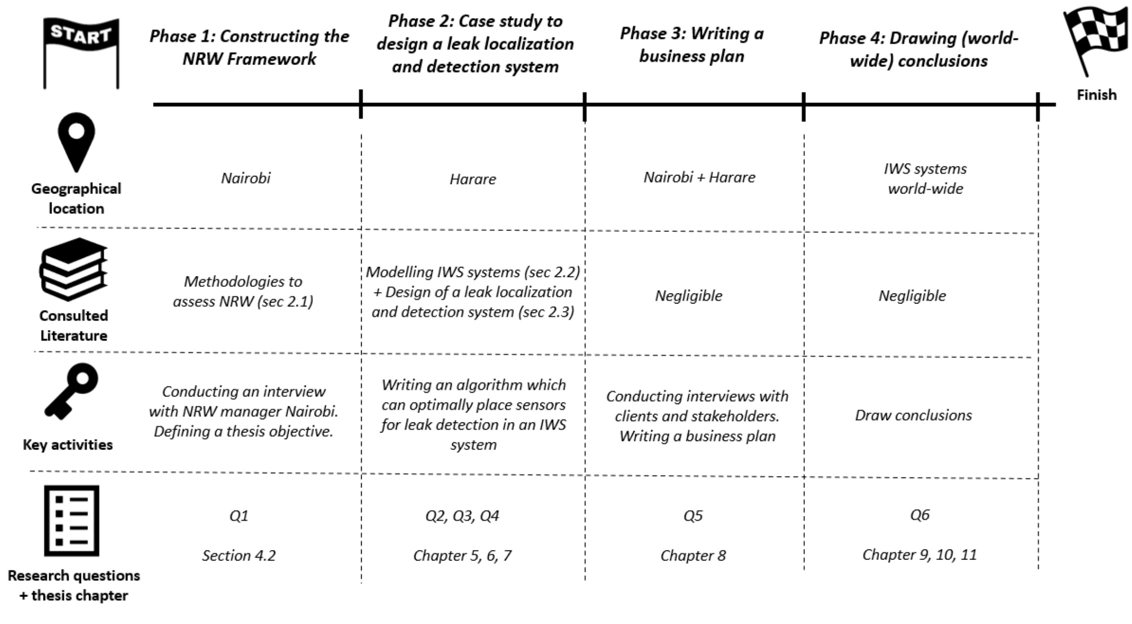

Figure 4.1: Summary of the methodology and planning of this thesis.

The main reason for the shift from Nairobi to Harare for the case study was that the received digital information from the water supply system in Nairobi contained contradictory information about pipe lengths and diameters in different files, making it hard to construct a trustworthy model. In Harare however, they had some DMA's which had complete GIS documentation, whose inlet pressure and flow at the entrance had been monitored over the last years and even had a complete House-to-House (H2H) survey in one of the DMA's. This provided the information needed to build a hydraulic model for an IWS area. Furthermore, at the beginning of this modelling study, there were some internal switches of staff within the NRW-department in Nairobi. This came along with the fact that Covid-19 had just reached Kenya, which added challenges to life in the city and prevented me from visiting Nairobi to validate information. A full account of the planning and events around this switch, as well as resulting advantages and disadvantages, can be found in appendix A.3.

4.2 Important data to monitor in Nairobi (Q1)

In order to identify which type of would be important to monitor in Nairobi, literature about different methodologies for assessing NRW was analyzed (section 2.1). The information was used to conduct a well-structured interview with the NRW-manager of NCWSC, identifying which type of data was lacking the most for their existing NRW strategies.

The interview used the same framework for distinguishing different types of NRW as the IWA and the above mentioned literature, as its international standards help in creating a common understanding and terminology amongst water utilities world-wide. To give visual guidance in how different volumes of the IWA water balance depend on data, different "data-dependency trees" were constructed. An example of a data-dependency tree can be found in Figure 4.2 and all trees are listed in appendix A.2.

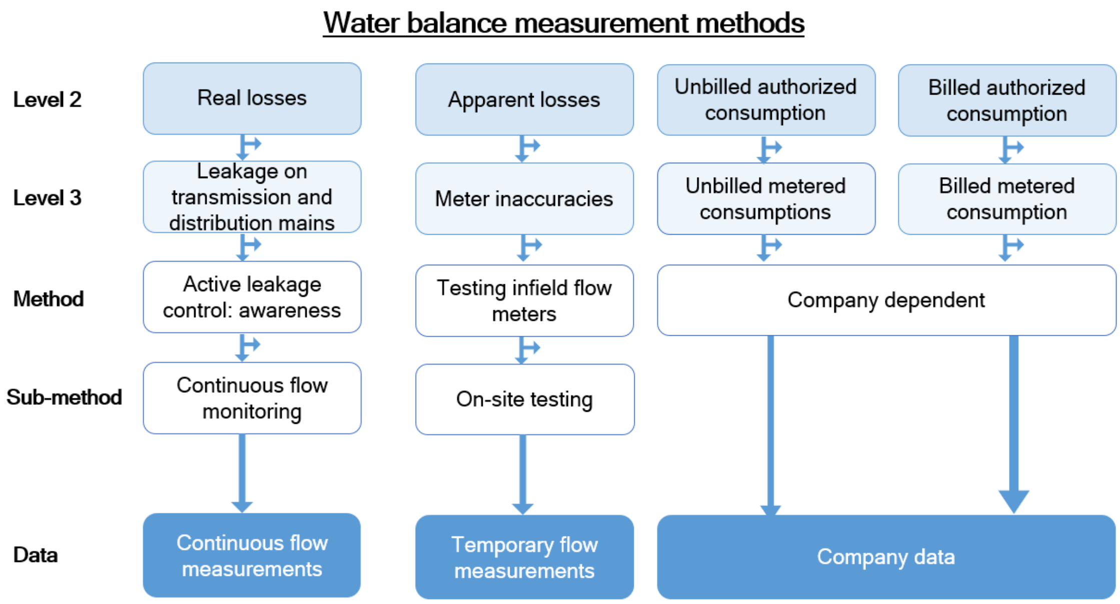

Figure 4.2: Examples of a data dependency tree. The different "levels" are explained in table 2.1

The purpose of the interview was to identify which volume of the IWA water balance (Table 2.1) is most difficult to measure for the utility of Nairobi. These volumes were identified by gradually discussing all data-dependency trees. The purpose of the monitoring system would be to give the local utility a useful tool to be better able to measure this specific volume, which turned out to be the hidden leaks in the system (appendix A.1).

4.3 Hydraulic modelling of IWS systems (Q2)

This section describes the methods which were used to construct the hydraulic model which can be used to simulate IWS conditions. It explains how the hydraulic model was constructed and calibrated with field measurements. The results of the hydraulic model can be found in chapter 5.

Strategy for constructing the hydraulic model

This thesis will make use of the demand patterns that were deducted from actual measurements and not use standard patterns, as demands tend to very different in IWS systems (De Marchis et al., 2010) and no standard IWS patterns could be found. The demand patterns that were deducted from analyzing the flow into Ashdown Park are shown in section 5.1.

It is advised to use pressure dependant outflow (PDO) modelling for IWS systems (Vairavamoorthy et al., 2007), as low pressures often influence the demands in these areas. Another advantage of PDO-modelling is that the flow through leaks can be modelled as a function of pressure (equation 2.1, section 2.3). Therefore, it was decided to use WNTR to construct the hydraulic model. WNTR is a Python-package, which can be built upon existing EPANET models and allows for pressure dependant outflow modelling (NTESS, 2019). It has previously been used to design leak detection and localization systems in research of Van Lagen (2020).

Harare Water Department and VEI shared the geographical outline of the drinking water system in Ashdown Park, by sharing its EPANET model. This EPANET model was loaded into WNTR. The larger distribution network outside the DMA was modelled by using a reservoir, whose head can be adapted to simulate the effects of changes in inlet pressure on the system (see Figure 4.3). The roughness coefficient of the pipes was the only physical parameter to be adjusted, since EPANET used the Darcy-Weisbach coefficient (measured in feet or meter) for calculating frictional energy losses and WNTR uses the unitless Hazen-Williams coefficient Rossman (2000). The roughness coeffecient was changed from 0.6×10^{-3} feet to 125 (-). The explanation behind these numbers can be found in appendix A.7.

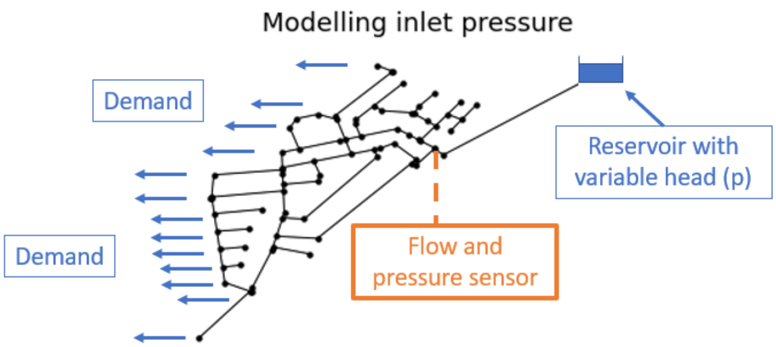

Figure 4.3: The model can use input values (in blue) to simulate the flow and pressure at the location of the sensor (in orange).

Calibrating the hydraulic model to field measurements

The two main parameters in the model that can be changed for the simulations are the nodal demand and the reservoir head, as can be seen in Figure 4.3. The hydraulic model in WNTR goes through an iterative process of solving mass and energy balances, until the remaining mass and energy deficits of the balances approach 0. As a solution, the model can calculate the flows that occur in pipes (noted by vector $\mathbf{q_t}$) and the pressure that occurs at nodes (noted by vector $\mathbf{h_t}$) that result from the demands in the nodes (noted by vector $\mathbf{d_t}$) and the pressure in the reservoirs and storage of the model (in this case noted by $\mathrm{h0_t}$, since only one reservoir is used). Hence, equation 4.1 shows the model with its output ($\mathbf{q_t}$ and $\mathbf{h_t}$) on the left side and its input ($\mathbf{d_t}$ and $\mathrm{h0_t}$) on the right side.

The resulting pressure at any node i ($h_{i,t}$) and flow at any pipe j ($q_{j,t}$) at time t can be read from the indices of vectors $\mathbf{h_t}$ and $\mathbf{q_t}$. The devices that measure flow and pressure at the DMA inlet (Figure 4.3) were estimated to be closest to pipe P73 and node N24. Therefore, the field measurements (measured flow noted by $\mathrm{q_t^*}$ and measured pressure noted by $\mathrm{h_t^*}$) should be compared with the computed flow at pipe P73 ($\mathrm{q_{j=P73,t}}$) and the computed flow at node N24 ($\mathrm{h_{i=N24,t}}$). Unfortunately the model's computations, using the initial demand settings in EPANET ($\mathbf{d_{init}}$[1]) and the average measured pressure at the reservoir ($\mathrm{h0_{init}}$), did not coincide with the measured values (equation 4.2). Therefore, the model needed to be calibrated.

$$f(\mathbf{q_t}, \mathbf{h_t}; \mathbf{d_t}, \mathrm{h0_t}) = 0 \tag{4.1}$$

$$f(\mathbf{q_t}, \mathbf{h_t}; \mathbf{d_{init}}, \mathrm{h0_{init}}) = 0 \implies \mathrm{q_{j=P73}} \neq \mathrm{q_t^*}, \quad \mathrm{h_{i=N24}} \neq \mathrm{h_t^*} \tag{4.2}$$

The calibration method introduces parameters δ_d and δ_{h0}, as well as variables $x1_t$ and $x2_t$. The demands ($\mathbf{d_{init}}$) and the pressure in the reservoir ($\mathrm{h0_{init}}$) are multiplied by the former parameters, to make the model's output of pressure at node N24 and flow at pipe P73 approach the measured values (equation 4.3). Parameters δ_d and δ_{h0} are input values of the model that influence the accuracy of the calibration (section 5.3). Variables $x1_t$ and $x2_t$ are defined by the perturbations caused by adding extra demand to the model (vector **c** [c1, c2], varies with δ_d) and extra pressure to the model's reservoir (vector **d** [d1, d2], varies with δ_{h0}), as well as the difference between the measurements ($\mathrm{h_t^*}$ and $\mathrm{q_t^*}$) and the calculated pressure at node N24 and flow at pipe P73 ($\mathrm{h_{i=N24,t}}$ and $\mathrm{q_{j=P73,t}}$). They are the unique solution to the linear set of equations shown by equation 4.4. A derivation of this calibration method, reported in detail for the example with δ_d=1 and δ_{h0}=1, can be found in appendix A.4. Its results are shown in section 5.3.

$$\begin{gathered} f(\mathbf{q_t}, \mathbf{h_t}; \mathbf{d_{init}} \times (1 + \delta_d \times x1_t), \mathrm{h0_{init}} \times (1 + \delta_{h0} \times x2_t)) = 0 \implies \\ \mathrm{q_{j=P73,t}} \approx \mathrm{q_t^*}, \quad \mathrm{h_{i=N24,t}} \approx \mathrm{h_t^*} \end{gathered} \tag{4.3}$$

$$\begin{bmatrix} c1 \\ c2 \end{bmatrix} \times x1_t + \begin{bmatrix} d1 \\ d2 \end{bmatrix} \times x2_t = \begin{bmatrix} (\mathrm{h_t^*}/\mathrm{h_{i=N24}} - 1) \\ (\mathrm{q_t^*}/\mathrm{q_{j=P73}} - 1) \end{bmatrix} \tag{4.4}$$

[1] In the EPANET-file received from VEI there was already a demand pattern inserted to change the demands over time. However, using this demand pattern, the output of the model did not approach the measurements at any time. Therefore, the calibration method will "override" the inserted demand pattern and only the initial base demands $\mathbf{d_{init}}$ are taken into account at this stage.

Accuracy of calibration

The accuracy of the model output can be estimated by calculating the Root Mean Squared Error (RMSE). Since this RMSE can be computed for both the pressure and the flow output, which both have different units, the Normalized Root Mean Squared Error (NRMSE) is used. This allows for a better comparisons between the two errors. The expressions of these NRMSE's are given in equation 4.5 and 4.6.

$$NRMSE_{flow} = \frac{\sqrt{\frac{\sum_{t=0}^{143}(\mathrm{q}_\mathrm{t}^{*}-\mathrm{q}_{\mathrm{j=P73,t}})^2}{N}}}{\mathbf{avg(q^{*})}} \tag{4.5}$$

$$NRMSE_{pressure} = \frac{\sqrt{\frac{\sum_{t=0}^{143}(\mathrm{h}_\mathrm{t}^{*}-\mathrm{h}_{\mathrm{i=N24,t}})^2}{N}}}{\mathbf{avg(h_t^{*})}} \tag{4.6}$$

4.4 Leak detection with Boolean matrices (Q3)

This section describes the strategy that is used for leak detection, which is based on threshold values of flow and pressure that can be expected in a water supply system. It proposes two methods to model different demand realizations that can be used to model these thresholds. At last, it explains how Boolean matrices can be used to store information whether a sensor at a certain location can detect leaks that emerge in the DMA. The results of applying these methods can be found in chapter 6.

Threshold based leak detection strategy

The monitoring system within the DMA should give the water utility information whether a new hidden leak has emerged in the area. The measurements need to be easily interpretable, so that system operators can respond fast when the leaks emerges. Since this thesis is conducted in cooperation with Waternet, the water utility in Amsterdam, it was first investigated how awareness of hidden leaks is created in Amsterdam. Here, a system is used which is called "Dynamische Bandbreedte Monitor", which translates to "Dynamic Bandwidth Monitor (DBM)" in English. It consists of several flow meters which are installed to monitor the flow into several DMAs (e.g. Figure 4.4 and 4.5).

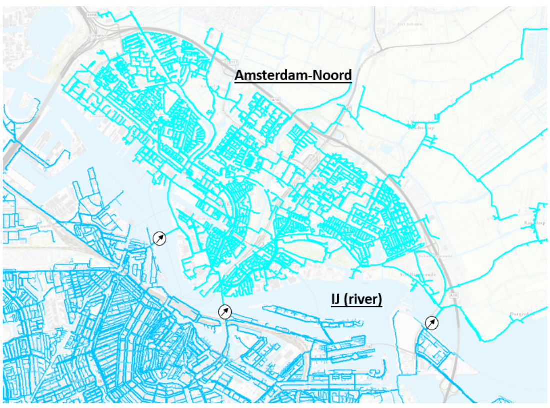

Figure 4.4: Location of sensors (with arrows) for the DBM that monitors the flow to Amsterdam-Noord (light-blue DMA).

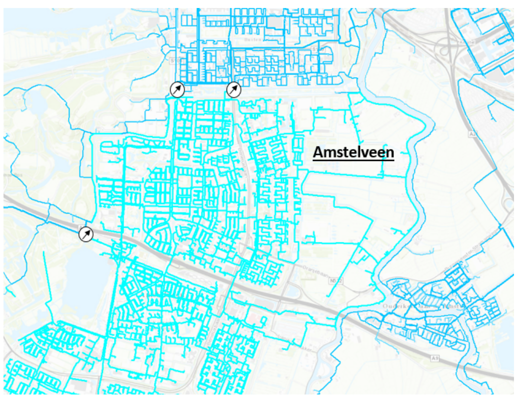

Figure 4.5: Location of sensors (with arrows) for the DBM that monitors the flow to Amstelveen (light-blue DMA).

The devices send the measured values to the control room of the water utility, where they are compared to an alarm value that has been set based on historical measurements. When the real-time measurements exceed this alarm value for a certain period of time, it is assumed that a leak has emerged. An example of a display of the real-time measurements (in red), a prediction of expected values (in black) and the alarm value (in blue) is given in Figure 4.6. Carefully a leak time, a specific time frame that the measurements need to be above the alarm value, can help in preventing the utility from raising false alarms. In the Netherlands, it is advised to choose smaller detection times during the night than during the day since the nocturnal flow has less fluctuations (Van Steen, 2020).

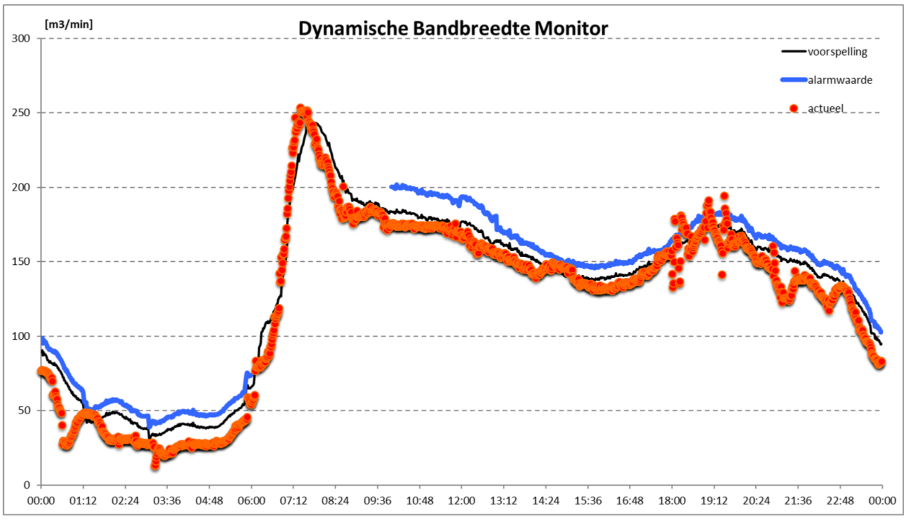

Figure 4.6: Example of applying the Dynamic Bandwidth Monitor.

The above system was used as an inspiration to design the monitoring system in Harare[2] for leak detection purposes. However, the DBM uses a historical data-set at a location where the sensor has already been installed (Van Vossen-van den Berg, 2017). When a new leak detection system is designed, the sensors are not yet located in the network. Therefore, there is no historical data set available and the expected flows and pressures have to be modelled at the locations which are suitable for sensor placement. The range of calculated flows and pressures at a certain location can be referred to as "artificial DBM's" as they predict the ranges of flow and pressure that would have been measured by a flow or pressure logger at that location. In this thesis artificial DBM's were constructed for all pipes, since every pipe was considered as a potential sensor location. The performance of each artificial DBM is measured by how many leaks it detects. The DBM's which can detect the most leaks are marked as the optimal location to place sensors.

[2]The DMA that is used for the case study in Harare is way smaller than the DMAs in Amsterdam (+/- 450 connections in Harare vs. +/- 40.000 connections in Amsterdam), but the same threshold based method for leak detection (the Dynamic Bandwidth Monitor) can be applied in both cases.

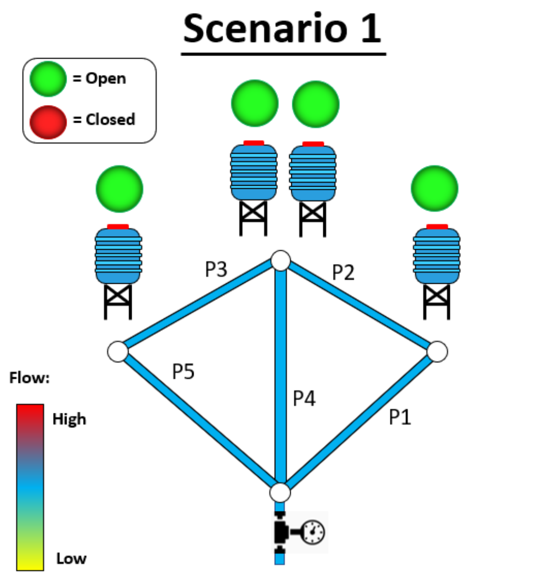

Figure 4.7: Scenario 1: Flow in system when all households subtract water.

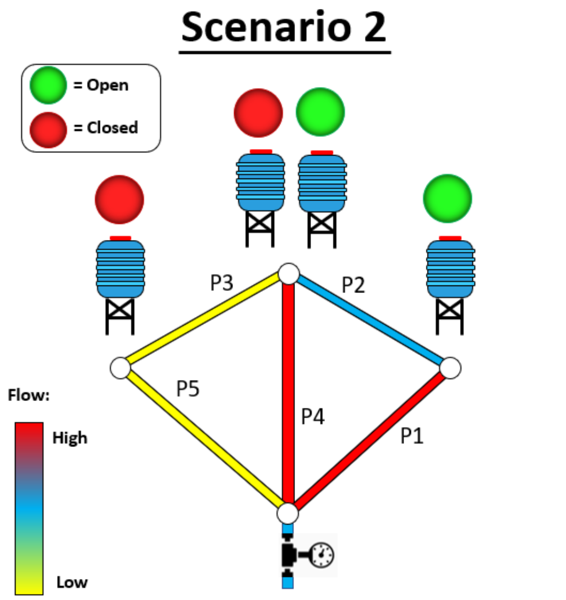

Figure 4.8: Scenario 2: Flow in system when two of the four households subtract water.

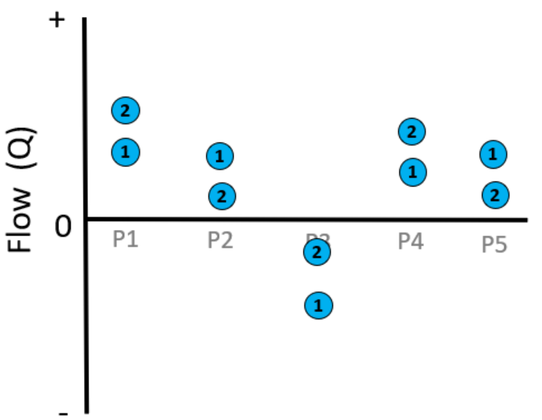

Figure 4.9: Example of possible flow measurements from scenario 1 and 2.

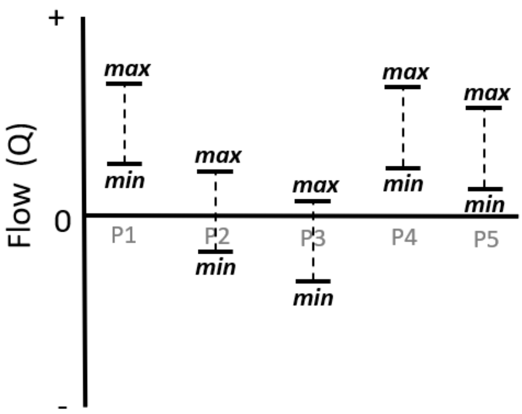

Figure 4.10: The flow measurements from all scenarios are converted into alarm values.

Figure 4.7 up to 4.10 show how artificial DBM's can be created for the pipes in a simple exemplary network. Scenario 1 and 2 show the flow in the network during different demand configurations. These flow values are plotted in Figure 4.9. When all possible flow values in the pipes are recorded, by modelling all possible demand scenarios, the range within which the flow values can be expected to be are given in Figure 4.10. These ranges can be considered as artificial DBM's for pipe 1 up to 5. In this example the range is constructed from flow values, but also pressure-based DBM's can be constructed when the possible pressure in the nodes is modelled. Whereas other research focuses mostly on pressure sensors for leak detection and localization (Perez et al., 2009)(Khorshidi et al., 2020), this thesis design two systems, one with flow and one with pressure sensors, allowing for a comparison between the two. Furthermore, it will design the monitoring system specifically for IWS systems.

Methods for modelling demand realizations to construct alarm values for the DBM

The proposed system for leak detection, the DBM, compares real time data with a range of expected values in "normal" situations, so situations without leaks, and attributes deviations from the expected values to a potential leak. The key to a successful DBM is the construction of the boundaries for the expected range of values, the so called "alarm values". These alarm values represent the maximum deviations from the average scenario, which still can be expected to happen. The boundary scenarios for flow in pipes often occur when the demand concentrates locally. Especially a system which has a small amount of households connected to its nodes, as is the case for Ashdown Park (Figure 4.11), can be sensitive to situations where all or none of the houses at a single node demand water at the same time. In this thesis, two methods were proposed to model this demand allocation.

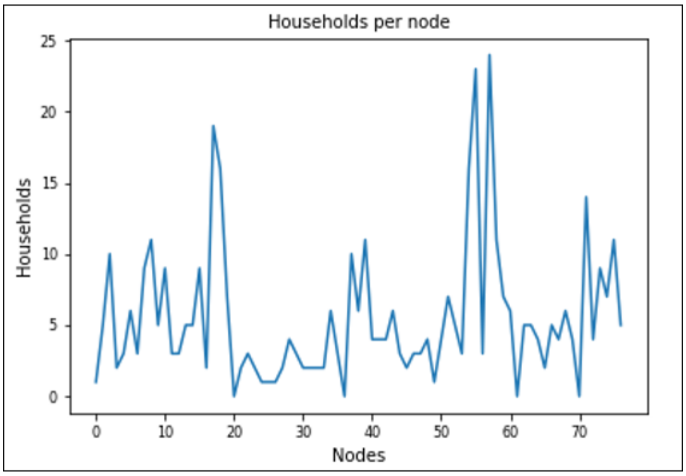

Figure 4.11: Number of households per node in Ashdown Park

Both approaches consist of running several Monte Carlo simulations, thereby saving the most extreme flows and pressures from these experiments as alarm values. The first method that was designed includes creating demand realizations by random draws from a probability density function, based on the mean value and deviation of the incoming flow. This method was eventually not considered feasible for Ashdown Park and shall only explained shortly. The second method uses a weighted random choice and adds a single parameter with a physical meaning in IWS systems, the tap capacity. This method deemed to be more feasible and shall be explained in more detail.

Method 1: Monte Carlo simulation with demand realizations as random draws from a normal distribution

First, a normal distribution is constructed by using the mean and standard deviation from the historical data of the flow that enters the DMA. An extensive explanation of how the mean and standard deviation was constructed from this historical data-set can be found in appendix A.11. A similar, scaled normal distribution[3] is constructed for the

[3] The normal distribution for the demand in node i is scaled by multiplying the standard deviation and average of the normal distribution of the inflow ($N(Q_{avg,t}, \sigma_{Q,t})$) by the ratio of the demand i to the inflow

demand at every node. Then, a Monte Carlo simulation is performed, with each nodal demand being determined by a random draw from its scaled normal distribution. Running several experiments with this model allows us to calculate the different flows that can occur in the network. For a more detailed explanation of this method please read appendix A.12.

This method gives a small chance that the demand is much lower than the average demand and a small chance that it is much higher. However, it does not at any time return the possibility of having no demand. Figure 4.11 shows us that some nodes only contain few households. The chance that none of the households uses water is therefore quite realistic. So, the flows that are calculated using this method are likely to be over- or underestimated, making this method unsuitable for constructing the artificial DBMs. If the modelled system would be be a water transport network of an IWS system, whose DMAs would be presented as nodes (suggested in section 5.5), then this method could be appropriate since its nodes would always have some demand.

Method 2: Monte Carlo simulation with demand as weighted random choice with tap capacity parameter

Adding a parameter with a physical meaning[4] has resulted into developing this second method. The method assumes that every household has one tap that is directly connected to the distribution system. This assumption is specifically suitable for IWS systems, since many people have their own water storage. The only tap that is directly connected to the distribution system is the tap that is used for filling their storage (see Figure 4.12). Usually, the storage is filled with manually operating the tap above the storage (MSc-Student University of Bulawayo, 2020).

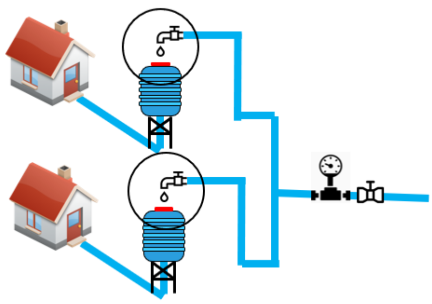

Figure 4.12: One household uses one tap to withdraw water from the distribution network.

Since most people would want to fill their storage as quickly as possible, it is assumed that the tap will mostly be fully open or closed. The volume that flows through a tap that is fully open will be referred to as the Tap Capacity (TC). This flow rate can differ a lot geographically, as shown by the different flows that were recorded flowing through the tap in different places in the world (Table 4.1). Developed countries regularly put water saving features in their taps, resulting in relatively low flow rates. The absence of such features, as well as the urgency to collect water as quickly as possible during times

($\frac{D_{i,t}}{Q_{avg,t}}$), resulting in a scaled normal distribution for the demand in node i ($N(D_{i,t}, \frac{D_{i,t}}{Q_{avg,t}} \times \sigma_{Q,t})$).

[4]This is often done in in the stochastic end-use model SIMDEUM, that is being used to model drinking water demand in the Netherlands (Blokker, 2020).

of water availability, could all be reasons for higher flow rates in Ghana and South Africa. However, due to the limited pressure and water supply in IWS zones, flow rates could also be low, as was the case in India. Two experiments performed in Ashdown Park, explained in more detail in appendix A.9, showed an average flow rate from the tap of 0.29 m3/h, which will be rounded to 0.3 m3/h. Therefore, a TC of 0.3 m3/h will be assumed for modelling demand realizations in Ashdown Park.

Tap Capacity (TC)	Value (m3/h)
Standard indicator TU (Tapping Unit) the Netherlands (TU Delft, 2020)	0.3
South Africa field survey; tap fully open (Jacobs et al, 2015)	1.1
University Ghana (Oduro-Kwarteng et al, 2009)	0.95
IWS region India field survey (Kumpel et al, 2017)	0 – 0.33
Ashdown Park field experiments	0.27 – 0.3

Table 4.1: Flow rates from a tap in different researches (TU Delft, 2020b)(Jacobs et al., 2015)(Oduro-Kwarteng et al., 2009)(Kumpel et al., 2017)

The total demand of the DMA can be estimated once the hydraulic model has been calibrated using the flow and pressure measurements at the DMA inlet (section 5.3). It can be estimated how many storage taps are open by dividing the total demand at that moment by the TC. For example, at 00:00 the incoming flow and the total demand are 35.5 m3/h on average, implying that $\frac{35.5}{0.5} = 71$ of the 424 taps are open. Modelling all possible demand realizations with this amount of open taps would result in 8.2×10^{81} possible combinations (appendix A.16) and is therefore regarded as impossible. It is however possible to run a Monte Carlo simulation, modelling the demand at every household as a weighted random choice between a demand of 0.3 m3/h (implying an open tap) or 0 m3/h (implying a closed tap). The weight for choosing an open tap equals the ratio of open taps to the total amount of taps (e.g. 71/424) and the weight for choosing a closed tap equals the ratio of closed taps (e.g. 353/424). An example of applying this random weighted choice in a fictive network with five households is shown in Figure 4.13.

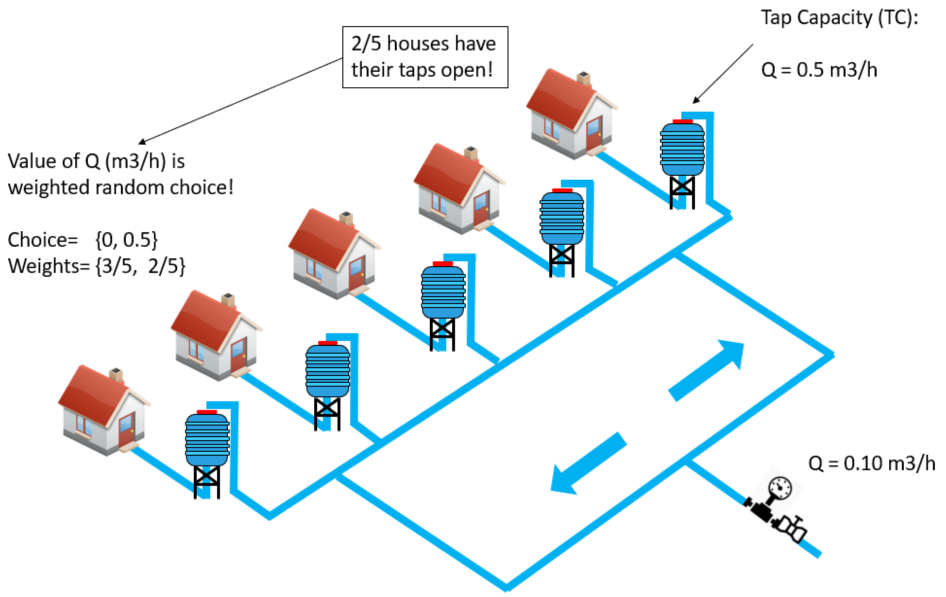

Figure 4.13: Example of demand allocation using weighted random choice and a tap capacity (TC).

Using Boolean matrices for leak detection

The scenario with a leak in a certain pipe can be modelled and its result compared to the constructed alarm values in the previous chapter. For this comparison, only situations where a single leak occurs in the network were considered. It is important that, except from the added leak, the other conditions remain equal to the conditions during which the alarm values were constructed. Comparable leak sizes as mentioned in Casillas Ponce et al. (2013) will added to hydraulic model to simulate scenarios with a leak. However, the leak size in the hydraulic model of this thesis will be defined by its area and dependent on the occurring pressure, as these physical relations are important for the leak flow rate (Marchis and Milici, 2019). The resulting flow in the pipes of the scenario with leak can be compared to the alarm values, of which an example can be seen in Figure 4.14. A sensor in the pipes where the flow exceeds the alarm values could have detected the leak, whereas a sensor in the pipes where the flow remains within the alarm values could have not. The same principle holds for the new pressure at nodes that result from a leak in a certain pipe segment.

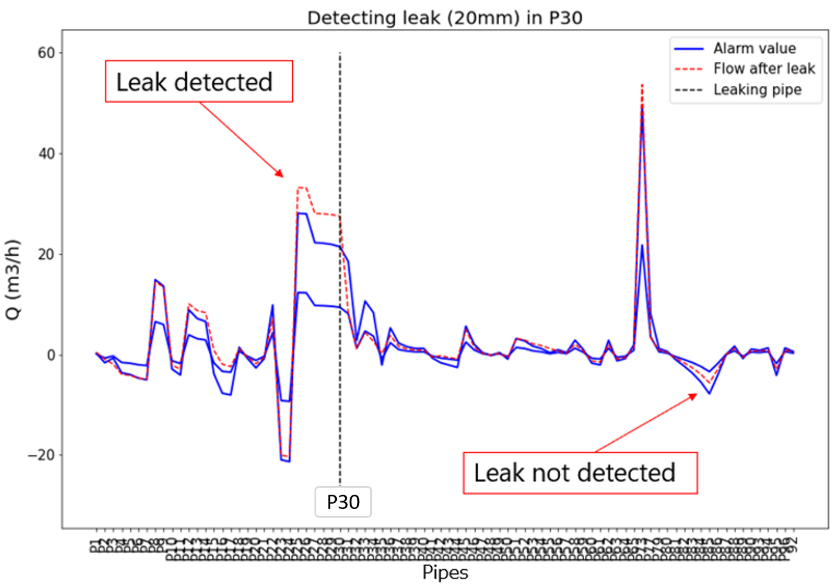

Figure 4.14: Example of leak detection with flow sensors at 00:00.

This process, adding a leak in a certain pipe and deciding whether a sensor at a certain pipe or node could have detected a leak, can be repeated by adding a leak at a different pipe every time. The ability of the network to detect leaks can then be stored in a Boolean matrix, which is often done in leak detection methods (Perez et al., 2009)(Khorshidi et al., 2020). When designing a monitoring system with flow sensors, the potential locations to place sensors equal the number of pipes in the network. So, if there are n pipes in the network, there are n potential locations to place sensors. Therefore, it is possible to construct a nxn Boolean matrix $A_{i,j}$, which stores information whether a leak in pipe i can be detected ($A_{i,j} = 1$) by a flow sensor in pipe j or not ($A_{i,j} = 0$), see equation 4.7.

$$A_{ij} = \begin{cases} 1, & \text{A leak in pipe i } \textit{can} \text{ be detected by a sensor in pipe j} \\ 0, & \text{A leak in pipe i } \textit{cannot} \text{ be detected by a sensor in pipe j} \end{cases} \tag{4.7}$$

A similar Boolean matrix $A_{i,j}$ can be constructed when designing a monitoring system with pressure sensors, only this design has m potential sensor locations if the network has m nodes. Therefore this matrix has size nxm. It stores the information whether a leak in pipe i can be detected ($A_{i,j} = 1$) by a pressure sensor in node j or not ($A_{i,j} = 0$), see equation 4.8.

$$A_{ij} = \begin{cases} 1, & \text{A leak in pipe i } \textit{can} \text{ be detected by a sensor in node j} \\ 0, & \text{A leak in pipe i } \textit{cannot} \text{ be detected by a sensor in node j} \end{cases} \tag{4.8}$$

However, different demand realizations play a large role as well which can be explained by expanding the example of Figure 4.7 and 4.8. Figure 4.15 shows the same network, with a leak added in pipe 5. All nodes subtract water and the leak in pipe 5 could have been detected by flow sensors in pipe 3 and 5 (Figure 4.16). However, if the demand realization changes, this will also change the ability of the network to detect leaks. The leak in Figure 4.17 forces the flow to the left, but the demand forces the flow to the right since only the houses on the right subtract water. This dampens out the effect of the leaks and results in the situation that the leak can not be detected anymore (Figure 4.18).

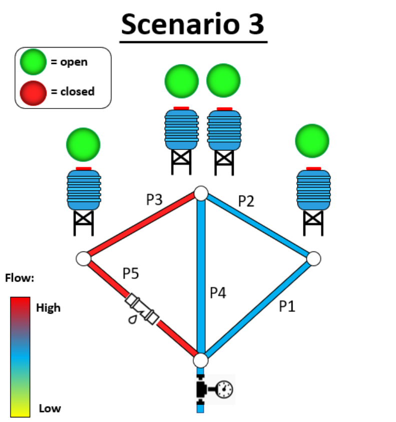

Figure 4.15: Scenario 3: A leak in pipe 5 and all houses subtract water.

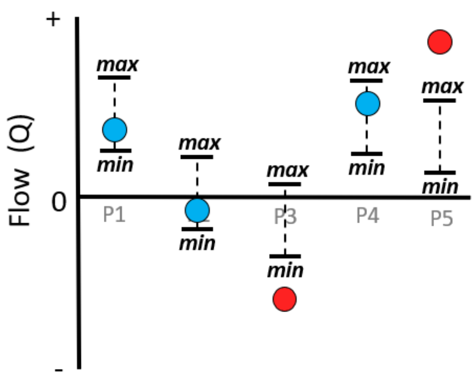

Figure 4.16: Results Scenario 3: The leak can be detected by flow sensors in pipes 3 and 5, since they measured value exceeds the alarm values.

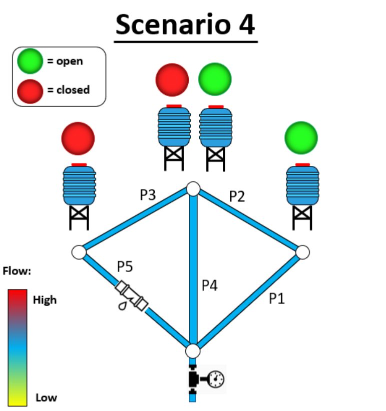

Figure 4.17: Scenario 3: A leak in pipe 5 and the houses on the right subtract water.

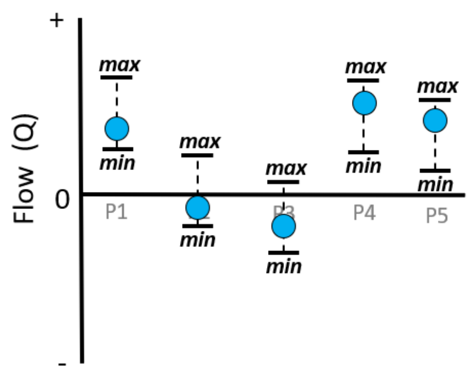

Figure 4.18: Results Scenario 4: The leak cannot be detected by any flow sensors.

So, a different demand realization changes the matrix A, which stores the ability of the network to detect leaks. As previously explained, it is possible to model a range of possible demand realizations by using a Monte Carlo simulation, letting each demand realization be determined by a random weighted choice. When constructing the matrix which stores the detectability of leaks, the same Monte Carlo simulation is used to model the different demand realizations, only this time leaks are added to the network at every realization. Every experiment within the Monte Carlo simulation yields a different matrix A. The result is a large 3D-matrix A ($A_{i,j,k}$), which stores information of whether a leak in pipe i can be detected by a sensor in pipes j in experiment k, as can be seen in Figure 4.19. Taking this third dimension into account in the Boolean matrix gives insight in how demand allocation influences leak detection and can be regarded as a new feature in leak localization and detection methods, which usually only take a single demand realization into account (Perez et al., 2009)(Khorshidi et al., 2020).

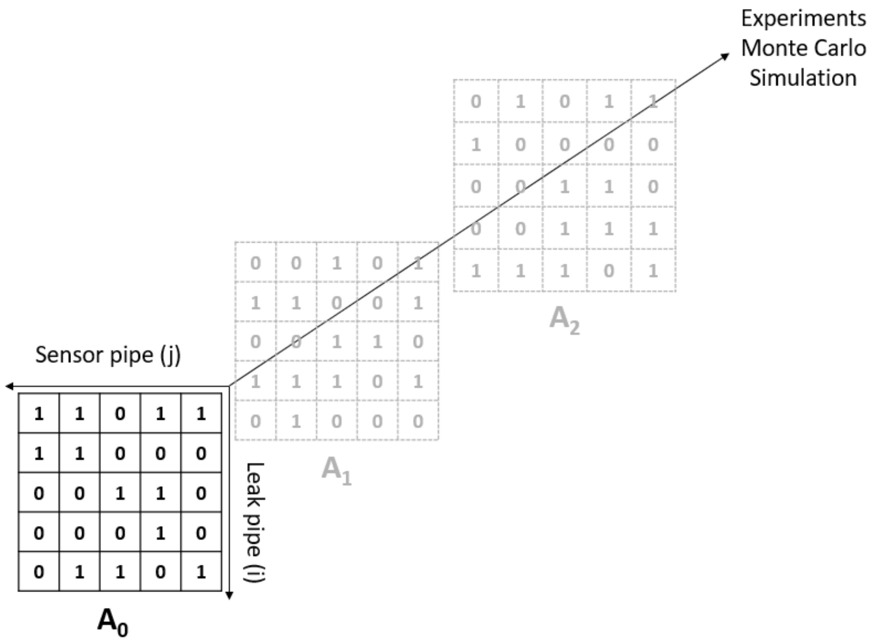

Figure 4.19: Different Boolean matrix with every new experiment.

Key Performance Indicators

The performance of the hydraulic model to detect leaks can be measured by using key performance indicators (KPIs), which result from analyzing the Boolean 3D-matrix. A well-functioning DBM should return a low *percentage of false alarms* and a high *detectability*. Both KPIs shall be explained below.

Percentage of false alarms

The performance of the model can be measured by computing a certain % of false alarms. This can be best explained by taking the example of adding an extreme small leak to the model, for instance with a diameter of 0.1mm. Using common sense, the leak would not be detectable for any sensor. However, depending on the number of experiments chosen to construct alarm values on one side and to model demand realizations with leaks on the other, the constructed 3D-matrix still shows some entries being "1". This then does not happen because of the leak, but because an extreme demand realization has occured in constructing the 3D-matrix that was not taken into account when constructing the alarm values. The % of false alarms is defined as the amount of 1's in the 3D matrix, when applying a leak of 0.1mm, divided by the total amount of entries in the 3D-matrix.

Detectability

The ability of the network to detect leaks at a certain time can be expressed as the "average number of leaks detected per sensor per scenario". Let us refer to this KPI as *Detectability (Dt (t))*. For a network with m nodes, n pipes, where the demand allocation is modelled with a Monte Carlo simulation with p simulations, the detectability of flow sensors at a certain time is then given by equation 4.9.

$$Dt(t) = \frac{1}{np} \sum_{j=1}^{n} (\sum_{i=1}^{n} (\sum_{k=1}^{p} A_{i,j,k}(t))) \tag{4.9}$$

The detectability of pressure sensors at a certain time for the same network is given by equation 4.10.

$$Dt(t) = \frac{1}{mp} \sum_{j=1}^{m} (\sum_{i=1}^{n} (\sum_{k=1}^{p} A_{i,j,k}(t))) \tag{4.10}$$

4.5 Optimizing sensor placement (Q4)

This section describes the optimization techniques that were applied to calculate the optimal sensor configuration, that can detect a *maximum* amount of leaks using a *minimum* amount of equipment. It starts with describing the optimization technique which can be used for a single demand realization and from there expands the method to make it applicable for multiple demand realizations. At last it shows how the performance of different sensor configurations can be assessed.

Single scenario optimization

First a closer look is taken how the optimal sensor allocation of a single matrix, so of one demand realization, shall be taken. This 2D-matrix ($A_{i,j}$) can be right-multiplied by a vector ($\mathbf{x}$) that represents whether a pipe contains a sensor ($x_j = 1$) or not ($x_j = 0$). Vector $\mathbf{x}$ is a Boolean vector, so it only contains the entries 0 or 1. This results in a matrix-vector product ($\hat{\mathbf{y}}$) which stores information how many times a leak in a certain pipe i is detected by the installed sensors. This vector $\hat{\mathbf{y}}$ will be referred to as the estimator. Ashdown Park contains 84 pipe segments, resulting in an estimator of equal size. To give a simple overview of the explanation above, Figure 4.20 is created. This is not a result of the analysis for Ashdown Park, but just an example to explain the story.

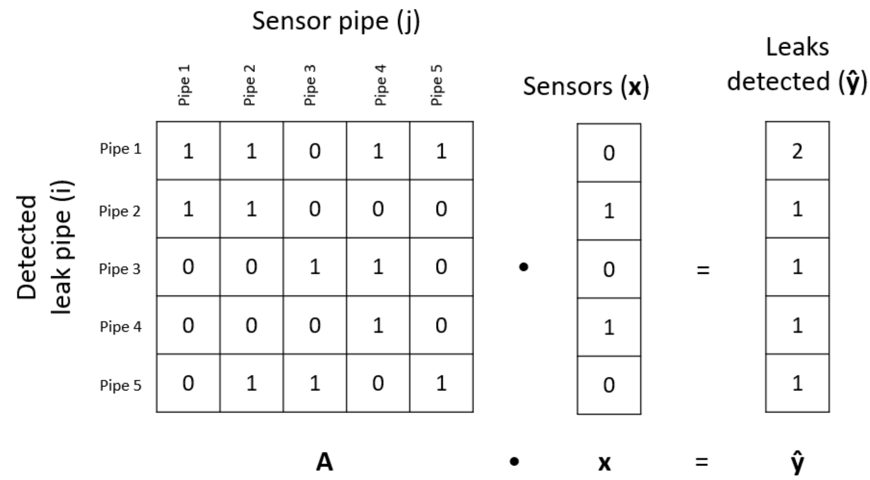

Figure 4.20: Example of Matrix calculation.

Objective for single scenario optimization

The unknown vector to solve in this problem is vector $\mathbf{x}$, which contains the information at which pipe segments sensors should be installed. This problem can be solved by applying a minimizing objective and constraints. The main objective is to have every entry of vector $\hat{\mathbf{y}}$ as close to 1 as possible. The entries should be above 0, because it then can detect a leak a certain pipe. Entries above 1 are less favorable since you then have two sensors detected the same leak, which can be considered as double work. This objective is shown in equation 4.11.

$$\begin{aligned} \min_{x} \quad & ||\hat{\mathbf{y}} - \mathbf{y}||^2 \\ \text{s.t.} \quad & \hat{\mathbf{y}} = \mathbf{A}\mathbf{x} \qquad \text{, with } \mathbf{x}=\{0,1\} \\ & \mathbf{y} = \mathbf{1} \end{aligned} \tag{4.11}$$

There is a second objective, which concerns the unknown vector $\mathbf{x}$ directly and states to use as few sensors as possible. Using a minimum amount of sensors is a requirement for the optimal allocation as defined at the beginning of this chapter. This objective is shown in equation 4.12.

$$\min_{x} \quad ||\mathbf{x}||^2 \qquad \text{, with } \mathbf{x}=\{0,1\} \tag{4.12}$$

These two objectives show the trade-off that has to be made. By favouring the objective of equation 4.11 over 4.12, it would be implied that it is very important to find all the leaks, so a lot of sensors can be used. By favouring the objective of equation 4.12 over 4.11, it is very important to use a little amount of sensors and it is fine to compromise on leak detection. This kind of trade-offs can, in multi-objective functions, be expressed by introducing a parameter λ (equation 4.13). If λ would be larger than 1, this would imply that achieving the second objective is preferred over achieving the first objective. If λ would be smaller than 1, the first objective would be preferred over the second objective.

$$\min_{x} \quad ||\hat{\mathbf{y}} - \mathbf{y}||^2 + \lambda ||\mathbf{x}||^2 \qquad \text{, with } \mathbf{x}=\{0,1\} \tag{4.13}$$

Another way of adding the numbers of sensors to the optimization function, would be to add the objective that minimizes $\mathbf{x}$ as a constraint, which depends on the number of sensors that can be used in the network. For example, if no more than two sensors are allowed in the network, the constraint that is added will be: $||\mathbf{x}||^2 \leq 2$. Solving this optimization problem for a different number of sensors (i) shows us which amount of water can be saved by using a certain number of sensors, as well as where they should be placed (see equation 4.14). The optimization solver Gurobi (Gurobi Optimization LLC, 2020) can be used to solve the optimization problem.

$$\begin{aligned} \min_{x} \quad & ||\hat{\mathbf{y}} - \mathbf{y}||^2 \\ \text{s.t.} \quad & \hat{\mathbf{y}} = \mathbf{A}\mathbf{x} \qquad \text{, with } \mathbf{x}=\{0,1\} \\ & ||\mathbf{x}||^{\mathbf{2}} \leq i \\ & \mathbf{y} = \mathbf{1} \end{aligned} \tag{4.14}$$

Weighted optimization

As previously described, the water volume that is lost during leakage is different for places with different pressures. In order to make sure that the sensors in the designed network prefer detecting pipes where high losses occur over pipes where low losses occur, weights are added to the optimization problem. This is done by constructing a vector $\mathbf{w}$, whose individual entries (w_i) represent the fraction of the volume lost from a leak in pipe i, compared the the sum of volumes lost from leaks in all pipes (see equation 4.15).

$$w_i = \frac{d_{leak,i}}{\sum_{j=1}^{84}(d_{leak,j})} \tag{4.15}$$

These weights are applied to the previous objective (equation 4.14). Adding a stronger weight to pipe i, will prefer detecting a leak at pipe i over the other pipes. The new weighted objective function can be seen in equation 4.16[5].

$$\begin{aligned} \min_{x} \quad & ||\sqrt{\mathbf{w}}(\hat{\mathbf{y}} - \mathbf{y})||^2 \\ \text{s.t.} \quad & \hat{\mathbf{y}} = \mathbf{A}\mathbf{x} \qquad \text{, with } \mathbf{x}=\{0,1\} \\ & ||\mathbf{x}||^{\mathbf{2}} \leq i \\ & \mathbf{y} = \mathbf{1} \end{aligned} \tag{4.16}$$

Multi scenario optimization

The objective function from equation 4.16 can be solved for a single scenario, since it includes one Boolean A-matrix. However, different demand realizations lead to a 3D-Boolean matrix ($A_{i,j,k}$), where the third dimension k represents the number of experiments from a Monte Carlo Simulation. Finding the optimal solution for the sensor allocation with this multitude of experiments looks similar to the previous objective function (equation 4.14), but is expanded on some points.

[5]Adding these weights can be written as $||\sqrt{\mathbf{w}}(\hat{\mathbf{y}} - \mathbf{y})||^2$, since this can be broken down into: $||\sqrt{\mathbf{w}}(\hat{\mathbf{y}} - \mathbf{y})||^2 = w_1(\hat{y_1} - y_1)^2 + w_2(\hat{y_2} - y_2)^2 + ... + w_n(\hat{y_n} - y_n)^2$, with n being the number of pipe segments in the DMA.

This optimization objective can be explained by Figure 4.21. Right-multiplying the different 2D Boolean matrices $(\mathbf{A_0}, \mathbf{A_1}, ..., \mathbf{A_k})$ by the same $\mathbf{x}$-vector, results in different estimators $(\hat{\mathbf{y}}_\mathbf{0}, \hat{\mathbf{y}}_\mathbf{1}, ..., \hat{\mathbf{y}}_\mathbf{k})$. The indices in the different estimators should still be as all as close to 1 as possible, since leaks should be detected in as many experiments as possible. The $\mathbf{x}$-vector remains the same in all scenarios, since only one sensor configuration can be chosen. Another difference with the single scenario optimization, is that the weight vector $(\mathbf{w})$ changes with every scenario[6], resulting in different weight vectors $(\mathbf{w_0}, \mathbf{w_1}...\mathbf{w_k})$.

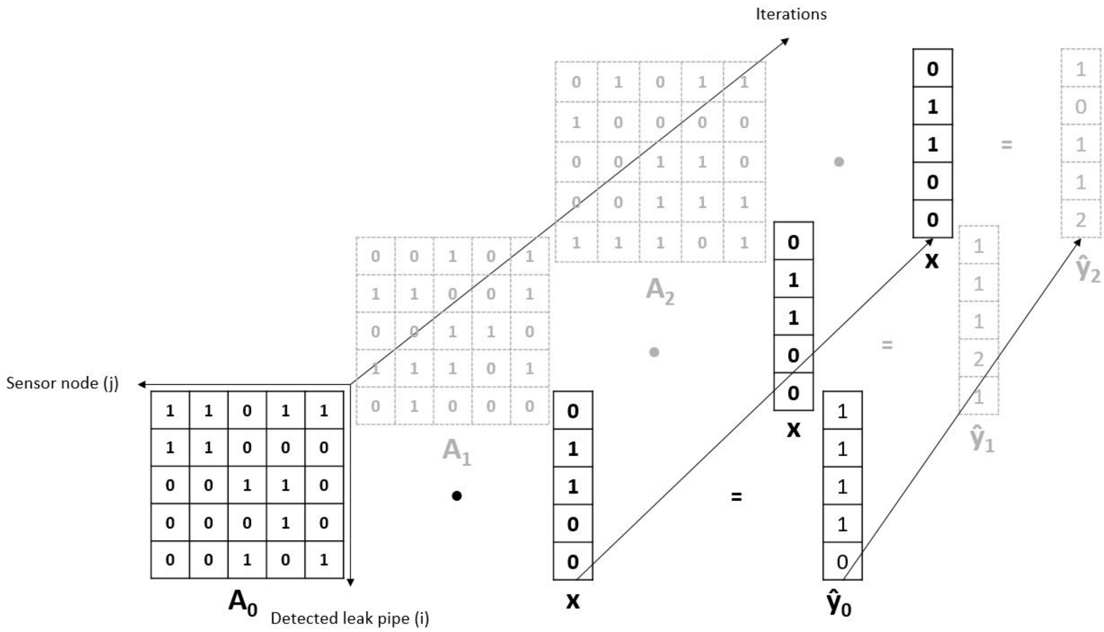

Figure 4.21: Example of right-multiplying several scenarios.

The different estimators $(\hat{\mathbf{y}}_\mathbf{0}, \hat{\mathbf{y}}_\mathbf{1}, ..., \hat{\mathbf{y}}_\mathbf{k})$ are added as constraints to the objective function. The objective itself is expanded, so that all the estimators approach the 1-vector $\mathbf{y}$. The weights by which the individual entries of the estimators approach 1 are determined by weight vectors $(\mathbf{w_0}, \mathbf{w_1}...\mathbf{w_k})$. The objective function, shown in equation 4.17, can be solved for a different numbers of sensors (i) using optimization solver Gurobi.

$$
\begin{aligned}
\min_{x} \quad & \sum_{i=0}^{k} ||\sqrt{\mathbf{w_i}}(\hat{\mathbf{y}}_\mathbf{i} - \mathbf{y})||^2 \\
\text{s.t.} \quad & ||\mathbf{x}||^\mathbf{2} \leq i \\
& \mathbf{y} = \mathbf{1} \\
& \hat{\mathbf{y}}_\mathbf{0} = \mathbf{A_0 x} \qquad , \text{with } \mathbf{x}=\{0,1\} \\
& \hat{\mathbf{y}}_\mathbf{1} = \mathbf{A_1 x} \qquad , \text{with } \mathbf{x}=\{0,1\} \\
& ... \\
& \hat{\mathbf{y}}_\mathbf{k} = \mathbf{A_k x} \qquad , \text{with } \mathbf{x}=\{0,1\}
\end{aligned}
\tag{4.17}
$$

[6]This change is caused by a changing leak flow rate, due to a slightly different nodal pressure in each demand realization.

Performance of optimal sensor configurations

The output of the algorithm that optimally places flow and pressure sensors should be able to state whether a configuration of sensors is able to detect leaks in certain pipes. The performance of a certain configuration ($\mathbf{x}$-vector) can be assessed by interpreting 2D-matrix $\hat{\mathbf{Y}}$, with $\hat{\mathbf{y}}_\mathbf{0}$, $\hat{\mathbf{y}}_\mathbf{1}$, ..., $\hat{\mathbf{y}}_\mathbf{k}$ as columns. This matrix $\hat{\mathbf{Y}}_{\mathbf{i,k}}$ represents how many times a leak in pipe i can be found in experiment k. The nonzero-function is used to see if a leak is detected in a certain scenario. This function returns whether a value ($\hat{\mathbf{Y}}_{\mathbf{i,j}}$) is **not** zero, in which case it returns a 1 ($\mathbf{nnz}(\hat{\mathbf{Y}}_{\mathbf{i,j}})$=1) and if the entry is zero, it returns a 0. With this function it is possible to express the percentage of leaks in pipe i that can be found by the sensor allocation ($\mathbf{x}$) in the different scenarios. This "percentage of leaks found in pipe i" is expressed as pf_i. It formula is given by equation 4.18.

$$pf_i = \frac{\sum_{i=0}^{k}(\mathbf{nnz}(\hat{\mathbf{Y}}_{\mathbf{i,k}}))}{k} \times 100 \tag{4.18}$$

However, the above probability can not be translated one to one to the daily chances of finding a leak. For example, it is very likely that if the results above show a 1% chance of finding a leak, that this is a false alarm. For this thesis, it is assumed that a leak should be found at least once a day. This translation of probability involves quite some stochastic assumptions and is therefore merely an estimation of what can be expected in reality. It also makes use of some results from chapter 7, which might add difficulties in understanding when these results have not been read yet.

In a different MSc-thesis of Van Steen (2020) a leak detection time of 25 minutes was proposed to distinguish a demand deviation from a actual leak. Since the algorithm in this thesis distinguishes time steps of 10 minutes, this implies that a leak should be able to detected for 3 time-steps in a row. The probability that the first of the three steps finds a leak equals the percentage found (pf_i). From there, it is assumed that the probability that a leak can still be found in the sequential step is somewhere in between the correlated scenario, in which case the sequential probability will be expressed by using Pearson's correlation coefficient (r_{t+1}), and the uncorrelated scenario[7], which results again in the same probability (pf_i). Doing so, the chance that a sensor detects a leak for three consecutive time steps ($P_{sens,leak}(t:t+2)$) can be expressed, as shown in equation 4.19. A more detailed explanation of constructing this formula can be found in appendix A.18.

$$P_{sens,leak}(t:t+2) = pf_i \times \frac{r_{flow,t+1} + \frac{pf_i}{100}}{2} \times \frac{r_{flow,t+2} + \frac{pf_i}{100}}{2} \tag{4.19}$$

The formula above assumes that the Boolean 3D-matrix does not change over time. However, this assumption can not hold for the entire day, as results in section 6.3 show a clear difference between the detectability of the system during night and during day (Figure 6.7 and 6.8). Therefore, the ranges at which the detectability was considered as "good" (in the same figures) will be used as the time frame during which the 3D-matrix does not change. For flow sensors this time frame was between 21:40 and 05:20,

[7] The flow in pipes at two sequential time steps can be considered to be correlated, since the flow into the DMA is very correlated at two sequential time steps. This correlation does however decrease for pipes within the DMA, as shifts in demands can cause significant flow changes within the DMA. The correlation of flow in pipes in the DMA was therefore considered to be somewhere in between the correlated and the uncorrelated scenario.

containing 44 different options of three consecutive time steps (t=0,1,2; t=1,2,3; etc.). Now, it is possible to calculate a probability limit (p_{limit}), which would result in a leak that is being found once a day. This probability limit has to satisfy equation 4.20. It was found that the probability limit for detecting leaks once a day with flow sensors was 0.079. Due to the high uncertainties that arose in estimating this probability limit, it is rounded up to 0.1 (10%).

$$p_{limit} \times \frac{r_{flow,t+1} + p_{limit}}{2} \times \frac{r_{flow,t+2} + p_{limit}}{2} \times 44 \geq 1 \tag{4.20}$$

The range with "good" detectability for the monitoring system with pressure sensors was shown to be between 00:00 and 06:30 and between 18:30 and 23:50[8] resulting in 68 different possible time steps (section 6.3). Furthermore, the pressure measurements that are used as input values only occured during 5% of all Fridays (explained later in section 6.1). Therefore, this probability of 5% should be added to equation 4.19, resulting in a probability limit of 55% for pressure sensors (for details, see appendix A.18). This probability limit would prove to be hard to obtain.

So, leaks which have a higher probability of being detected than the probability limit are expected to be detected once a day. For the monitoring system with flow sensors this implies that leaks should be detected in 10% of the experiments in the Monte Carlo simulation, whilst the monitoring system with pressure sensors should be able to detect leaks in 55% of the experiments.

When it is clear which leaks can be detected once a day, the *coverage* of a certain configuration of sensors can be calculated. The coverage is defined as the percentage of pipes in which the sensor configuration can detect leaks once a day. So, if the sensor configuration is able to detect leaks in 50% of all pipes, than the coverage equals 50%. The actual spatial sensor placement of the optimal solution can be found by reading which entry of vector $\mathbf{x}$ equals to 1, as a sensor was placed in the coherent pipe or node.

4.6 Constructing a business model (Q5) and assessing the applicability of the monitoring system for IWS areas around the globe (Q6)

The input for the business model consisted of a small amount literature, but most information was gathered by talking to the engineers and operators of the local utilities. These interviews were conducted through online platforms, as it was not possible to visit the utilities. Different analyses and the resulting business model canvas are shown in chapter 8.3.

To conclude whether the monitoring system is applicable in other IWS areas around the globe, the results and conclusions of all chapters (the results from answering research questions 1 to 5) were combined in an advice for IWS systems worldwide.

[8]These are two separate time frames since leak detection with pressure sensors in this design would only occur on Fridays.

5. Hydraulic modelling of IWS systems (case study Harare)

In this chapter, three DMA's in Harare are analyzed to showcase demand and pressure patterns in IWS systems. One DMA (Ashdown Park) is analyzed in more detail, since a hydraulic model was created for this DMA. The hydraulic model was calibrated with flow and pressure measurements conducted at the DMA entrance.

5.1 Demand and pressure in three DMA's in Harare

Flow and pressure data from 08-06-2019 until 08-06-2020 has been analyzed for three DMA's in Harare. These DMA's are Ashdown Park, Sunningdale and New-Marimba Park. Their geographical locations are shown in Figure 5.1. The flow and pressure were registered every 10 minutes, resulting in daily data-sets of 144 points. Daily sets with less than 144 measurements were marked as inconsistent and removed from the overall data-set. After this "clean-up", a set with 223 days in Ashdown Park, 343 days in Sunningdale and 347 days in New-Marimba Park remained. The results of the analysis will be described below.

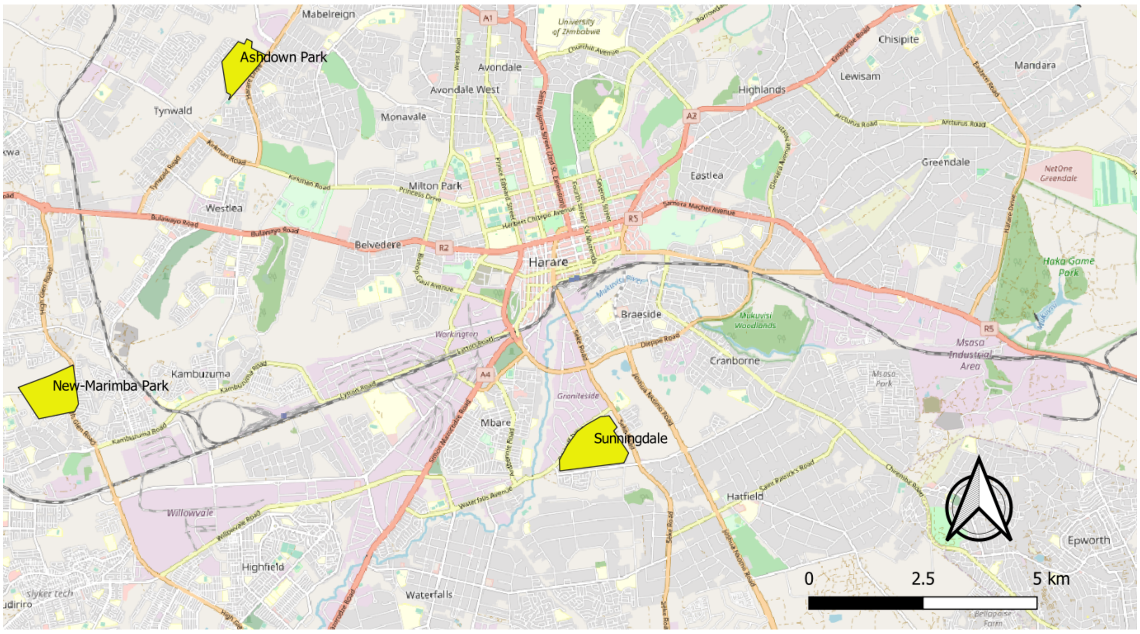

Figure 5.1: Locations of the analyzed DMA's in relation to the city centre of Harare (in the middle of the figure).

Intermittency

In Figure 5.2, 5.3 and 5.4, the average weekly flow patterns can be found. All these DMA's are fed by a single feed, at which the flow is monitored. It can be noted that Ashdown Park receives water from Thursday afternoon until Tuesday afternoon, Sunningdale receives water during the entire week and Marimba Park receives water from Monday until Friday.

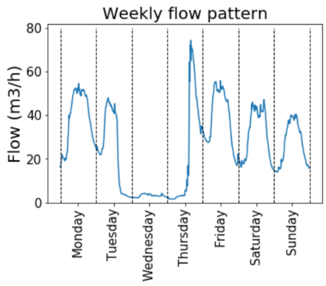

Figure 5.2: Weekly flow pattern Ashdown Park

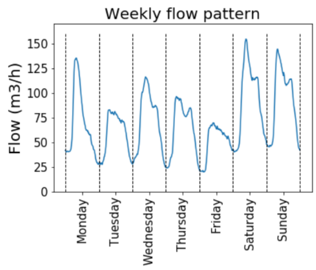

Figure 5.3: Weekly flow pattern Sunningdale

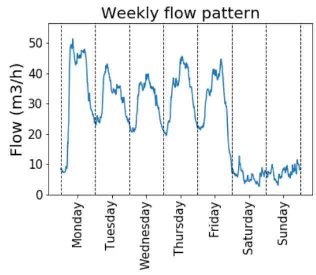

Figure 5.4: Weekly flow pattern New-Marimba Park

Pressure patterns

The pressure patterns are given in Figure 5.5, 5.6 and 5.7. Again, all these pressures are measured at the DMA inlet. Ashdown park is located behind a pressure reducing valve (PRV), which aims to keep pressure in the DMA constant during the week (Kureva and Moors, 2019). The PRV is located at the inlet of the DMA (see Figure 5.11) and monitors the downstream and upstream pressure. The pressure in Sunningdale shows a high value around the weekend and a low value during the other days. New-Marimba Park shows a contrary pattern. However, the layout of the system shows that they are not directly connected to the same source, so no conclusions can be drawn from this contrary pattern (City of Harare, 2019). The pressure in the entrance can not be a measure for the pressure in the rest of the system, since Harare has differences in elevation throughout the city.

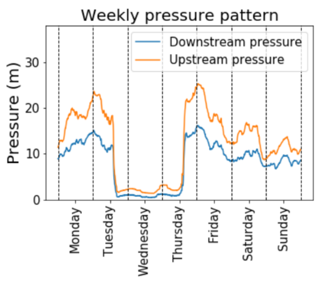

Figure 5.5: Weekly pressure pattern Ashdown Park

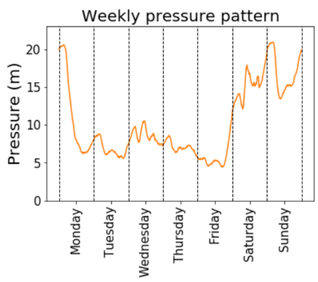

Figure 5.6: Weekly pressure pattern Sunningdale

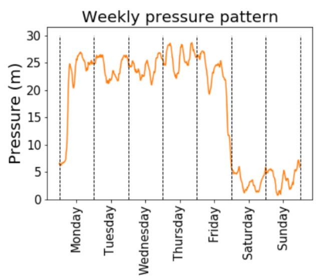

Figure 5.7: Weekly pressure pattern New-Marimba Park

Base demands

To get an estimate of the water consumption behaviour of Harare's inhabitants, the water flow through the meters was translated to a personal water use (in L/p/d). It should be noted that the DMA inflow is used for this, so possible leaks within the DMA are included in this amount.

The available population data from Ashdown Park was quite accurate due to the conducted customer survey (the data from this survey is used in this analysis). For the population data in Sunningdale and New-Marimba Park, a choice had to be made between the number of properties that were in the customer database or the cities registry. For this analysis the customer database is used, since this gives proof that the property is connected to the drinking water system. To estimate the water use per person, a similar household size was used as concluded from the H2H survey (5.8 persons per household), but estimates could range between 4 and 8 persons per household for these DMA's (Moors, 2020).

The base demands that were found were 238 L/p/d, 198 L/p/d and 313 L/p/d for Ashdown Park, Sunningdale and New-Marimba Park respectively (Table 5.1). This shows a way higher personal water use than for instance in the Netherlands, where usually around 130 L/p/d is used as a guideline (TU Delft, 2019). This high per capita water use seems to be in line with the high per capita water use which was found in an earlier modelling report of the Republic of Zimbabwe (2017), which was elaborated upon in the literature review (section 2.2). This high demand can be due to leaks in the system or meter inaccuracies, but also due to a different human behaviour. Washing, for instance, takes place by using many different buckets instead of a water re-using washing machine which is likely to result in a higher personal water use.

DMA (08/06/2019 – 08/06/2020)	Daily water use (m3/d)	No. properties (customer database)	Avg. Household size	Personal water use (L/p/d)
Ashdown Park	632	458	5.8	238
Sunningdale	1712	1487 (1201 in registry)	*5.8 (4 - 8)*	*198 (144 - 288)*
Marimba Park	594	327 (387 in registry)	*5.8 (4 - 8)*	*313 (227 - 454)*

Table 5.1: Demands per person in DMA's.

Daily demand patterns

The daily demand patterns on days with pressure in the two intermittent DMA's (Ashdown Park and New-Marimba Park) were compared with an average Dutch daily demand pattern. As can be seen in Figure 5.8, an average Dutch daily demand pattern has a peak in the morning when people wake up and a peak at night around dinner. When the daily patterns of the intermittent DMA's were considered, no such strong peaks could be distinguished. The patterns for all the days of the week can be found in appendix A.5. Most of the demand patterns were like the flow pattern on Monday at Ashdown Park (Figure 5.9); quite constant over the day and a drop of demand during the night. The demand pattern which came closest the Dutch pattern was the pattern on Tuesdays in New-Marimba park (Figure 5.10). Still this pattern is more gradual, which could have various reasons.

From the H2H-survey at Ashdown Park, it could be concluded that 10 out of 261 questioned households had their tap constantly open (Kureva and Moors, 2019). This creates a constant demand over the day. Most people use a certain type of storage to fulfill their water need on days without supply. If these people fill their storage during the day, their demand remains constant over the day as well. The demand fluctuation caused by a single customer then cannot be measured, since they tap water directly from their storage tanks instead of from the distribution network. Also cultural differences can play a role in water use behaviour and therefore result in a different demand patterns. For example, if people in Harare are used to washing during the day and use more water for cooking lunch at home than cooking dinner at home, their demand pattern will deviate from the Dutch one. All the aforementioned behaviours could be reasons for the difference in demand pattern in the Netherlands and Harare, but giving proof for certain relationships is outside the scope of this thesis.

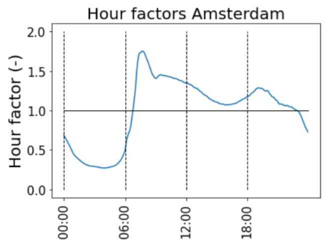

Figure 5.8: Example demand pattern Amsterdam (Waternet, 2020)

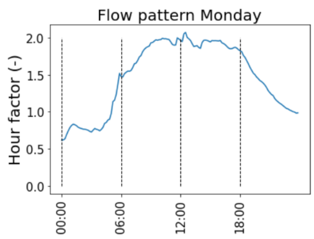

Figure 5.9: Flow pattern Monday Ashdown Park

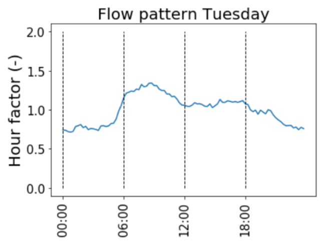

Figure 5.10: Flow pattern Tuesday New-Marimba Park

5.2 Ashdown Park (Harare, Zimbabwe)

In this chapter, some more detailed information about the DMA Ashdown Park is given, as the hydraulic model shall be built for this DMA.

Introduction

Ashdown Park consists of 458 properties and has about 2656 inhabitants, of which 99% are residential (Kureva and Moors, 2019). 261 houses were accessible (61%) for the previously mentioned H2H-survey. The DMA is located in the North-West of Harare. Its elevation ranges from 1472m to 1487m. At the DMA inlet there are sensors that measure the incoming flow and the pressure. The outline of the DMA is given in Figure 5.11.

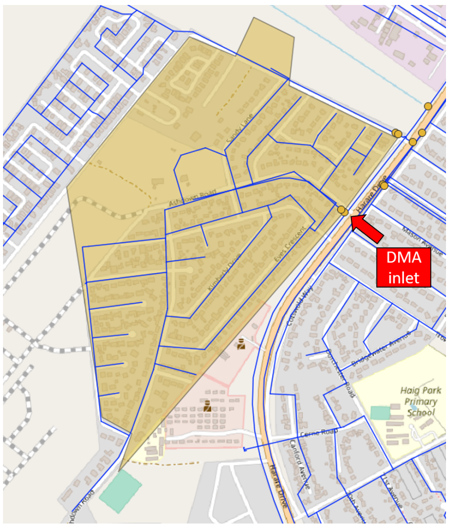

Figure 5.11: DMA Ashdown Park and inlet

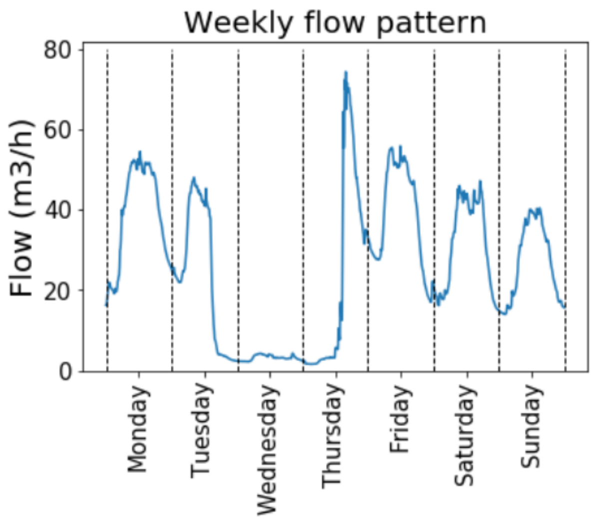

Figure 5.12: Weekly flow pattern Ashdown Park

Water demand

The water demand in the DMA can be described in terms of quantity and in terms of spatial variation.

Quantity

Of the respondents of the H2H-survey, 27% had access to water supply alternatives. 11% because they had a bore hole on the property, 14% because they had a well on the property and 2% through delivery. As can be seen in Figure 5.12, the DMA usually receives water on 5 days per week. During Tuesday evenings, Wednesdays and Thursday mornings, there is no water. As mentioned before, the calculated water demand from the inflow data is 238 L/person/day. Seven of the 261 houses had a swimming pool that they used, which increases the water demand of these properties.

Spatial variation

The spatial variation of demand has a significant impact on the flows in the network and therefore on the detectability of leaks. In the received EPANET-file from VEI, the nodes were already given a certain demand. This demand had been based on a standard procedure within VEI to calculate demands based upon the number of connections at a node and a standard demand of 20 m3/month per connection (Moors, 2020). The resulting division of demand among the nodes is visualised in Figure 5.13. It shows a strong spatial variability within the DMA, mainly since some parts of the DMA are more densely inhabited than other areas.

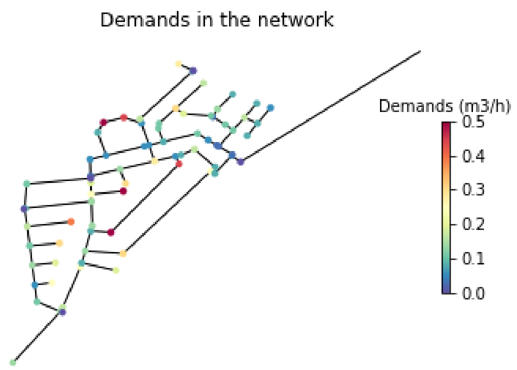

Figure 5.13: Demands in Ashdown Park

Storage

The H2H-survey also questioned whether people have the option to store drinking water at home. An overview of the results of this survey can be found in Table 5.2. The storage option of “some buckets” was not further defined. From personal experiences, I estimated that approximately 50L could be stored in “some buckets”. The total storage capacity was found to be 207.170 L. This is about 794 L storage per property. Extrapolating this number to the 458 properties[1], the total storage in the area would be around 364.000 L. This equals the estimated daily water demand of about 1500 people in the area (56% of the total population).

Storage?	No	1.000 L	2.500 L	5.000 L	10.000 L	Some buckets (50L)	Other	Total
Households	101	14	3	22	6	111	6	263
Storage capacity	0 L	14.000 L	7.500 L	110.000 L	60.000 L	5.550 L	10.120 L	207.170 L

Table 5.2: Overview of storage in Ashdown Park, according to the House-to-House survey

5.3 Constructing the hydraulic IWS model

The strategy and software that is used for constructing the hydraulic model is explained in section 4.3. Here, the accuracy of the hydraulic model is shown and an estimation of the nominal pressure in Ashdown Park is made. The nominal pressure is included as an important parameter in the hydraulic model.

Calibrating the hydraulic model with flow and pressure measurements

The hydraulic model simulates the hydraulic behaviour of water flows resulting from the nodal demands in the network, added pressure from the head of a reservoir and the physical parameters of the system (Figure 5.14). The demand and reservoir head in the hydraulic model were adjusted, to make sure that the calculated flow and pressure at the DMA inlet approached the values of the measurements at this location. The calibration method that was used is shortly explained in section 4.3 and explained in more detail in appendix A.4.

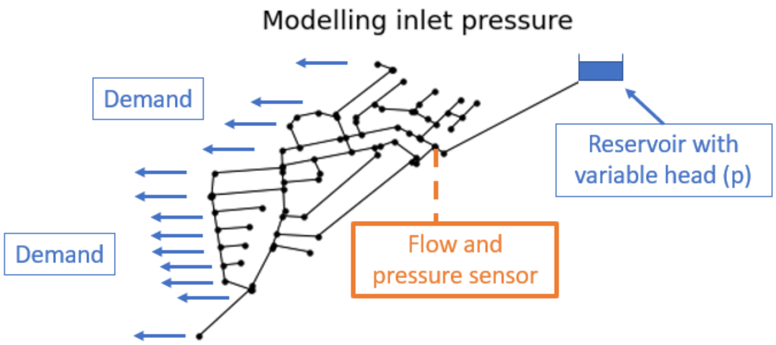

Figure 5.14: The model can use input values (in blue) to simulate the flow and pressure at the location of the sensor (in orange).

The flow and pressure values that were calculated by the model at the DMA inlet and the performed measurements can be visualized before the calibration (Figure 5.15) and after the calibration (Figure 5.16). It can be seen that the calibration method strongly increases the accuracy of the model.

[1]The H2H-survey only covered 61% (263 of the 458 properties) of the number of houses in the DMA.

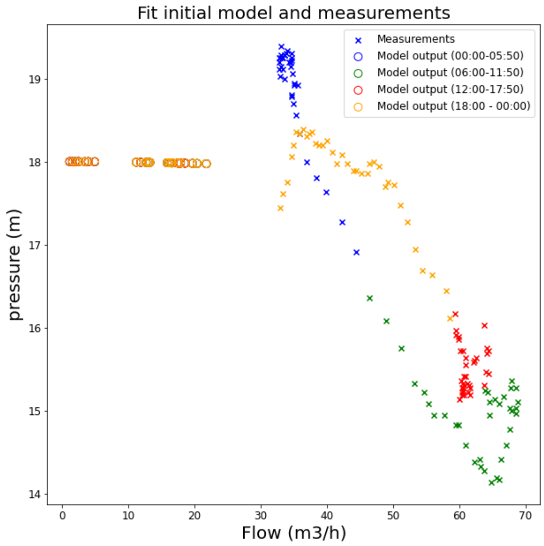

Figure 5.15: Output of the model (circles) compared to the measurements (x's) before calibration.

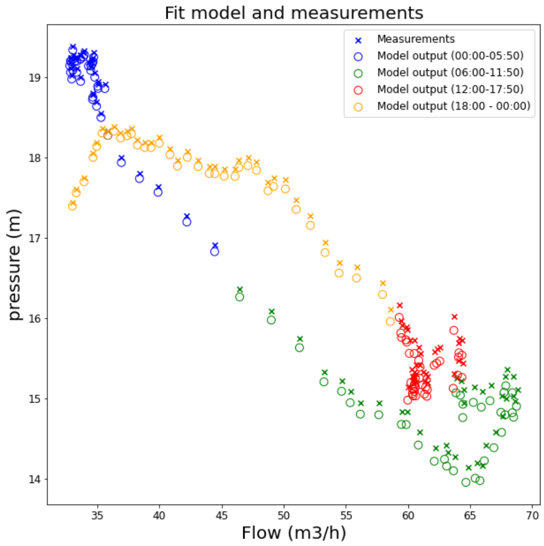

Figure 5.16: Output of the model (circles) compared to the measurements (x's) after calibration.

This accuracy can be expressed in numbers and is influenced by the choice of δ_D and δ_p in the calibration method. The nominal pressure influences the accuracy as well, since a lower nominal pressure value will increase the non-linear relationship between pressures, demands and flows. In an iterative process, different values for δ_D and δ_p were used at different nominal pressures to determine which values for δ_D and δ_p would result in the lowest NRMSE's (for the definition of NRMSE, see section 4.3). The entire process can be found in appendix A.8 and its results can be seen in Table 5.3. All individual NRMSE values are below 0.05, which indicates that the proposed method can perform quite well in all scenarios given that the δ_D and δ_p values are adapted to the right values.

Nominal pressure (m)	δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
20	18	1	0.048	0.012	0.060
12	0.1	1	0.012	0.0070	0.019
1	13	1	0.0051	0.0071	0.012

Table 5.3: Overview of best results from sensitivity analysis.

Estimating the nominal pressure in Ashdown Park.

Since the PDO-modelling is applied for the hydraulic model, a certain nominal pressure will need to be estimated for Ashdown Park. This nominal pressure (explained in more detail in section 2.2) determines whether the demand of the people can be fully met, given the pressure that occurs in the system. When applying a nominal pressure of 12m in the hydraulic model, it was found that the total demand was almost the same as the inflow measurements on Monday (Figure 5.17). The fact that nearly everyone received their total demand with a nominal pressure of 12m, implies that pressure in the hydraulic model at all nodes with demand was above 12m on Monday. Applying a nominal pressure of 20m resulted in a total demand that was higher than the inflow measurements (Figure 5.18).

This implies that due to the nominal pressure inhabitants received only a part of their total demand. This can only occur if the pressure at the nodes drops below 20m, which occurred. This way, the nominal pressure has a significant influence on the modelled total demand.

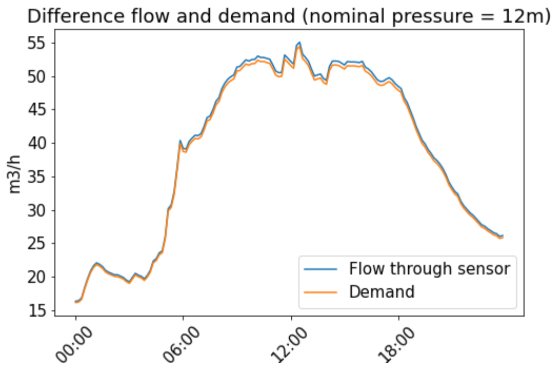

Figure 5.17: The difference between flow and demand at a nominal pressure of 12m on Monday.

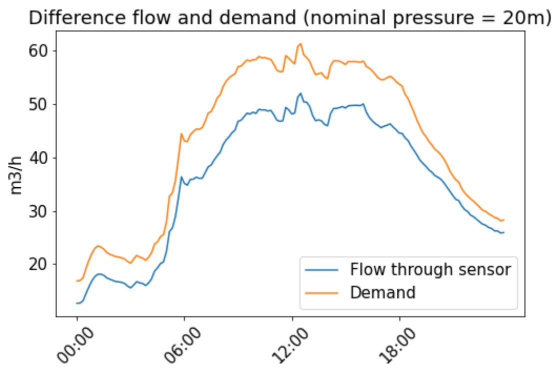

Figure 5.18: The difference between flow and demand at a nominal pressure of 20m on Monday.

The nominal pressure can be estimated by taking into account the energy losses that occur at the house connection and the elevation that the drinking water has to overcome. If the pressure in the system is insufficient to overcome these losses, the demand of the people can not be satisfied. Calculations to estimate these losses in Ashdown Park, described in appendix A.9, showed a potential head losses of 6 to 8m. Since 12m is well above the calculated head losses at the household level (6-8m), it can be concluded that there is sufficient pressure in the DMA. Therefore, it can be stated that for Ashdown Park the inflow into the DMA is likely to equal the total demand of the DMA (including leaks).

5.4 Output of the hydraulic IWS model

This chapter shows the output of the hydraulic model, describing which pressures and flow are expected to occur within the DMA.

Pressures

The main pressure losses in the system are due to changes in elevation. The daily changes in pressure (+/- 15m difference within one week) are most likely caused by the pump operation scheme, which is explained in detail in appendix A.6. The critical point, the point with the lowest pressure in the system, is shown in Figure 5.19. The pressure difference between the feed and the critical point is around two meter at night and three meter during the day, according to the results of the simulation in Figure 5.20. The elevation difference between the feed (1484m) and the critical point (1486m) is also two meter, which indicates that only minor head losses at night. Figure 5.21 shows that the highest elevations (above the feed's elevation) all occur around the same area as the critical point.

Flows

Figure 5.22 shows the modelled flows that occur on average on Monday at 00:00. The flow output can be used to identify segments where large frictional losses occur, or where high residence times occur.

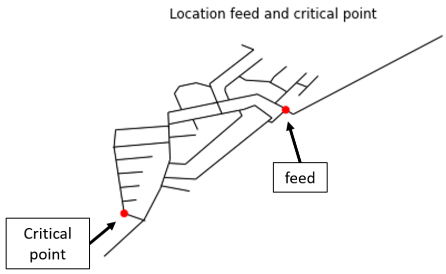

Figure 5.19: Location of the feed and the critical point N72

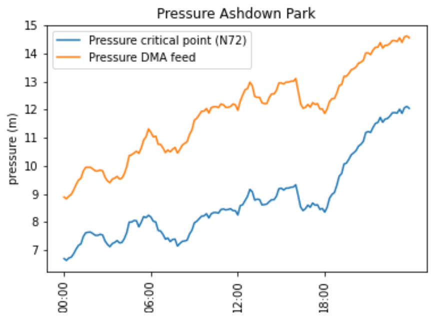

Figure 5.20: Pressure on Monday at 00:00 at the feed and critical point

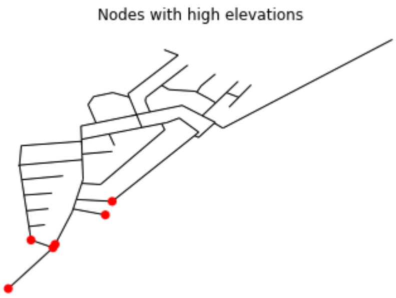

Figure 5.21: The area of the nodes with high elevation (above 1484m).

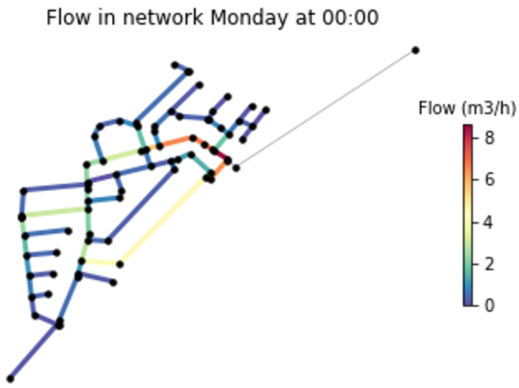

Figure 5.22: The flows in the network on Monday at 00:00

5.5 Conclusions chapter 5

This chapter answers the following research question:

Q2: How can intermittency be included in a hydraulic model?

A number of conclusions can be drawn from constructing the hydraulic model. First of all, the daily demand patterns in IWS systems show a different pattern than is continuously supplied systems. The IWS patterns show less strong peaks. This can be due to a constant water demand for filling storage, leaks in the system or different consumer behaviour.

This different daily demand pattern, subtracted from analyzing flow data at the DMA inlet, can be used to calibrate the hydraulic model. The complicated relations between demand and pressure in the model can be approached by using assuming linear relationships between the demands, inlet pressure as input variables and the pressure (at a node) and flow (at a pipe) as output. When the expected pressures are below the estimated nominal pressure, the accuracy of the method decreases as the relationship between demands and pressure becomes more non-linear.

Lastly, the inlet pressure fluctuates considerably over the week. As shown in the modelling study, the inlet pressure is almost entirely dependent on the pressure in the main drinking water network to which the DMA is connected and therefore by the supply scheme of the utility. This should be well taken into account in the hydraulic model. Depending on the nominal pressure in the area, this inlet pressure can determine whether demands are entirely met or not.

Suggestion

If the aim would be to model an entire city with IWS, the main factors that determine the flow would be the pressures in the main network and the demands in the DMA. Therefore, a suitable way of modelling these systems would be to represent the DMAs as nodes connected to a main network (see Figure 5.23). By connecting and disconnecting these nodes from the main network at given days, the pressure in the main network changes. This behaviour could be modelled by estimating the right demands for the DMAs and adding demand patterns and a distribution scheme. Unfortunately, a city with IWS does usually not have the necessary equipment to construct or monitor the inflow and pressure at all connected DMAs or the geographical information (GIS) of the system is missing, making this way of modelling not possible *yet* for Nairobi and Zimbabwe.

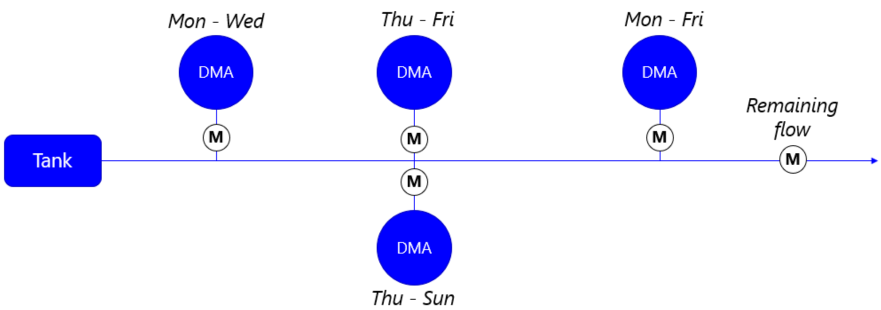

Figure 5.23: Ideal way of modelling IWS systems.

6. Boolean leak detection

The method that is used to detect leaks is the "Dynamical Bandwidth Monitor (DBM)", whose concept was explained in section 4.4. This chapter highlights differences between designing a DBM system with flow sensors and pressure sensors. Furthermore it substantiates which method for modelling demand realizations can best be used to construct alarm values for the DBM. Finally, it identifies the sensitivity of the leak detection model to certain important parameters.

6.1 Differences between the design with flow sensors and pressure sensors.

There are a number of differences between the design of the monitoring system with flow and pressure sensors. These differences arose from several analyses, whose conclusions and emerging assumptions were used to make a distinction between the two types of design. These analyses are shown below, as well as the differences that arose.

Filtering days with continuous supply

The DBM compares ranges of expected values to real-time measurements, after which an anomaly in measured value is marked as a potential leak. To simulate reliable ranges of expected values, days with a regular flow and supply pattern have to be used. Days with irregular patterns (or no flow and pressure at all) can not be used, as anomalies would occur without new leaks emerging in the system, resulting in false alarms. So, the DBM can only be used on days which have a continuous supply for the entire day. Removing days with intermittent supply during the day resulted in a significant reduction of days whose measurements could be used to construct the monitoring systems (Table 6.1). Irregularity of supply by the water utility so decreases the functionality of the monitoring system.

Day of the week	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Days with a complete dataset (144 measurements)	31	32	32	32	32	32	32
Days with continuous supply	10	X	X	X	14	3	5
% of days with continuous supply	**33%**	**0%**	**0%**	**0%**	**44%**	**9%**	**16%**

Table 6.1: Daily data-sets that were used to construct the DBM's.

This low number of days that are available to construct the DBM let to the following design goal: *"Construct alarm values which can be used for all days with supply, instead of creating different alarm values for each weekday. Weekly alarm values are more easy to operate and can be used more frequently.*[1]

[1]More detailed explanation: Creating different alarm values for each weekday would result in more

Relation between hydraulics and measurements at the DMA inlet

To distinguish between designing a monitoring system with flow sensors and with pressure sensors, it should be taken into account how the flows in pipes and pressure in nodes depend on the flow and pressure at the DMA inlet. These relationships are listed below and summarized in Table 6.2.

- The flow in pipes is assumed to be **dependent** on the inflow into the DMA, since the magnitude of the incoming flow determines the mass balance of water entering the system and therefore the flow rate that will go through the pipes.
- The flow in pipes is assumed to be **independent** on the pressure at the DMA inlet, as the pressure in Ashdown Park is estimated to be sufficient to allow people to fulfill their water demand (section 5.3). Therefore, it is assumed that a pressure deviation will not likely result in a change of flows in the system.[2]
- The pressure in nodes is assumed to be **dependent** on the inflow into the DMA, as this influences the flows in pipes, which influences the head losses over pipes and the pressures in nodes.
- The pressure in nodes is assumed to be **dependent** on the pressure at the DMA inlet as this determines the total amount of pressure that is added to the supply system and therefore directly influences the pressures in nodes.

	DMA inflow	Inlet pressure
Flows in pipes	Dependent	Independent
Pressure in nodes	Dependent	Dependent

Table 6.2: Relation between hydraulics and DMA inflow and inlet pressure.

Similarity of flow and pressure patterns on different weekdays

A design goal for the monitoring system is to construct alarm values that can be used for each weekday. Therefore, it was explored whether the flow and pressure patterns that were found after analyzing the measurements at the DMA inlet at different weekdays with continuous supply (the weekdays shown in Table 6.1) showed enough similarity to construct a weekly pattern.

Comparing the differences between weekly flow averages and flow averages of each individual weekday, resulted in a NRMSE of 0.0498, indicating that a weekly flow pattern **is able to** represent the flow which can be expected on every individual week-day. The

accurate ranges of expected flow and pressure on days like Monday and Friday. They could however not be used on Saturdays and Sundays, as 3 and 5 days were assumed to be too little days to construct a reliable range of hydraulic values. Using weekly alarm values (for days with supply) eliminates the need for an operator to switch alarm values on different weekdays, increasing the ease of operation. Furthermore, it allows the system to be used on Saturdays and Sundays as well. These advantages were chosen to outweigh the disadvantage of having a less reliable range of expected values on Monday and Friday.

[2]The pressure would influence the flow rate going through leaks in the system (equation 2.1, section 2.3). However, this is not taken into account at this point.

same analysis with the weekly pressure pattern, resulted in a NRMSE of 0.101, which was regarded to be too high[3]. Therefore, a weekly pressure pattern **is not able to** represent the pressure of individual weekdays with continuous supply. This full analysis can be found in appendix A.11 and the flow and pressure patterns of all individual weekdays can be found in Figure 6.3 and 6.4. These figures show that the flows at individual weekdays are more alike than the pressures at individual weekdays.

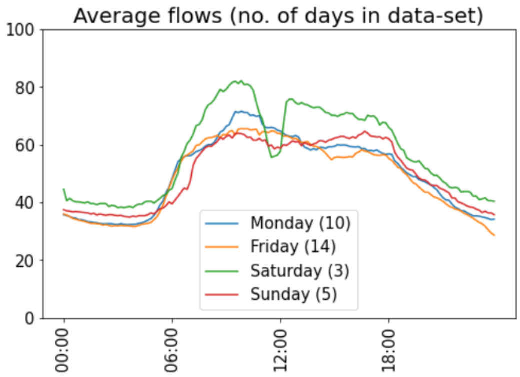

Figure 6.1: Average flows of days for with continuous pressure.

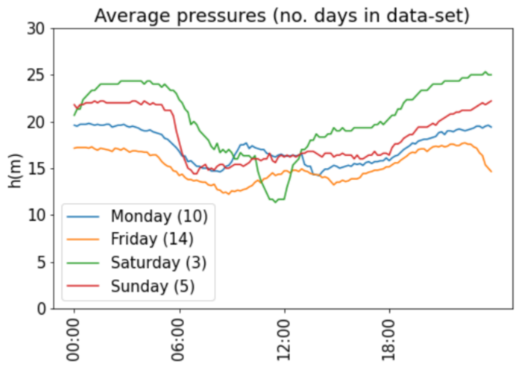

Figure 6.2: Average pressures of days with continuous pressure.

Resulting differences between the design with flow sensors and pressure sensors.

A relatively obvious difference between a monitoring system with flow sensors and pressure sensors is that flow sensors should be located at pipes and pressure sensors are usually located at nodes.[4] The number of suitable locations for flow sensors therefore coincides with the number of pipes in the network, whereas the number of suitable locations for pressure sensors coincides with the number of nodes. Furthermore, a monitoring system with pressure sensors only uses a lower alarm value, since each leak takes away pressure from the system and thereby can only reduce the pressure in nodes. Flow monitoring systems use both a lower and upper alarm value, as flows in pipes can increase or decrease when a leak emerges.

As previously explained, the hydraulic model is calibrated with the flow and pressure measurements at the DMA inlet. There are measurements from multiple days with continuous supply, so it is possible to calibrate the hydraulic model to different circumstances during these days. The arguments of the previous three chapters combined, result in a distinction in circumstances during which the monitoring system with flow sensors should be designed, compared to the monitoring system with pressure sensors. This distinction will be explained below and is summarized in Table 6.3.

- **Flow monitoring system:** *It is preferred to construct alarm values which can be used for all days with supply (goal).* A weekly pressure pattern is not able to represent pressure accurately during all days with continuous supply. However, the

[3] the threshold for an acceptable NRMSE was set at 0.05

[4] This is mainly in hydraulic modelling. In reality one could drill a hole in a pipe and install a pressure logger in that hole, which could be hydraulically modelled by adding a new node to the system.

flow in pipes is assumed to be independent from the inlet pressure (Table 6.2), so it can be assumed that an inaccurate inlet pressure has an insignificant influence on the ranges of flows that will be modelled.[5] Therefore, still weekly average values of the pressure measurements can be used. A weekly flow pattern is able to represent flows accurately during all days with continuous supply and can therefore be used as well. The pressure measurements and flow measurements need to be from the same days, so flow alarm values can be constructed using the average measured flow and pressure of all days with continuous supply (as preferred).

- **Pressure monitoring system:** *It is preferred to construct alarm values which can be used for all days with supply (goal).* However, a weekly pressure pattern is not able to represent pressure accurately during all days with continuous supply. Since the pressure in nodes is dependent on this inlet pressure (Table 6.2), pressure measurements of a single weekday should be used. For this design it was chosen to use pressure measurements from Friday, as Fridays contained most days to construct a reliable range of values (table 6.1). It should be taken account as well that only a lower alarm value will be constructed for the pressure monitoring system. As the inlet pressure determines the amount of pressure that is added to the system, the pressure alarm values should be constructed in the circumstances when a low inlet pressure occurs.[6] Therefore was chosen to use the 5-percentile pressure values from all Fridays as pressure measurements. The pressure measurements and flow measurements need to be from the same days, so the flow measurements should be from Friday as well. The average flow values for Friday can be used, as the demand allocation method will include possible flow deviations in the model (shown in section 6.2).

	Flow measurements	Pressure measurements
Flow monitoring system	**Weekly average** (Dependent)	**Weekly average** (Independent)
Pressure monitoring system	**Friday average** (Dependent)	**5-percentile Friday** (Dependent)

Table 6.3: Relation between hydraulics and DMA inflow and inlet pressure.

Especially the difference in pressure measurements at the DMA inlet that are used to calibrate the model has a significant impact on the performance of both types of monitoring systems. These performances are described in later sections. The different pressures that are used can be seen in Figure 6.3.

[5] If this relationship was assumed to be dependent, different flow alarm values should be constructed for different weekdays.

[6] If the alarm values would be constructed whilst modelling an average inlet pressure, the monitoring system alarm value would be exceeded, and leaks would be "detected", every time that the inlet pressure drops below the average. A lot of false alarms will occur, since the alarm values is exceeded because the inlet pressure at the DMA is low and not because a leak has emerged.

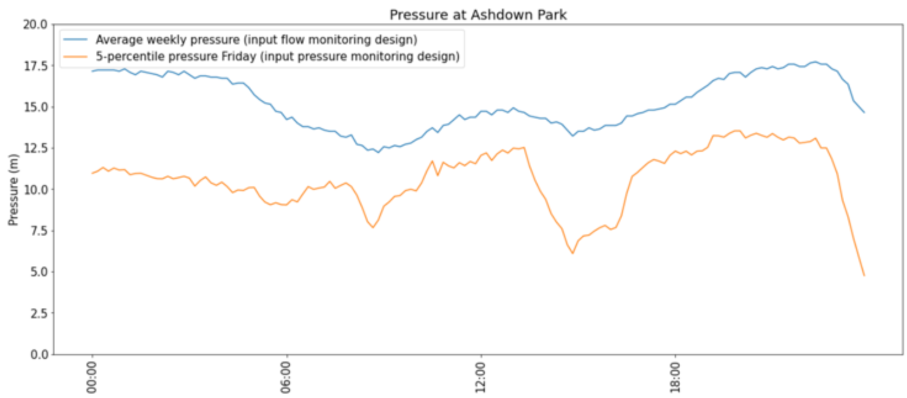

Figure 6.3: The inlet pressures that were used in the model to construct artificial DBM's with flow sensors (in blue) and artificial DBM's with pressure sensors (in orange).

6.2 Method for constructing alarm values for the DBM

Two methodologies were proposed in section 4.4 to model random demand realizations, which could be used in a Monte Carlo-simulation during which the most extreme flows and pressures are saved as alarm values. The different methods will be referred to as "method 1" (using random draws from a scaled normal distribution) and "method 2" (using weighted random choice with a tap capacity parameter). This section substantiates which method shall be used in the leak detection model.

Alarm values from different methods for modelling demand realizations

The result of calculating the alarm values by using flow sensors and optimal input values[7] for both methods can be seen in Figure 6.4. It can be seen that method 2 (the tap capacity method) yields a wider range of flows than method 1 (drawing randomly from the scaled normal distribution).

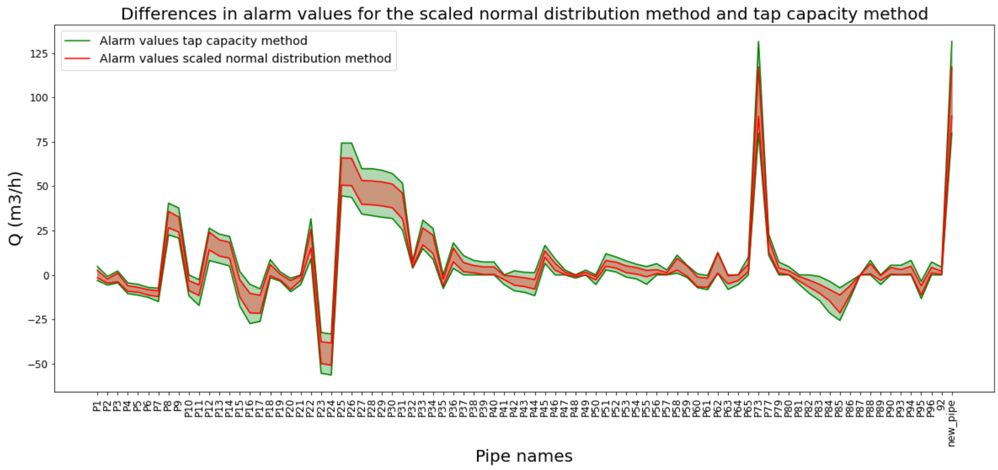

Figure 6.4: Differences in resulting alarm values for a monitoring design with flow sensors between method 1 (using the scaled normal distribution) and method 2 (the tap capacity method).

[7]These optimal input values imply that the alarm values are constructed by running 1000 different demand realizations at 01:20. These values will be elaborated upon in section 6.3

This result was expected, since the tap capacity method does not model demand everywhere, resulting in more locally concentrated demands and in a wider range of flows. This wider range also results in lower pressure alarm values for method 2, since higher flows in pipes occur and more pressure is lost due to frictional losses.

Spread in inflow resulting from different methods for modelling demand realizations

Since the analyzed data-set covers a year of flow and pressure measurements, a data spread at a specific weekday and a specific time can be constructed. The boundaries of this data-spread represent the boundaries of the expected flow and pressure values at the DMA entrance. The 5-percentile and 95-percentile values of weekly flow values at the DMA inlet that can be expected to occur are shown in blue in Figure 6.6. By constructing different demand realizations with random choices, the random scenario can deviate from the average scenario. The modelled inflow into the DMA, which is a result from modelling the different demand realizations at a single time-step according to method 1, is shown in Figure 6.5 with box plots[8].

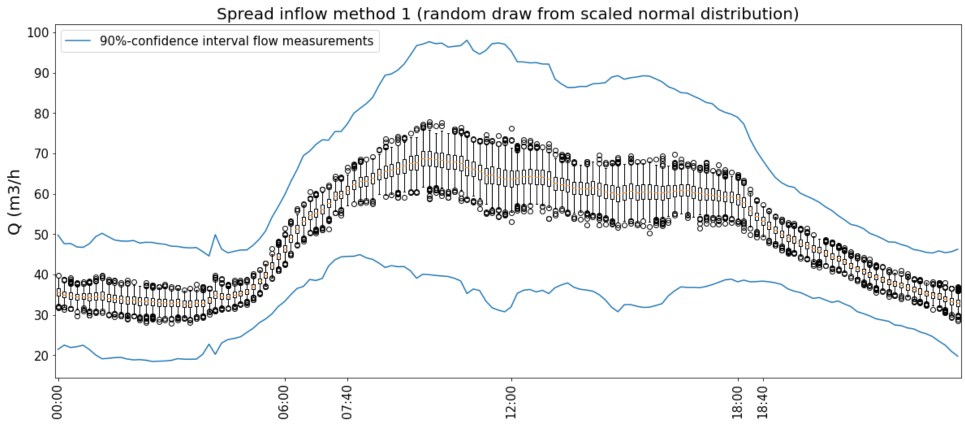

Figure 6.5: The spread of method 1.

Comparing the box plots with the spread of the measurements shows us that the spread in inflow that was calculated by the model was significantly smaller than the spread in the measurements. So, method 1 seems not include the expected range of inflow scenarios. Making the same graph for method 2 resulted in Figure 6.6. This shows us that, especially during night-time (18:40 - 05:20), the modelled spread of inflow is comparable to the spread in measurement data. This comparison does not hold at daytime (05:20 - 18:40), since the measured spread is larger then.

[8]These box plots show the the average value (small orange line in the middle) of all experiments, an interquartile range (the small rectangle shows the range of values between the 25th and the 75th quartile), a minimum and maximum (point up to which the continuous stripe from the interquartile range reaches) and outliers (small circles at the edges of the box plots).

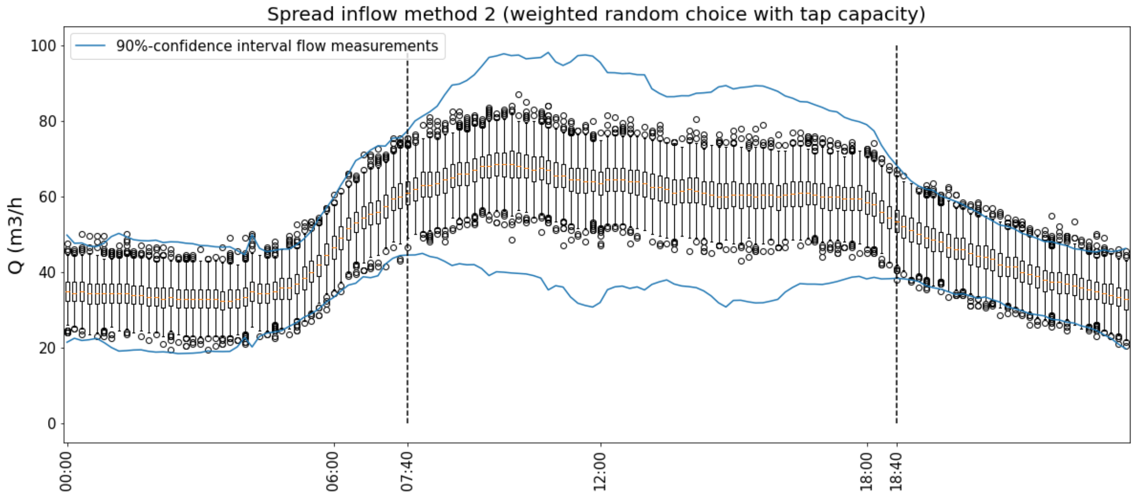

Figure 6.6: The spread in random weighted choice.

Since weekly flow values are used in the above examples, Figure 6.6 and 6.5 show scenarios that occur when designing a monitoring system with flow sensors. Appendix A.13 shows a similar graph, comparing the range of modelled inflows with method 2 to the measurements that are used to design the monitoring system with pressure sensors (measurements on Friday, as explained in section 6.1). This graph also shows that method 2 is able to produce a realistic spread during night-time on Fridays.

Final method for constructing alarm values for the DBM.

It was shown that method 2 is better able to model the range of inflows that can be expected at the DMA inlet than method 1. Especially during night-times, method 2 seems to produce an accurate spread of flows at the inlet. Furthermore, method 2 models a wider range of flows in pipes within the DMA. This results in higher (in absolute terms) flow alarm values and lower pressure alarm values, which are less likely to be exceeded. Since method 2 produces an more accurate spread of inflow at the DMA inlet and produces alarm values which are less likely to be exceeded[9], method 2 is chosen to model the demand realizations and construct the alarm values.

At this point it is clear how the demand realizations ought to be modelled in order to construct alarm values for the artificial DBMs. An overview of all the steps and formulae that are needed to construct the flow alarm values is shown in Algorithm 1, the same overview for pressure alarm values in shown in Algorithm 2.

[9]This "strict" alarm values increase the safety of the design, since they are less likely to be exceeded.

Algorithm 1 Method that is used to construct flow alarm values.

Steps	**Formulae**
Step 1: *Create a model with the demand at nodes* (**D_{init}**) *and the reservoir pressure* ($p_{init,res}$) *as input and gives the flow* ($Q_{init,sens}$) *and pressure* ($h_{init,sens}$) *at the location of the DMA-inlet as output.*	$model_{init,sens}\begin{pmatrix}\mathbf{D_{init}} \\ p_{init,res}\end{pmatrix} = \begin{pmatrix}h_{init,sens} \\ Q_{init,sens}\end{pmatrix}$
Step 2: *Adjust the demand and pressure with a factor* ($D_{factor,t}$ *and* $p_{factor,t}$)*, so that the model output approaches the average weekly flow and pressure from the historical measurements* ($Q_{avg\_week,t}$ *and* $h_{avg\_week,t}$)*.*	$model_{sens,t}\begin{pmatrix}\mathbf{D_{init}} \times D_{factor,t} \\ p_{init,res} \times p_{factor,t}\end{pmatrix} \approx \begin{pmatrix}h_{avg\_week,t} \\ Q_{avg\_week,t}\end{pmatrix}$
Step 3: *Determine the tap capacity (TC) in the DMA and calculate the number of open and closed taps* ($n_{taps\_open,t}$ *and* $n_{taps\_closed,t}$)*.*	$n_{taps} = no.households$ $\quad TC = 0.3\, m^3/h$ $n_{taps\_open,t} = \frac{\sum_{i=1}^{n} D_{i,t} * D_{factor,t}}{TC}$ $n_{taps\_closed,t} = n_{taps} - n_{taps\_open,t}$
Step 4: *Construct flow alarm values for a single time step (t) with a Monte Carlo simulation.*	**Monte Carlo simulation**
Step 4.1: *Assign demand in node i* (D_i) *with random weighted choice and model the pressure in the reservoir* ($p_{reservoir,t}$)*.*	D_i: $\text{Random choice} = \{0, TC\}$ $\text{Weights} = \left\{\frac{n_{taps\_closed,t}}{n_{taps}}, \frac{n_{taps\_open,t}}{n_{taps}}\right\}$ $p_{reservoir,t} = p_{init,res} * p_{factor,t}$
Step 4.2: *Run the simulation; calculate the flow in every pipe*	*Pipe* | P1 | P2 | … | Pj *Flow* | Q_1 | Q_2 | … | Q_j
Step 4.3: *Save extreme flows in pipe j* (Q_j) *as lower or upper alarm value* (Q_{j,al_low} *or* Q_{j,al_high})*.*	If $Q_j < Q_{j,al\_low}$: $Q_{j,al\_low} = Q_j$ If $Q_j > Q_{j,al\_high}$: $Q_{j,al\_high} = Q_j$ *Pipe* | P1 | P2 | … | Pj *Lower alarm value* | Q_{1,al_low} | Q_{2,al_low} | … | Q_{j,al_low} *Pipe* | P1 | P2 | … | Pj *Higher alarm value* | Q_{1,al_high} | Q_{2,al_high} | … | Q_{j,al_high}
Step 4.4: *Run Monte Carlo Simulation by applying 1000 different demand realizations*	Run 1000 experiments

Algorithm 2 Method that is used to construct pressure alarm values.

Steps	**Formulae**
Step 1: *Create a model with the demand at nodes ($\mathbf{D}_{init}$) and the reservoir pressure ($p_{init,\ res}$) as input and gives the flow ($Q_{init,sens}$) and pressure ($h_{init,sens}$) at the location of the DMA-inlet as output.*	$model_{init,sens}\begin{pmatrix}\mathbf{D}_{\mathbf{init}}\\ p_{init,res}\end{pmatrix} = \begin{pmatrix}h_{init,sens}\\ Q_{init,sens}\end{pmatrix}$
Step 2: *Adjust the demand and pressure with a factor ($D_{factor,t}$ and $p_{factor,t}$), so that the model output approaches the average flow and 5-percentile pressures on Friday from the historical measurements ($Q_{avg\_fri,t}$ and $h_{5\_perc\_fri,t}$).*	$model_{sens,t}\begin{pmatrix}\mathbf{D}_{\mathbf{init}} \times D_{factor,t}\\ p_{init,res} \times p_{factor,t}\end{pmatrix} \approx \begin{pmatrix}Q_{avg\_fri,t}\\ h_{5\_perc\_fri,t}\end{pmatrix}$
Step 3: *Determine the tap capacity (TC) in the DMA and calculate the number of open and closed taps ($n_{taps\_open,t}$ and $n_{taps\_closed,t}$).*	$n_{taps} = no.households \qquad TC = 0.3\ {m^3}/{h}$ $n_{taps\_open,t} = \dfrac{\sum_{i=1}^{n} D_{i,t} * D_{factor,t}}{TC}$ $n_{taps\_closed,t} = n_{taps} - n_{taps\_open,t}$
Step 4: *Construct pressure alarm values for a single time step (t) with a Monte Carlo simulation.*	**Monte Carlo simulation**
Step 4.1: *Assign demand in node i (D_i) with random weighted choice and model the pressure in the reservoir ($p_{reservoir,t}$).*	D_i: Random choice $= \{0, TC\}$ Weights $= \left\{\dfrac{n_{taps\_closed,t}}{n_{taps}}, \dfrac{n_{taps\_open,t}}{n_{taps}}\right\}$ $p_{reservoir,t} = p_{init,res} * p_{factor,t}$
Step 4.2: *Run the simulation; calculate the pressure in every node*	*Node* | N1 | N2 | … | Nj *Pressure* | h_1 | h_2 | … | h_j
Step 4.3: *Save extreme pressures in node j (h_j) as lower alarm value (h_{j,al_low}).*	If $h_j < h_{j,al\_low}$: $h_{j,al\_low} = h_j$ *Node* | N1 | N2 | … | Nj *Lower alarm value* | h_{1,al_low} | h_{2,al_low} | … | h_{j,al_low}
Step 4.4: *Run Monte Carlo Simulation by applying 1000 different demand realizations*	Run 1000 experiments

6.3 Sensitivity of the proposed model to construct the Boolean matrices

After having constructed the alarm values for the DBM, Boolean matrices can be constructed that store information whether a leak can be detected by a certain sensor. The ability of the hydraulic model to create this Boolean matrix accurately is determined by several parameters in the model and can be indicated by using KPIs (section 4.4). The type of KPI which can best be used differs per parameter. Calculating the effect of changing a parameter on the KPI gives an indication of the sensitivity of the hydraulic model.

Sensitivity analysis 1: Number of experiments

A single experiment from the Boolean 3D-matrix contains 84 scenarios (since there are 84 pipes in the network), each scenario simulating a leak in a different pipe. So, if there are k experiments in the 3D-matrix, k x 84 scenarios need to be calculated. Therefore, the computational time increases fast when the 3D-matrix grows with more experiments. The computation of the alarm values also uses a certain number of experiments, assigning a new maximum or minimum to the stored alarm values (see Algorithm 1 and 2). In this case, each experiment contains only one scenario since no leaks need to be added to the model, which results in less computational time for a large number of experiments. Different combinations of the number of experiments for constructing the alarm values on one side and constructing the 3D-matrix on the other side were compared by computing a percentage of false alarms (KPI, section 4.4).

The number of experiments that are used to construct alarm values and to construct the 3D-matrix, can be seen as a trade-off between computational time and a certain % of false alarms which is allowed in the model. For example, a number of 10 experiments in the 3D-matrix uses a computational time around 2 minutes, whereas it takes 1 second to compute 10 experiments for the constructing of alarm values. The resulting performance of the monitoring design with flow meters during different times of the day can be seen in Table 6.4.

Percentage of false alarms with a different numbers of experiments for constructing **flow** alarm values and the Boolean 3D-matrix

Calculated by simulating leaks of 0.1mm and calculating the percentage of 1's in the 3D-matrix

Number of experiments in 3D-matrix

Time	06:00			12:00			22:00		
Number of experiments	10	100	1000	10	100	1000	10	100	1000
10	20%	2.3%	0.24%	17.4%	2.1%	0.24%	13.0%	2.0%	0.12%
100	13.9%	2.1%	0.12%	15.9%	1.6%	0.18%	21.6%	1.9%	0.11%

Number of experiments for constructing alarm values

Table 6.4: Percentage of false alarms for the design with flow meters and a different number of iterations.

It was chosen to construct the alarm values using 1000 experiments and to construct the 3D-matrix using 100 experiments. These 100 samples should present a reasonable image of the flows that can occur in the DMA (Blokker, 2020). This resulted in a computational time of around 25 minutes, which was feasible to further investigate the influence of other parameters on the model. The same analysis for the monitoring system with pressure sensors (appendix A.14), showed that 1000 experiments for constructing alarm values and 100 experiments for constructing the 3D-matrix was also a good choice for this system.

Sensitivity analysis 2: Time

The Boolean 3D-matrix is constructed for a specific time step, since demand realizations and modelled inlet pressure are time dependent. Given the fact that the inlet pressure and incoming flow are measuring every 10 minutes, the number of open taps (which influences the demand realizations) and the inlet pressure change every 10 minutes. Therefore, the above described Boolean 3D-matrix changes every 10 minutes. The performance of the monitoring system can be expressed by calculating the detectability (KPI, section 4.4) at every time step.

Plotting the detectability for the monitoring network with flow meters at different times during the day[10], shows us that the leaks are best detectable between 21:40 and 05:20. This time range is chosen, since the detectability during these times is clearly higher than the detectability during the day (Figure 6.7). From now on, in order to calculate the optimal sensor allocation, all scenarios shall be run on 01:20. This time has a detectability that can be seen as average within the period with good detectability and can well be used for estimating how many leaks can be detected within this period.

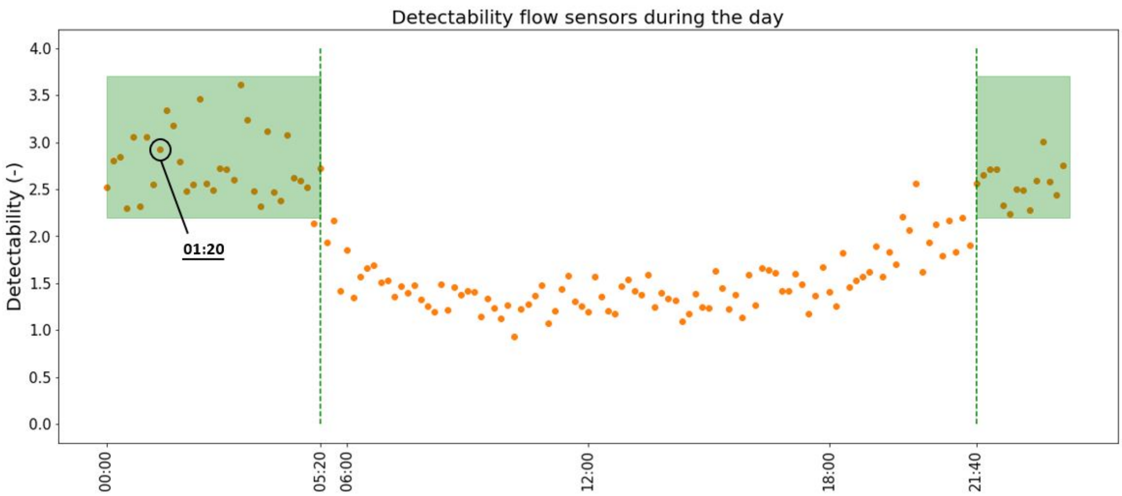

Figure 6.7: Detectability of leaks with flow sensors at different times of the day.

Making the same plot for the detectability of pressure sensors shows a slightly different outcome. As one can see in Figure 6.8, the precision of the detectability is lower than

[10]For this calculation, the 3D-matrix needed to be calculated 144 times, resulting in a calculation time of 64 hours. In this simulation, leaks of 20mm were used, which will later show to be too large. In the final model leaks of 15mm will be used. This would decrease the exact detectability in Figure 6.7 and 6.8, but its shape is likely to remain the same. Since the aim of this chapter is merely to identify which time should be used to run scenario's, which can be done by just using the shape of the graphs, the simulations were not redone.

the precision of the network with flow sensors. The results show a slightly improved detectability during the night, but the distinction between night and day is not as strong as was the case with flow sensors. Therefore, the distinction of the "good detectability" range is made based on different grounds. The previous section showed that the hydraulic model's output at the DMA inlet (using method 2) during the night-time was comparable to the spread of the inflow measurements (Appendix A.13 for the design with pressure sensors). The range during which the model produced accurate results was from 00:00 until 06:30 and from 18:30 until 23:50[11] (also appendix A.13). These ranges were used in Figure 6.8 to distinguish a time during which the model performs best. Calculations for the allocation of pressure sensors shall be performed by using the flow and pressure input data at 03:30 (average value within the range of "good detectability"), albeit that this time preference is less strong than was the case for designing with flow sensors.

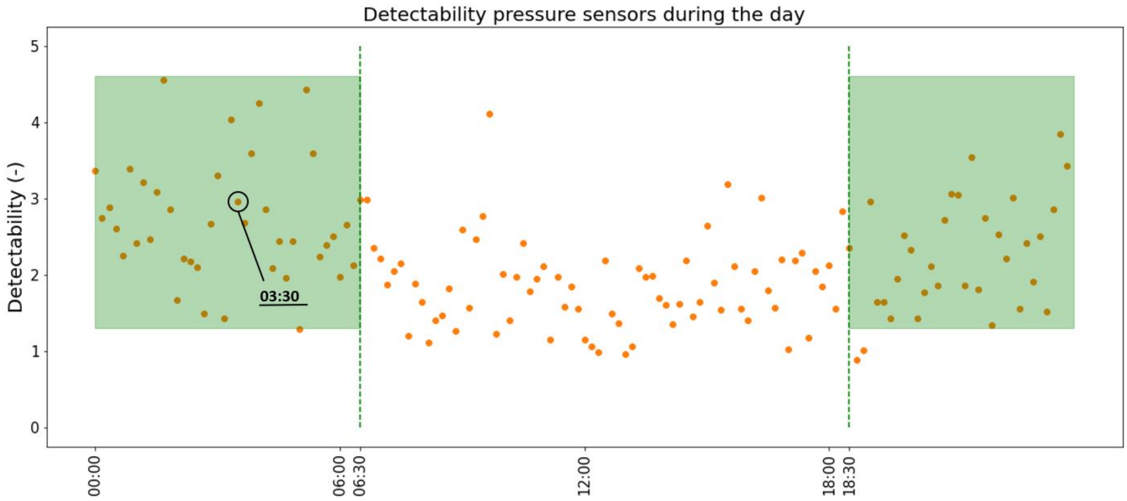

Figure 6.8: Detectability of leaks with pressure sensors at different times of the day.

An extensive comparison of the detectability of leaks at different times of the day for the flow and pressure monitoring system can be found in appendix A.15. Its conclusions are that in both cases the best leak detectability occur at night. This can be explained by the increased pressure during the night, which increases the leak's flow rate, in combination with the low flows at night, which increase the relative impact of the leak on the flow rates and pressures in the network. Furthermore, the precision of the detectability is also higher for the flow monitoring system. This can be since the modelled inlet pressure is higher when designing with flow sensors (resulting in higher leak flow rates), the daily pressure pattern for using flow sensors has a smoother curve (since it is calculated with more daily data sets) or since flows in pipes are more sensitive to random changes in demand allocation than nodal pressures. The difference in precision is likely to be due to a combination of these factors.

[11]The design with pressure sensors distinguishes two ranges, since it will only be used on Fridays and has to distinguish between Friday morning and Friday night.

Sensitivity analysis 3: Leak size

The water utility of Harare does register its leaks, but unfortunately not its sizes. The literature study has shown that leaks that can be expected to occur in the drinking water system are in the range of 2.52 - 10.8 m3/h (section 2.3). However, the WNTR-package defines leaks in terms of their area and not in terms of their flow rate (NTESS, 2019). So, different leak size were tested to see which area would give leak flow rates within the desired range. Applying leaks with a diameter of 10, 15 and 20 mm resulted in leak flow rates ranging from 3 to 18 m3/h (Figure 6.9[12]).

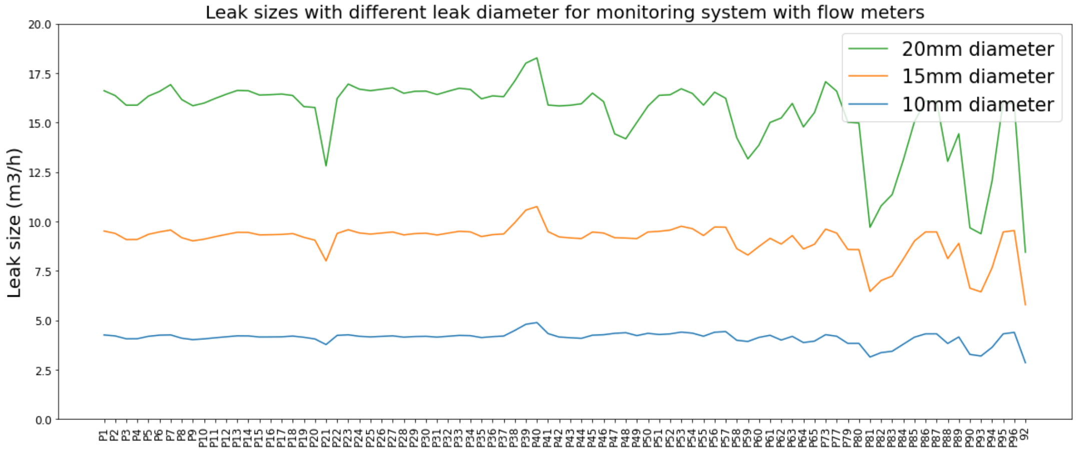

Figure 6.9: Different leak sizes (in terms of m3/h), which result from different leak sizes (in terms of diameter (mm)) in the algorithm for allocating flow sensors.

The hydraulic model that is used for designing the pressure monitoring system uses lower pressure values at the DMA inlet (section 6.1). This results in smaller leak flow rates with the same diameters sizes (10mm, 15mm and 20mm). The flow rates which occur in the leaks are between 0.5 and 12 m3/h (Figure 6.10).

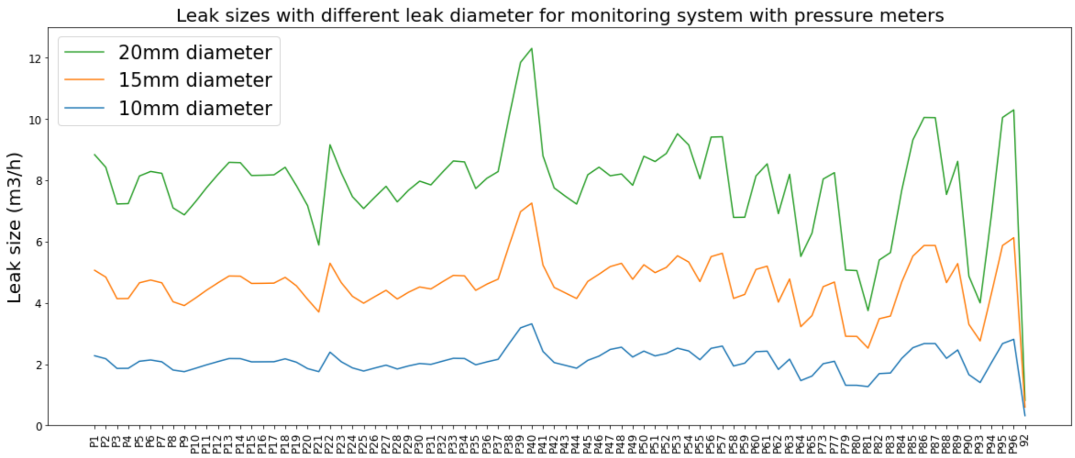

Figure 6.10: Different leak sizes (in terms of m3/h), which result from different leak sizes (in terms of diameter (mm)) in the algorithm for allocating pressure sensors.

[12]Considering that a new demand realization yields in a different pressure and therefore a different leak flow rate, the results of this graph show the average leak flow rate of 100 demand realizations with leaks.

The flow into Ashdown Park can range between 20 and 90 m3/h, with an average value of around 50 m3/h during days with continuous supply (Figure 6.6). The leak size should be realistic in relation to this inflow. A leak size of 9-10 m3/h would account for about 20% of the average inflow, which is a quite a large fraction. Therefore, the leak size of 10m3/h was chosen as the maximum leak size to be detected. This leak size is within the desired range of 2.52 - 10.8 m3/h. Figure 6.9 shows that this leak size can be included in the model for the monitoring design with flow meters by adding a leak with a diameter of 15mm. One of the goals of this thesis is to compare the performance of a monitoring system with flow sensors to the performance of a system with pressure sensors. Therefore, the leaks that are included in the model should be similar in size. So, also for the model that is used for the monitoring system with pressure meters, a leak with a diameter of 15mm is used, although the resulting leak flow rate is only around 5m3/h (Figure 6.10).

Sensitivity analysis 4: Tap capacity

The final parameter which is an input for the hydraulic model is the tap capacity. Section 4.4 (Table 4.1) has shown that the flow rate from a tap can differ a lot at different locations throughout the globe. For this thesis, a tap capacity of 0.3 m3/h was used, as this resulted of a field experiment in Ashdown Park.

To validate the performance of the model with different tap capacity parameters, the detectability KPI (equation 4.9, section 4.4) is used again. Figure 6.11 shows the detectability of a network with flow sensors under the previously determined “optimal” conditions. [13] It can be seen that a low tap capacity yields a high detectability. Earlier, in Figure 6.4, it was noted that alarm values are less extreme when the demand is better spread over the DMA. This is likely the cause of the high detectability for low tap capacities. As the alarm values at the sensors become less strict, leaks are more likely to surpass these alarm values and be detected.

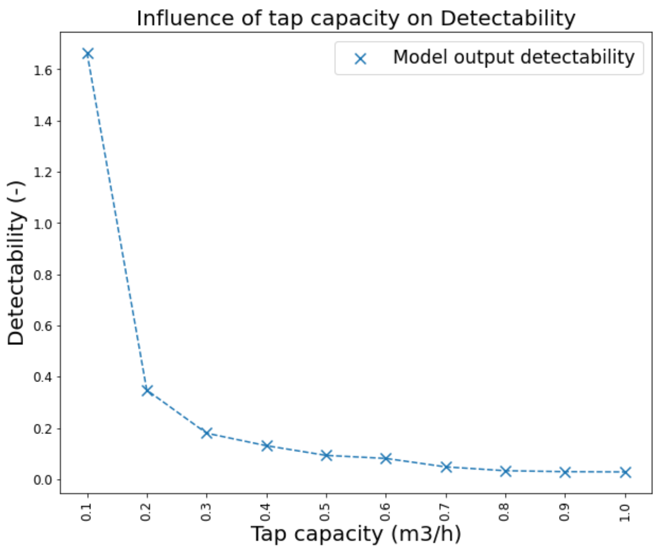

Figure 6.11: Influence of different tap capacities on the model’s output of detectability.

[13] These optimal conditions imply that the detectability was calculated at 01:20 with leaks of 15mm, 1000 demand realizations for constructing alarm values and 100 different experiments in the Boolean 3D-matrix.

6.4 Conclusions chapter 6

This chapter answers the following research question:

Q3: How can flow and pressure meters be used to detect hidden leaks?

This chapter showed that a threshold-based strategy, in this case the "Dynamic Bandwidth Monitor", can be used to identify single leaks in a DMA with flow and pressure sensors. Furthermore, it has identified differences and similarities for constructing alarm values for the two monitor systems.

During the analysis of historical data of the flow into Ashdown Park and the inlet pressure, it was found that the weekly flow pattern on days with continuous supply was more constant than the weekly pressure pattern. Therefore, it can be concluded that the water use behaviour of inhabitants in Ashdown Park has been more constant than the supply behaviour of the water utility in the concerned area. As a result, different values for the pressure and flow at the DMA inlet were used for the flow monitoring system than for the pressure monitoring system.

It was shown that modelling different demand realizations by allocating demand among the nodes with a random weighted choice and a single tap capacity, showed promising results for modelling flows that occur in Ashdown Park at night. These results were promising, since the spread of the modelled inflow was well comparable to the spread in the inflow measurements. Using a standard tap capacity is especially suitable for IWS areas, since people in IWS areas usually only have one tap directly connected to the water supply system and water ende-use devices are not directly connected to the network.

At last, the performance of the proposed monitoring systems depends on the number of experiments that are used in the Monte Carlo simulation to construct alarm values, on the time of the day during which the simulation is performed, on the leak sizes that should be detected and on the tap capacity that is included in the hydraulic model.

7. Optimizing sensor placement

The main objective of the optimization study is to place the sensors efficiently throughout the DMA, so that with a *minimum* amount of sensors a *maximum* amount of water can be saved. In this chapter it will be explained how this is achieved.

7.1 Results of optimizing sensor placement

This section shows the performance of optimally placing flow and pressure sensors. This performance of a certain configuration of sensors can be expressed by calculating a certain percentage of leaks that can be found by the sensors and calculating the areal coverage of the sensors. Both terms have been explained previously in section 4.5.

Percentage of leaks found

Figure 7.1 shows what percentage of leaks can be found in which pipe when optimally placing four flow sensors, as well as the pipes at which these sensors are placed. It shows also the probability limit, the percentage above which leaks should be found in order to be detected once a day (section 4.5).

Figure 7.2 shows the same percentages and probability limit for applying four pressure sensors. The reason for the higher probability limit (55% compared to 10%) is mainly since the DBM with pressure sensors is designed for conditions with a low inlet pressure which occurs during only 5% of all Fridays (further explained in section 4.5 and 6.1).

The results of the percentages of leaks found for optimally allocating one, two, three, four and five flow and pressure sensors can be found in appendix A.17.

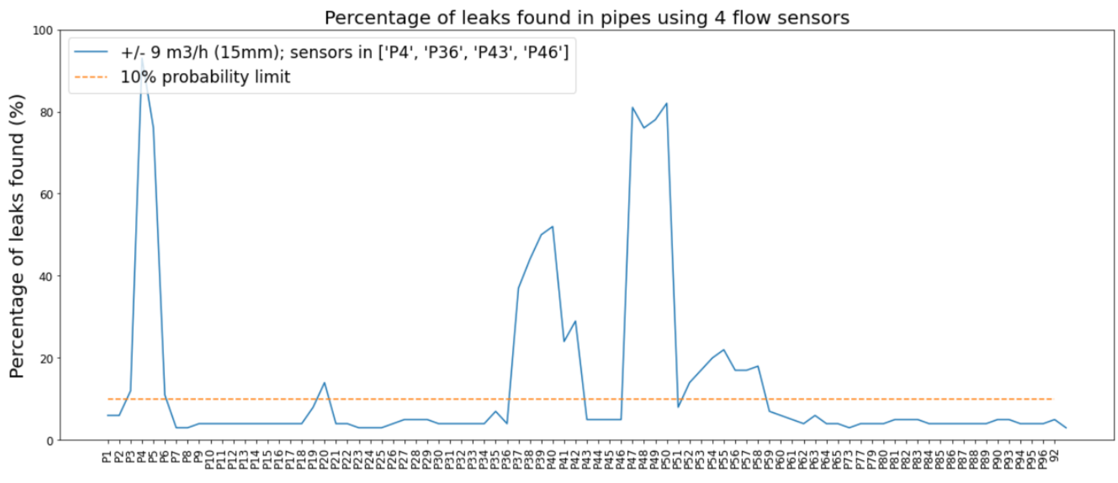

Figure 7.1: Percentage of leaks (y-axis) in certain pipes (x-axis) that are detected using four flow sensors, calculated at 01:20 for leaks of 15 mm.

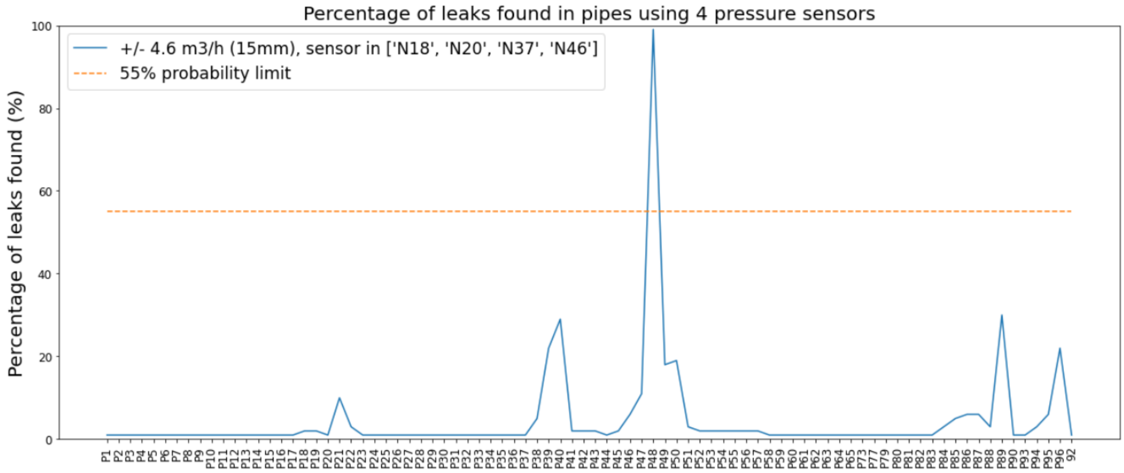

Figure 7.2: Percentage of leaks (y-axis) in certain pipes (x-axis) that are detected using four pressure sensors, calculated at 03:30 for leaks of 15 mm.

Coverage

By optimally allocating 4 flow sensors, leaks in 25% of the pipes in the DMA can be detected once a day. By optimally allocating 4 pressure sensors, leaks in only 1% of the pipes in the DMA can be detected once a day. The optimal placement of different number of sensors[1] and their coherent coverage can be seen in Table 7.1. From this result it can be concluded that a monitoring system with pressure sensors as designed in this thesis is less applicable for Ashdown Park than a flow monitoring system.

Number of sensors	Coverage flow monitoring system (%)	Coverage pressure monitoring system (%)	Coverage pressure monitoring system with pressure dependent DBM (%)
1	4.8 %	1.0 %	2.4 %
2	8.3 %	1.0 %	6.0 %
3	17.9 %	1.0 %	8.3 %
4	25.0 %	1.0 %	8.3 %
5	27.4 %	1.0 %	8.3 %

Table 7.1: Coverage of a given amount of optimally placed sensors.

The results from Table 7.1 can be visualized to show at which locations the sensors can be optimally placed and at which pipes leaks can be found. Figure 7.3, 7.4, 7.5 and 7.6 visualize where flow sensors would be located in the optimal scenario and where leaks would be found. This visualization shows that the coverage of the DMA increases as the amount of placed flow sensors increase. With more than four flow sensors, the increase in coverage per added flow sensor seems to decline, since the coverage of five flow sensors is only 27.4% (Table 7.1).

[1] This amount excludes the number of sensors that are needed to measure pressure and flow at the DMA inlet, as will later be explained in section 7.4.

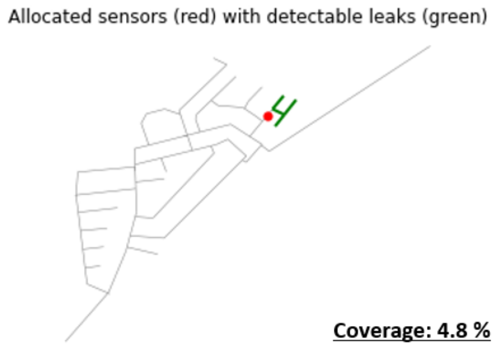

Figure 7.3: The optimal placement of one flow sensor (at pipe P4) for detecting leaks of 15mm.

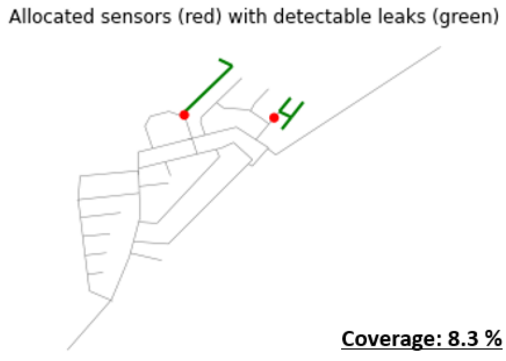

Figure 7.4: The optimal placement of two flow sensors (at pipes P4 and P46) for detecting leaks of 15mm.

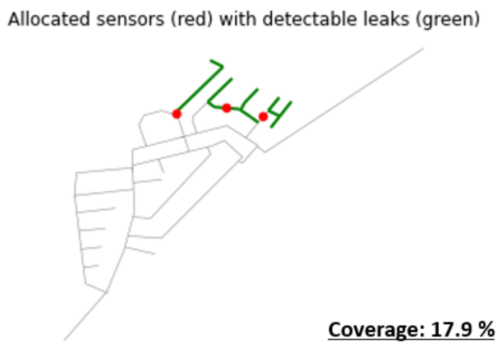

Figure 7.5: The optimal placement of three flow sensors (at pipes P4, P36 and P46) for detecting leaks of 15mm.

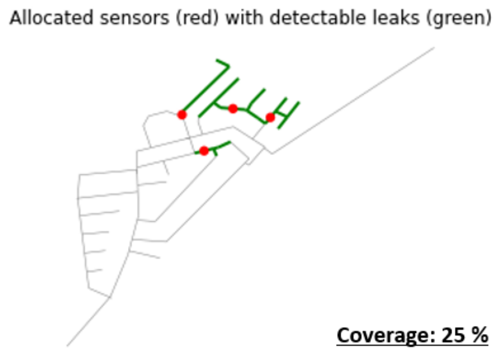

Figure 7.6: The optimal placement of four flow sensors (at pipes P4, P36, P43 and P46) for detecting leaks of 15mm.

Figure 7.7 and 7.8 show the placement of one and four pressure sensors. As shown, the coverage does not increase for using more than one pressure sensor. Hence, the monitoring system with pressure sensors is not very effective in Ashdown Park for detecting leaks.

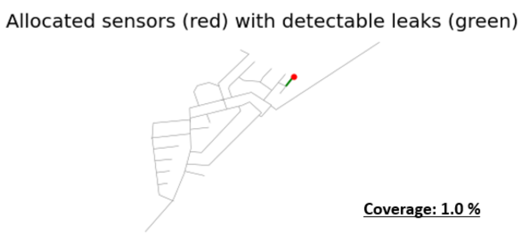

Figure 7.7: The optimal placement of one pressure sensor (at node N46) for detecting leaks of 15mm.

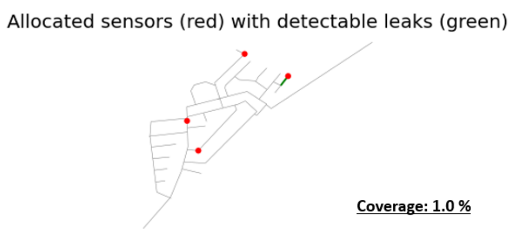

Figure 7.8: The optimal placement of four pressure sensors (at nodes N18, N20, N37 and N46) for detecting leaks of 15mm.

7.2 Pressure dependent DBM with pressure sensors

So far, it has been assumed that the alarm values of the DBM remained static at all times, since then the DBM is easily interpretative for the monitoring personnel. However, a consequence of this choice is that the artificial DBM's with pressure sensors needed to be calculated by using the 5-percentile pressure on Friday as inlet pressure at the DMA. This resulted in low leak volumes and a low chance of finding a leak once a day.

Now, assume that is possible to construct a DBM whose alarm values change as the measured inlet pressure changes. A visual example of this can be seen in Figure 7.9. The consequence of this pressure dependent DBM is that any measured inlet pressures can be used to construct the alarm values. So, when constructing the alarm values and calculating the 3D-matrix that stores the detectability of leaks, it is possible to use the average weekly pressure measurements and the average weekly flow measurements as input values for the monitoring system with pressure sensors. Doing so, gives us an estimation of how many leaks could be detected with pressure dependent DBM's[2]. The same values will be used for the parameters which are important to construct the flow and pressure monitoring systems (section 6.3) with non-changing alarm values. Also, the probability limit will be set similar to the monitoring design with flow meters (at 10%), since both designs use input values with an equal chance of occurrence.

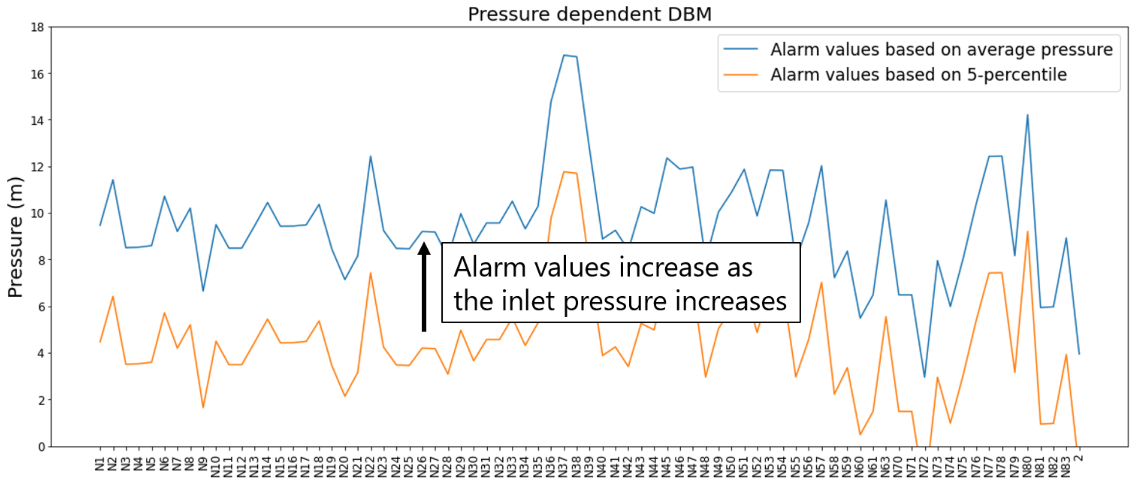

Figure 7.9: Example of how the pressure alarm values change with a different inlet pressure.

The amount of leaks that can be found using 4 pressure sensors in a pressure dependent DBM network, together with the probability limit can be found in Figure 7.10. By optimally allocating 4 pressure sensors and using a pressure dependent DBM network, leaks in 8.3% of the pipes in the DMA can be detected once a day, which is an improvement from the "normal" pressure monitoring system with non-changing alarm values. The percentages of leaks found for optimally placing one, two, three, four and five pressure sensors can be found in appendix A.17. Table 7.1 shows the coverage that can be reached when optimally placing a different number of pressure sensors with a pressure dependent DBM system. Figure 7.11, 7.12, 7.13 and 7.14 visualize where one, two, three and four pressure sensors with a pressure dependent DBM system would be optimally placed and in which pipes they would detect leaks.

[2]When the inlet pressure would be higher than the average, more leaks can be detected and with a lower inlet pressure less leaks. However, this will not be elaborated upon in this thesis.

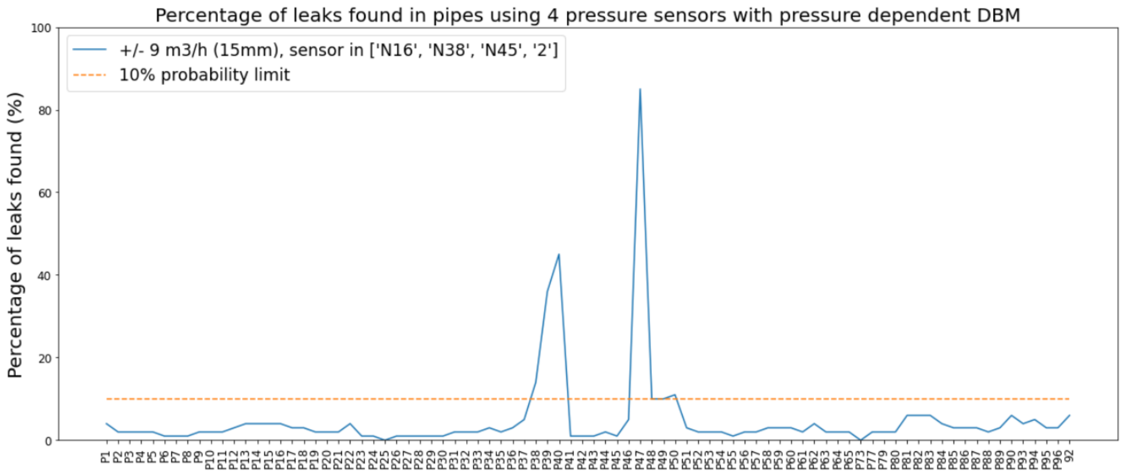

Figure 7.10: Percentage of leaks (y-axis) in certain pipes (x-axis) that are detected, using four pressure sensors and a pressure dependent DBM.

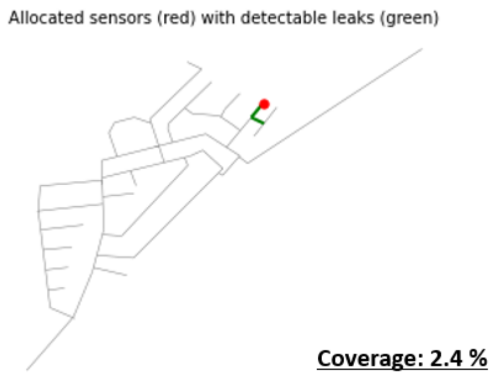

Figure 7.11: The optimal placement of one pressure sensor (at node N45) for detecting leaks of 15mm with a pressure dependent DBM.

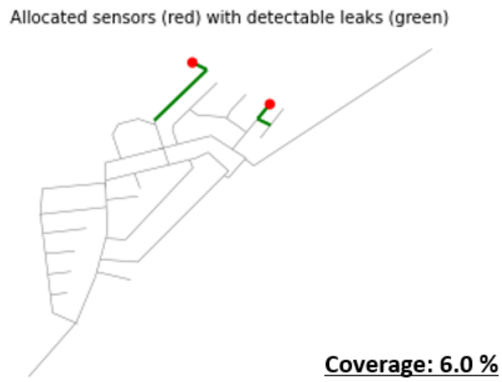

Figure 7.12: The optimal placement of two pressure sensors (at nodes N38 and N45) for detecting leaks of 15mm with a pressure dependent DBM.

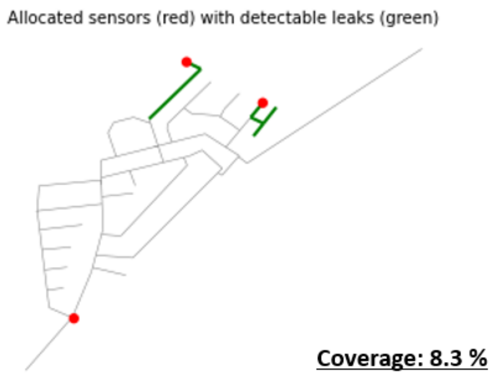

Figure 7.13: The optimal placement of three pressure sensors (at nodes 2, N38 and N45) for detecting leaks of 15mm with a pressure dependent DBM.

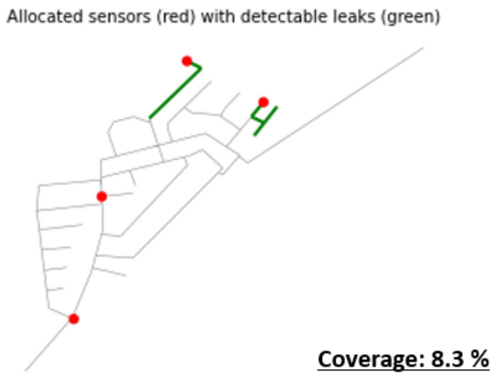

Figure 7.14: The optimal placement of four pressure sensors (at nodes 2, N16, N38 and N45) for detecting leaks of 15mm with a pressure dependent DBM.

These visualizations show that the coverage of the DMA increases as the amount of allocated pressure sensors increase, albeit its increase less than for the monitoring system with flow sensors. When the third pressure sensor is placed, a new pipe at the other side of the DMA becomes visible. This remarkable results occurs because a “false alarm” has pushed to probability that the pipe was found from 9% to 10%. This can be seen in the figures that show the percentages of leaks found in the pipes (appendix A.17). More than two pressure sensors will not increase the coverage of the DMA significantly.

7.3 Important factors for optimal sensor placement

Pipe and node characteristics of the locations that were “chosen” to place sensors were compared to the characteristics of unsuitable pipes and nodes, to see if certain factors influence the suitability of a location for sensor placement. The figures of this chapter can be found in appendix A.19.

Important factors when placing flow sensors

For this analysis, the pipes with flow sensors (Figure 7.6) were compared with the other pipes. The chosen pipes have an average elevation, so the elevation will probably have a limited effect on the optimal location for flow sensors (Figure A.67, appendix A.19). The chosen pipes also have an average diameter, so the diameter will probably have a limited effect on the optimal location for flow sensors as well (Figure A.68, appendix A.19). At last, the chosen pipes have a relatively low flow, but these low flows are not very unique compared to other pipes (Figure A.69, appendix A.19). The average flow in the pipes will therefore probably have a limited effect on the optimal location for flow sensors as well.

Important factors when placing pressure sensors

For determining the important factors for allocating pressure sensors, the characteristics of only two nodes were compared to the other nodes in the network. These two nodes (Figure 7.12) were the only nodes that significantly increased the coverage when optimally placing pressure sensors with a pressure dependent DBM system. The two chosen nodes have a very low elevation, compared to the other nodes (Figure A.70, appendix A.19). Therefore, it seems advisable to place pressure sensors for a monitoring system on locations with low elevations. The pressure, which is directly related to the elevations, in the two chosen nodes is high as well (Figure A.71, appendix A.19). Therefore, the importance of pressure should be seen in line with the importance of elevation, as nodes where high pressures occur seem very suitable for allocating sensors for a pressure monitoring system.

When looking at Figure 7.3 till 7.14, it can be noted that for monitoring systems with pressure sensors and for monitoring systems with flow sensors, favourable locations to place sensors seem to be at the branched part of the DMA. Parts in the DMA were the network is more looped seem less suitable for placing sensors. This is probably since nodal demands behind leaks can be relatively easy fulfilled in looped systems, as the flow can take detours in the system. In branched systems, more flow is “forced” to pass the leak, causing higher pressure and flow deficits in the area behind the leak.

7.4 Starting point of the monitoring system: The DMA entrance

The above monitoring system is designed for the Ashdown Park DMA in Harare. The reason to apply the modelling study to this specific DMA was that its incoming flow and inlet pressure had been monitored over the last years. These measurements were used to calibrate the hydraulic model and can therefore be regarded as the first measuring devices of the monitoring system. The locations where these devices were installed, at the entrance pipe and node of the network, were incorporated as potential locations where the allocation algorithm could place its pressure or flow sensor. However, the algorithm did not place sensors at these locations, probably because locations at the networks' branched ends were more favorable for leak detection. These could lead to the false perception that a monitoring system could be designed without measuring the inlet pressure and incoming flow. However, the devices that perform this measurements are very important for the system and should be taken into account (see Figure 7.15). In this chapter some arguments are aligned as to why these devices are important for the monitoring system.

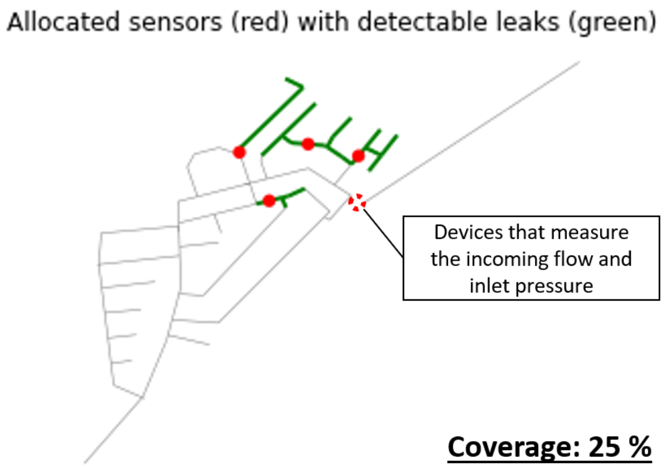

Figure 7.15: Measuring devices that measure the inlet pressure and incoming flow, which should be included into the design of the monitoring system.

Firstly and most importantly, the incoming flow and pressure allow us to construct and calibrate the hydraulic model of the DMA. Changing demands or the pressure of the reservoir in the model (previously shown in Figure 4.3) is always noted at the DMA inlet, making it such a valuable calibration point.

Furthermore, every NRW assessment method starts with estimating the system input volume (level 0, Figure 2.1). When making an attempt to estimate the spatial division of NRW throughout the system, it could be useful to address NRW per area. This allows the utility to identify which areas are "weak spots" in the system and where adjustments should be made. For the estimation of the NRW for a DMA, the system input volume would be measured first and used to calculate or estimate other volumes. So, the flow measurements at the DMA inlet are also very useful to make estimations of NRW per area.

Concluding, the pressure and flow meters at the entrance of the DMA are crucial for the monitoring system and to properly estimate NRW in different areas of the system. It does not necessarily have to be two separate meters, as integrated meters that measure both pressure and flow are currently finding its way into the market (Ernst Vink, 2020). However, in this thesis it is assumed that every monitoring system starts with a flow and pressure meter at the inlet of the DMA. If the monitoring system would be designed for a transport network between DMAs, instead of a distribution network, pumping data about the added flow and pressure to the system would also be sufficient for the same purposes.

7.5 Conclusions chapter 7

This chapter answers the following research question:

> **Q4:** How can the placement of the meters in the monitoring system be optimized?

A new method was shown that can be used to determine the optimal sensor placement for constructing monitoring systems with flow and pressure sensors, building upon the Boolean matrices that were constructed in the previous chapter.

Its results showed that a flow monitoring system in Ashdown Park with four sensors is able to detect leaks in a larger part of the DMA (25%) than a pressure monitoring system (1%). The performance of the pressure monitoring system slightly increased (from 1% to 8.3%) by making its alarm values dependent on the inlet pressure of the DMA, but it was still found to be below the performance of the flow monitoring system.

Furthermore, important factors that characterize the optimal locations for flow and pressure sensors were identified. In the case of the optimal locations for flow sensors, no specific pipe characteristics could be found that distinguished the chosen pipes from other pipes. The main characteristics of the sensor locations in this case was that they were placed in the branched part of the network. The optimally allocated pressure sensors were placed in the branched part of the network as well. Above that, the nodes that were considered as the optimal location for pressure sensors were characterized by low elevations and high pressures.

At last, every monitoring network should consist of devices that measure the flow and pressure at the DMA entrance, as this data is crucial for constructing and calibrating a hydraulic model.

8. Business model

The sequential step in this thesis is to construct a business case. The first section of this chapter describes the monitoring system in different contexts, which should be well considered when using the system for entrepreneurial purposes. Afterwards, potential financial benefits of the system shall be shown. Finally, the chapter is summarized by a sustainable business model canvas, which gives an overview of how the monitoring system would function in the existing dynamics within a water utility and potential partnerships.

8.1 Entrepreneurial contexts

A lot of cities with continuous supply are transiting to a "smart water system", where water quantity and quality are monitored in real-time, allowing the utility to making efficient repairs and interventions in the system when needed. This does not only concern flow and pressure meters, but also water quality measurements are crucial for ensuring a safe water supply. The first three sections in this chapter will focus on these "smart water systems". These sections contain (1) a roadmap for IWS to move towards smart water systems, (2) practical barriers for implementation of the monitoring system in a historical context and (3) different options for the local community to fulfill their drinking water demand. After these chapters, the focus shall be more on hydraulic monitoring systems, like designed in this thesis. These chapters will explain (4) the process chain from leak detection to leak repair and (5) the potential positive and negative effects that the monitoring system can have on its surrounding society.

Road map towards smart IWS systems

Whereas many continuous supply systems, such as Amsterdam, have moved towards a smart water system, there are a number of barriers that prevent the transition of an IWS system to a smart system. These barriers are visualized in Figure 8.1. Examples of how these barriers have affected this thesis will be elaborated upon below.

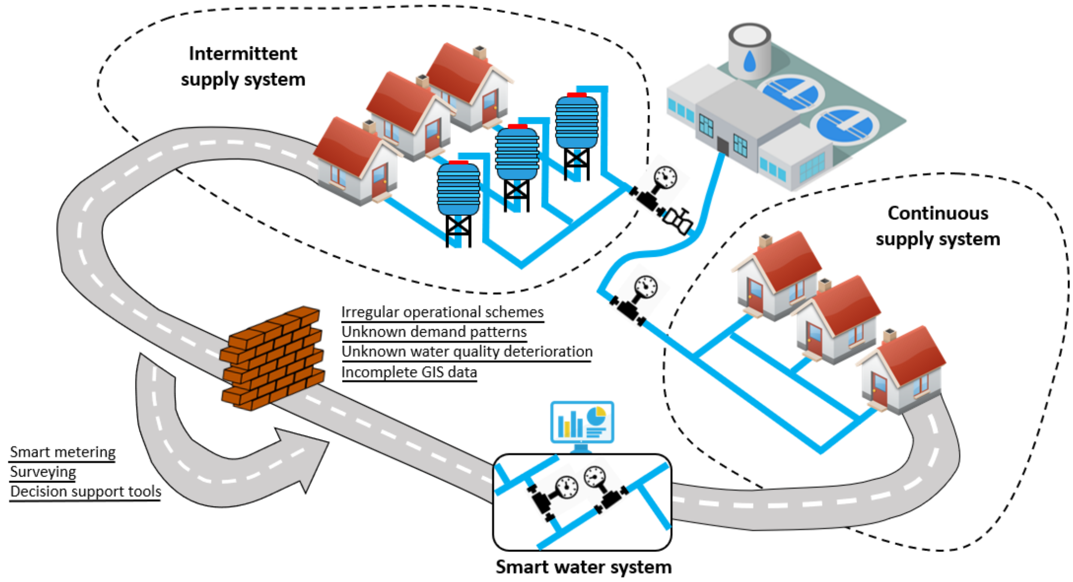

Figure 8.1: Roadmap for IWS systems towards a smart water system

Irregular operational schemes can be found when taking a look at the number of days that were removed when the days with anomalies were filtered out. This reduced the numbers of Saturdays and Sundays that could be used for constructing alarm values from 64 to 8, implying that 88% of the days in the weekend contained moments were the pressure or flow dropped to 0. Also, during one of the interviews it was mentioned that some service areas receive extra water at moments when extra water is available (Shana, 2020). Although the intentions of this extra provision are good, it adds irregularity to the supply scheme. This irregularity increases the number of anomalies that are not caused by leaks and so makes the construction of alarm values and the detection of leaks with the DBM a lot more difficult.

The *unknown demand patterns* result in not having a standard consumer demand patterns that can be used for modelling IWS systems. This makes it difficult to construct reliable hydraulic models, which often form a basis for smart water systems. This absence of standard demand patterns eventually led to the novel method of modelling demand realizations with the new tap capacity method.

Water quality deterioration is outside the scope of this research, but it has shown to be a problem for IWS systems. When performing some general research, little data was found that shows the water quality deterioration in IWS systems and its impact on water quality and health quantitatively. Using rapid quality sensors, combined with hydraulic modelling, water quality degradation in a distribution system could be easily mapped and interventions could be proposed accordingly, as shown by Sakomoto et al. (2020).

At last, *incomplete GIS data* turned out to provide some challenges as well. This thesis research started with the purpose to add value to the leak detection system in Nairobi, with the intentions of visiting Nairobi and performing field measurements as well. Due to Covid-19, this visit to Nairobi was cancelled, making the design of the monitoring network very dependent on digital information (GIS) of the water network, which could be shared by mail. Unfortunately, the digital information was often too inconsistent to allow for very reliable models. This is the reason why, during the design phase, the target area shifted from Nairobi to Harare. Although Harare had some DMA's whose digital archives were complete, it is still working on digitizing its entire network. Without this GIS data, no hydraulic models can be constructed and smart water systems can not be implemented.

Opportunities for IWS systems

Local utilities can overcome these barriers by starting with smart metering, which can initiate positive feedbacks and so increase its own efficiency (further explained at the end of this section). Furthermore, surveying and smart household metering can help in determining demand patterns. Decision support tools can help local utilities divide supply equitably and regularly. Finally, GIS systems will need to become up-to-date by a hard-working and skilled workforce.

Despite these barriers, it is also important to stress the unique opportunities that can be found when implementing smart water systems in IWS areas. Most utilities with continuous supply have strongly looped water distribution networks and struggle to construct DMA's within their networks. In IWS systems, closed off city sections are created on a daily basis. If these areas would be monitored at their entrances, you would instantly create DMA's with little effort. Digitizing these DMA's, in combination with a supply schedule which a consistent weekly pattern, would already be a sufficient environment to perform the calculations as described in this thesis. By determining the optimal sensor placement for a DBM-system, it would be possible to create leak awareness and reduce leakage efficiently.

Practical barriers in a historical context

The scope of this thesis focuses around the urban water system infrastructure of large cities with IWS conditions. The availability and quality in urban areas in general tends to worsen (Dos Santos et al., 2017). Ageing infrastructure and lack of finances for leakage repair are problems for the local water utilities in Nairobi and Harare (Shana, 2020)(Mugo, 2020). Mismanagement is a cause that Dos Santos et al. (2017) regularly mentions as one of the causes of this high NRW-percentage. Although corruption is a cause that still occurs in these counties, it is important to realize that the task for the local utilities is often very different than is the case in western countries.

Let us for instance consider the growth of some of the large cities in Sub-Saharan Africa. In times of colonization of England in Kenya, many local rural farmers were evicted from their farmland for not having the proper documents to prove that they were rightful owners of the land. The land was taken by the settlers and the local farmers, having no place to go, moved to the suburbs of Nairobi (Anderson, 2005). In the decades after colonization, the trend shifted from settlers owning the land towards selling farmland to large-scale investors. The poverty-reducing impact of this type of farming was little and the shift from rural towards urban land continued (Schutter, 2011). This shift from rural livelihoods towards urban living environments was way stronger in African cities than in Europe of the past decades, as shown by the growth of Amsterdam, Kenya and Zimbabwe in Figure 8.2 and the growth of the proportion of urban population in Sub-Saharan Africa in general (Figure 8.3). To express the growth from 1950 to 2020 in numbers, Amsterdam has grown by a factor 1.4, Harare has grown by a factor 10.7 and Nairobi has grown by a factor 34.4.[1]

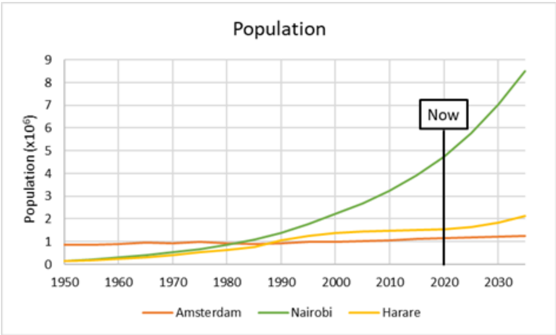

Figure 8.2: The population growth of Amsterdam, Nairobi and Harare World Bank (2020)

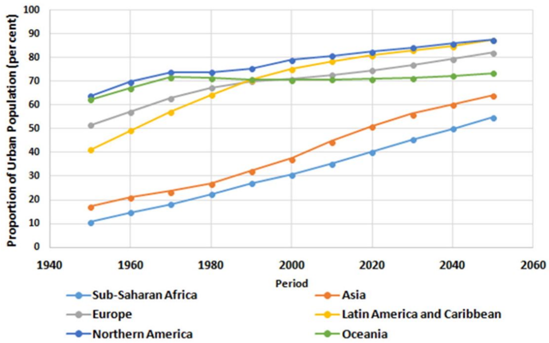

Figure 8.3: The population shift towards urban regions Dos Santos et al. (2017)

.

A rapid growth of the city, as can be seen in Nairobi and Harare, puts a lot of pressure on its drinking water infrastructure. Capital expenses should therefore be used to make proper investments. However, the economic growth of Kenya and Zimbabwe stayed far behind in comparison to the Netherlands, as can be seen whilst comparing the Gross Domestic Product (GDP; in dollars) per capita per country from 1960 to 2020 in Figure 8.4.[2] To express the GDP per capita growth in numbers, the GDP per capita in the Netherlands has grown by a factor 49.0, in Zimbabwe it has grown by a factor 5.3 and in Kenya it has grown by a factor 18.6.[3]

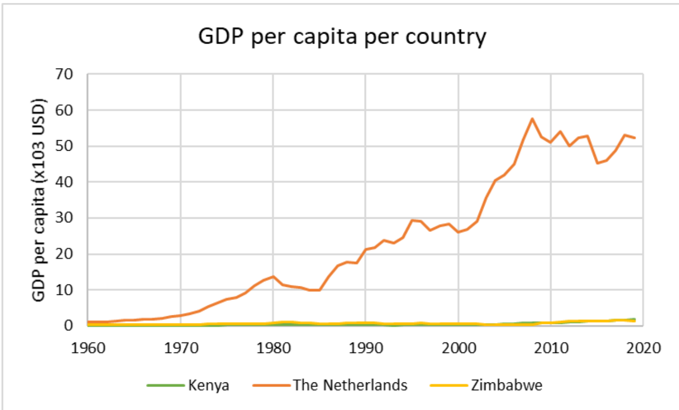

Figure 8.4: GDP per capita from 1960 until 2019 in the Netherlands, Kenya and Zimbabwe World Bank (2020).

[1]Amsterdam has grown from 850,777 to 1,148,972 inhabitants, Harare has grown from 142,652 to 1,529,920 and Nairobi has grown from 137,456 to 4,734,881. (World Bank, 2020)

[2]It is definitely the case that many more factors play a role in this economic comparison and the population growth in the large urban cities. This historical context should not be judged by the exact relations and numbers, but it tells a general tale about the challenges for local water utilities in cities like Harare and Nairobi.

[3]Exact numbers. Netherlands: from 1,069 to 52,331 USD. Zimbabwe from 279 to 1,464 USD. Kenya from 97 to 1,817 USD. (World Bank, 2020)

At last, the lingering water scarcity, one of the principal causes of intermittent supply, is a danger that comes along with urbanization (Vairavamoorthy et al., 2001). Putting all these issues together, local utilities in cities like Harare and Zimbabwe need to connect more people to their water supply with less capital. Although each region has its own challenges, the cause of mismanagement is easily put and should not be named as a standalone cause for insufficient access to drinking water, like Schutter (2011) argues that mismanagement in large-scale land investments is not the principal cause of rural poverty. Incomplete GIS-data or insufficient data from the system are all understandable situations taking into account the way some IWS cities have evolved. Efforts in knowledge-sharing, general cooperation and capacity building between Amsterdam and Nairobi, like the efforts of WorldWaternet, with the intention to equalize the quality of both supply networks, should therefore be encouraged. One might even call it a moral duty of former colonizing countries.

Drinking water alternatives in Harare

Water is a basic need for humans and for life is general, since no life can occur without water. So, in areas where water supply services limited, inhabitants still will need to find other ways to ensure themselves with water. In IWS systems, this can be done by storing water which was collected during days with supply. However, one can also ensure itself with water by turning to other alternatives than piped water. This chapter shows the results of interviews held with inhabitants from Harare. The price, quality, reliability and collection time of different alternatives for fetching water were scored from 1 to 5 (1 referring to a bad score and 5 referring to a good score). These scores are subjective judgements based on the available information. It should be realized that these comparisons can only be made for people who are connected to a water supply system, otherwise they would not have the option to acquire piped water.

In an attempt to view the water situation from the eyes of an inhabitant of Harare, it was chosen to interview a local inhabitant. To ensure that this person was not related to the local water utility, to prevent conflicting interests, it was chosen to approach someone from my pool of personal connections. An interview was held with this person through WhatsApp and additive information was provided by a MSc-student from the university of Bulawayo. Due to privacy reasons, both persons shall remain anonymous. A full account of these interviews can be found in appendix A.20.

The approached inhabitant of Harare lives in Budiriro, which is an area in the South West of Harare. He would describe Budiriro as an area with a high-density population, whereas he describes Ashdown Park as a low-density area. He says the condition of the piped drinking water system has been very good and reliable until about 10 years after the independence of Zimbabwe[4]. Currently, the supply of water is very irregular. When there is water, he uses the bathtub to store pipes water (Figure 8.5). Unfortunately, the quality of the water is insufficient to be used as drinking water, as sediments from the water start to settle after a few hours. He fetches drinking water from a borehole in the area, which is stored in 25L buckets (Figure 8.6) (Inhabitant Budiriro, 2020).

[4]Zimbabwe gained indepence in 1980

Figure 8.5: Storage of piped water in the bathtub in Budiriro.

Figure 8.6: Drinking water buckets of 25L.

Price: The tariff for piped drinking water at the 1st of November 2020 was pegged at 90 ZWL per cubic meter for the first five cubic meter that are consumed monthly (Magedi MacDonald, 2020). After the first five cubic meter, the price slightly increases. Drinking water buckets from the bore hole are priced at 7.50 ZWL per bucket and clean bottled water from the store ranges between 30 and 50 ZWL for a 500mL bottle (MSc-Student University of Bulawayo, 2020). Converting all prices to the price for 25L one would pay 1.25 ZWL for piped water (score:5), 7.50 ZWL for borehole water (score:4) and 1500-2500 ZWL for bottled drinking water (score:1).

Quality: Bottled water has been purified in a commercial drinking water treatment plant and has therefore a high quality (score:5). Inhabitants in Budiriro regard borehole water as a source with a higher water quality than piped water. However, research of Muzenda et al. (2019) has shown that also boreholes in Harare are often unfit for drinking water purposes. Unfortunately, no field research could be performed to determine the quality of the different sources individually. So, the judgement of the inhabitants was followed that borehole water (score: 3) was a source with a slightly higher water quality than piped water (score: 2).

Reliability of supply: Bottled water is always available in the stores (score:5). Borehole water is tapped from ground water. Ground water is available throughout the year, but its quality and level changes throughout the year (Muzenda et al., 2019). Also, an operator is needed to maintain the borehole. Therefore its reliability of supply is scored at 4. Intermittent supply conditions result in an irregular supply pattern, which results in inhabitants not knowing when or if they receive water. The reliability of supply of piped water was scored with a 2.

Collection time: When there is supply, piped water comes from the taps at the houses. Therefore, no time is needed to collect the water (score:5). Bottled water is sold at stores, which are numerous in Harare. However, people still will need to go to the store to buy water (score:3). Boreholes are less numerous than stores, which result in more time needed to collect the water (score:2). Of course, the collection time can very a lot per individual as the proximity of its house to shops or to a borehole differs per inhabitant.

The overall comparison of water alternatives in Harare can be seen in Figure 8.7. The main aspects that make people choose borehole water over piped water (when both are available to them) for drinking water purposes are its quality and the reliability of supply. Actions that can still be performed with water with a lower quality, such as washing and bathing, are usually performed with piped water, because of its pricing and better collection time (Inhabitant Budiriro, 2020). Based on the above comparisons, solving water quality issues and operating the water supply system according to a more regular schedule, would increase the competitive position of piped water against other water alternatives and increase the amount of piped water that is used by consumers. Therefore, a smart water system could be a beneficial tool for the local water utility to improve its competitive position against other drinking water sources.

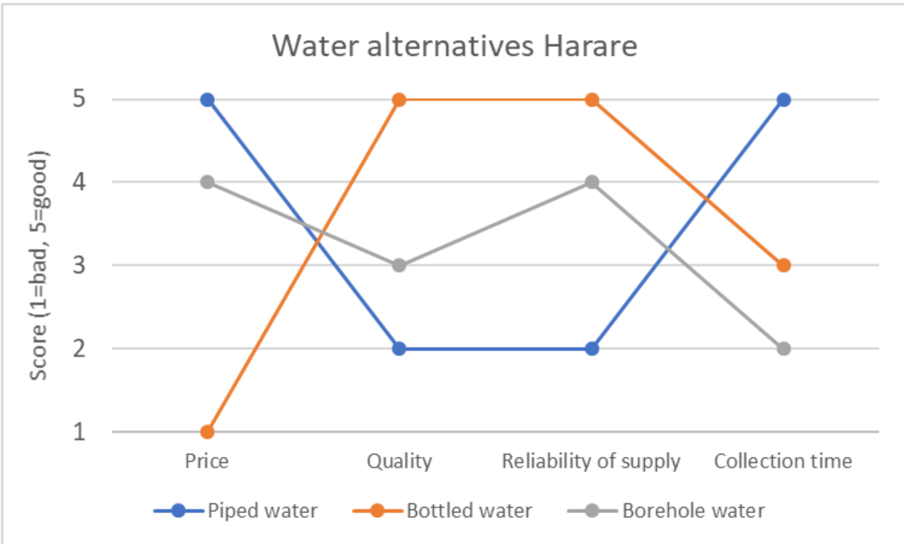

Figure 8.7: Comparing different aspects of water alternatives in Harare.

Process chain of leak repair

It takes several steps to repair a leak after is has occurred, which were identified to determine the position of the monitoring system among these steps. The duration of these steps, the leak repair time, determines the amount of water that is lost. These steps, together with the flows of information that take place after a leak emerges are shown in Figure 8.8. After a leak has occurred, awareness of the leak can be created through customer complaints, by using a monitoring system or with a minimum night flow (MNF) analysis (Ziegler et al., 2011). Further localization and repair are then performed by the leak repair team, after they have been notified.

The quality and quantity of these information flows can differ per region. For example, an inhabitant of Amsterdam would be very likely to report having low pressures or having no supply its local utility (line 3 in Figure 8.8), since it is an abnormal situation. An inhabitant of an area with IWS, for instance in Nairobi, will be less likely to report the same event, since this inhabitant is used to low pressures or no supply. If local inhabitants share information with visible leaks with the customer services, who on their shares the information with the leak repair team, who repairs the leak, then the issue is solved.

Most people in Harare report visible leaks to the municipality, who on their terms decide whether to respond quickly or not (MSc-Student University of Bulawayo, 2020). Therefore, the extra benefit of having a monitoring system in this case comes from detecting the leaks that are not visible nor reported. Another option to detect leaks that are not visible at the surface is by using an MNF analysis (its concept is explained in section 2.1). Using a real-time monitoring system to detect leaks has the added benefit of the leak being found within the same day as the operator uses the monitoring system. This detection time, and therefore the amount of water lost, depends on the frequency of which the operator checks the monitoring system. This is likely to be more frequent than the rate at which the utility conducts minimum night flow analyses.

So in this case, the monitoring system could be a useful tool for repairing leaks that are not visible at the surface. Whether a leak that is detected will be repaired depends largely on the work of the leak repair team and its communication with the operator of the monitoring system. In Harare the leak repair team has a lot of struggles, since they for instance lack the budget to buy diesel to travel to areas where leaks occur (Moors, 2020). Therefore, a lot of leaks are not repaired. This is not a problem which a monitoring system can solve, since it covers a different side of the chain of processes in leak repair.

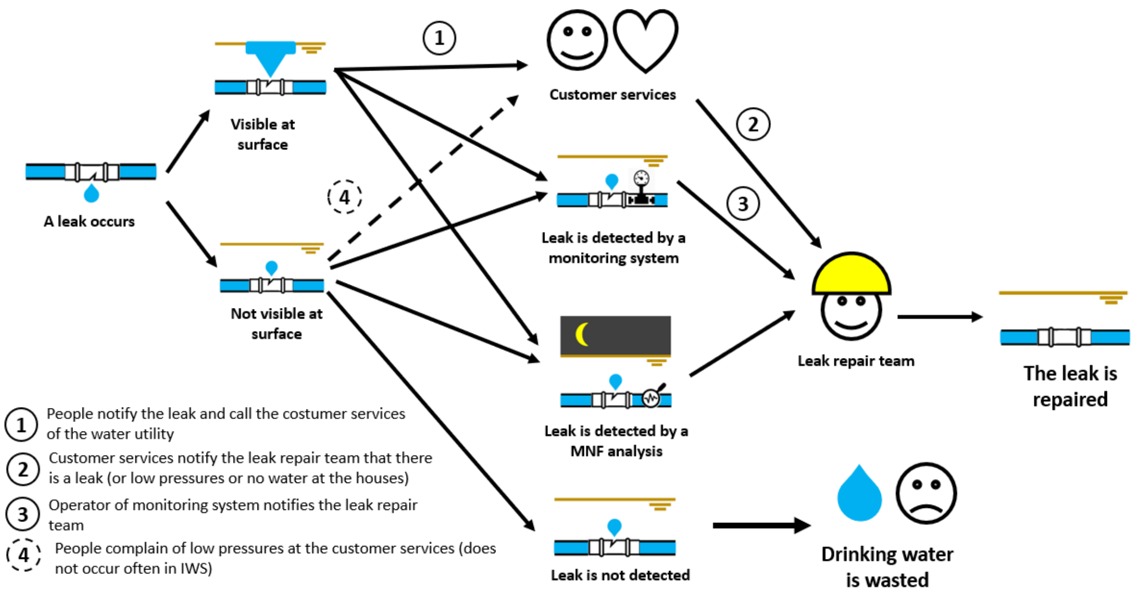

Figure 8.8: The chain of processes from leak occurrence to leak repair

Positive and negative effects of hydraulic monitoring on society

A smart water system can in many ways have positive (or negative) effects on the state of the water network and on the system's users. First of all, reducing leakage in the network will lead to less water quality degradation. Ground water intrusion at leaky places is a likely cause for this degradation (Sakomoto et al., 2020), which will be diminished if the leakage reduces. Reduced leakage and less water quality degradation will on their terms result in an improved quantity and quality of water in the network. With more quantity and quality, the utility can generate more revenues. One of the possible investments with this extra revenue could of course be more smart water systems, in which case this cycle can be sustained (Figure 8.9).

On itself, flow and pressure monitoring can lead to a circle of positive feedback relations which continuously improve each other, as shown in Figure 8.9. Less leakage will lead to a higher pressure in the water system, which on its term increases the volume of water that leaves through a single leak, which improves the detectability of leaks, so even more leaks can be found and repaired. These repairs result in a reduced number of leaks in the system and so on.

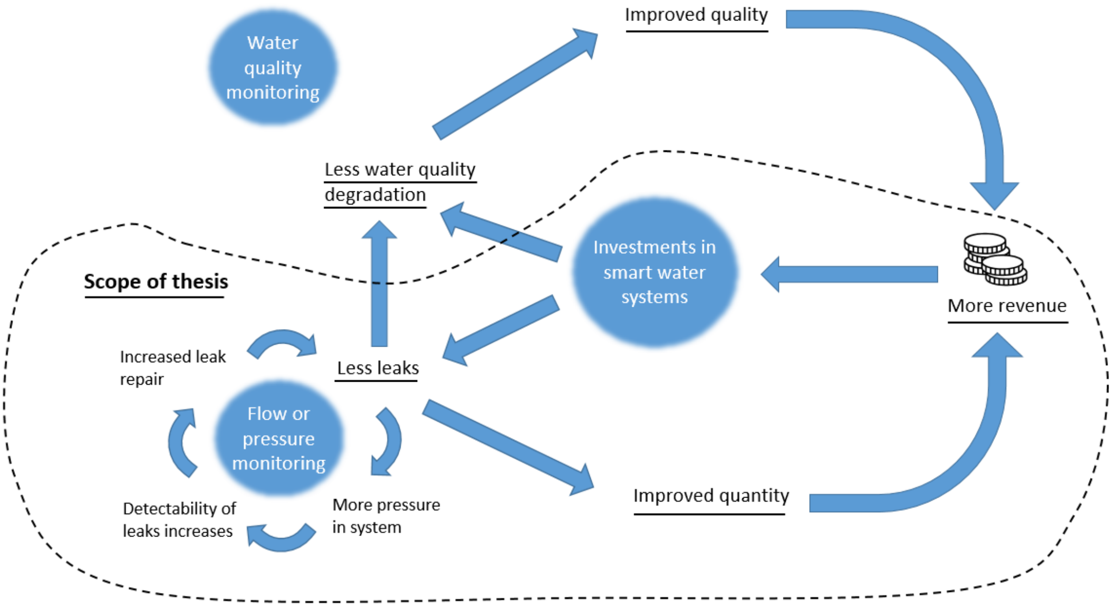

Figure 8.9: General cause and effect relations smart water systems.

This thesis focuses on hydraulic monitoring with flow and pressure meters, shown by its scope in Figure 8.9. Therefore, economical and social associated outcomes within this scope will be described. These outcomes often imply managerial choices. For instance, more revenue could mainly result in a reduced water tariff *or* be mainly invested in smart water systems. These choices are visualized by two arrows originating from the same stem (Figure 8.10). Effects which are not a choice per se are visualized by straight arrows. These positive (green) and negative (red) effects of associated outcomes on the economical and social situation of the users of the system are visualized in Figure 8.10 as well.

There are several positive associated outcomes which might occur. When the water quantity increases, the water utility can choose to buffer it and improve the reliability of supply. It can also allow more water to reach the customer and increase its revenue. This extra revenue could be then be invested in new infrastructure to increase water access in the city. It could also be used to reduce the water tariff and boost the local economy. If the extra revenue would be invested in more smart water systems, it will augment the amount of jobs in these systems. Less energy is needed for pumping when leakage reduces, so pumping costs can be reduced (Colombo and Karney, 2005). If there are customers at the end of a network that could not be reached due to insufficient pressure, an increased pressure might give them access to drinking water as well. Lastly, when there are more leaks to repair, this results in jobs for personnel in network maintenance and leak repair.

Negative associated outcomes might occur as well. First of all, the required capital investments might demand a large share of the financial resources of the local utility, if there are any available. Furthermore, people are known to use water with less care as its availability increases, resulting in more spilling of water. At last, the smart water meters add more valuable components to the water system, making it more vulnerable to theft. Theft can be discouraged by placing concrete chambers around the water meters, which is done in Harare, but this will increase the amount of required capital.

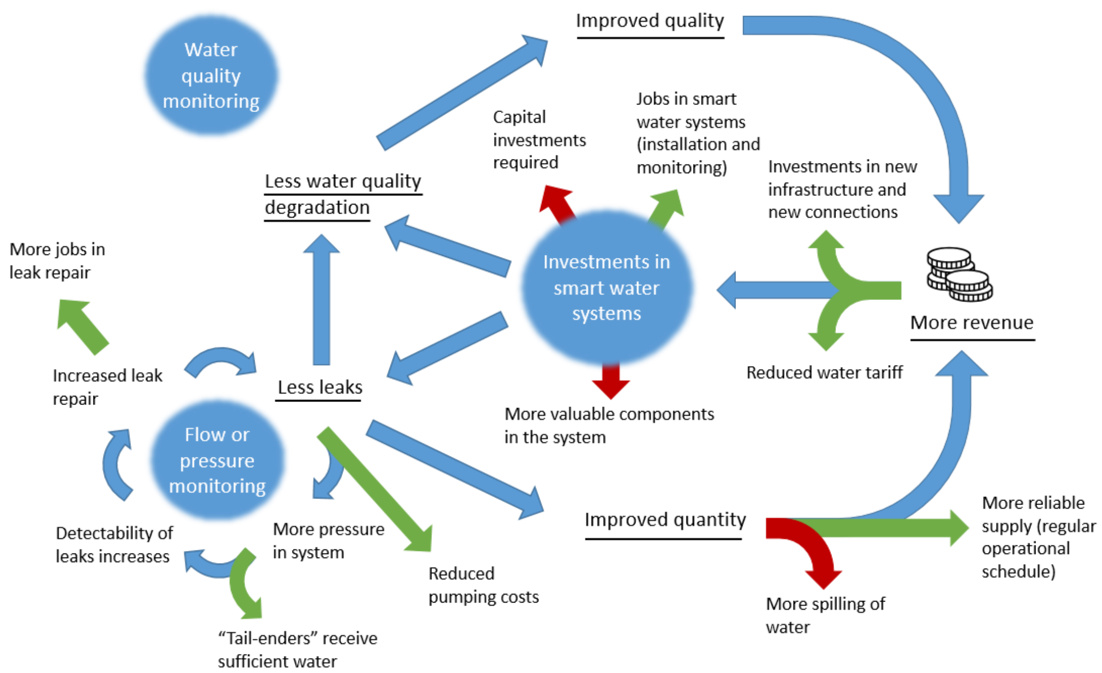

Figure 8.10: Positive and negative associated outcomes of hydraulic monitoring.

As one can see, there are many possible advantages for using a monitoring system to create leak awareness, but also some potential downsides. If well-founded and solid investment choices are made, the effects of using smart water systems to improve water quality and quantity in IWS networks can be very valuable and many more people will have the human right to water.

8.2 Finances

Translating the leak detection efficiency of the monitoring system to potential savings will inquire quite some stochastic assumptions, on top of the stochastic assumptions that have already been made for determining this leak detection efficiency. Therefore, the outcome of this financial analysis should not be judged by its exact numbers. The aim of this financial analysis is merely to give an estimation of the benefits that might be expected and to show the steps that are needed to make an estimation of the financial benefits.

Costs

The costs of the monitoring systems are divided into investment costs and variable costs. This thesis focuses on the situation in Harare and Nairobi, but is instructed and initialized by WorldWaternet. Retrieving the right costs of products or services has proven to be a time-consuming process, often delayed by the large bureaucracy of the utilities of cities with millions of inhabitants. Therefore, the costs for the required materials was asked for in all utilities, resulting in cost data from different places. So, this cost analysis shows an approximation which would change with time and place as these factors influence prices of products, transport fees and exchange rates[5]. The costs were all converted to US Dollars (USD) and can be seen in Table 8.1 and 8.2. Details of the reasoning behind these costs and their sources can be found in appendix A.21. The required measurements at the DMA inlet should be performed by a flow and a pressure meter (or an integrated meter). It was assumed that both meters were installed together, requiring only a single data transmitter, communication software and protection chamber for the measurements at the DMA entrance. In Harare, these devices are already installed, thereby significantly reducing the required investment costs (with a reduction of approximately 6350 USD).

Investment costs	Price (USD)
Measurement devices:	
Flow meter	729 - 846
Pressure meter	770
Supplementary materials (per installed meter):	
Protection chamber	2000
Data transmitter	1350
Connection cables	54
USB and communication software	166
Extra costs:	
Transport costs (excl. protection chamber)	+20%
Labour costs (per installed meter)	140
Total costs	
Total costs per flow meter	**4899 - 5039**
Total costs per pressure meter	**4948**

Table 8.1: Investment costs per flow and pressure meter.

Variable costs	Price (USD)
Maintenance costs:	
Operation online platform (per year)	121
Salary operator (per year)	405
Total maintenance costs per year	**626**
Leak repair	
Material costs (per leak)	101
Labour costs (per leak)	6.32
Fuel costs (per leak)	8
Total costs per leak	**115**

Table 8.2: Variable costs for a monitoring system.

[5] Especially Zimbabwe has experiences with dealing with strongly fluctuating exchange rates, as can be concluded from the fact that the price per cubic of piped water has increased from 20 Zimbabwean Dollar (ZWL) to 90 ZWL during the course of this thesis.

Savings

The previously calculated probability of detecting a leak does not provide any information about whether the leak will occur at all. This information has to be deducted from the historical registration of leaks, which is unfortunately quite poor.

Leak frequency

It was estimated by employees of Harare Water that in Ashdown Park a new leak occurs every two weeks. The sizes of new leaks are however not registered in Harare. Therefore, some assumptions had to be made made. It was assumed that small leaks occur more frequently than large leaks, so a gamma distribution with a maximum value around 2.52[6] was used to describe the spread in leak size, as can be seen in Figure 8.11.

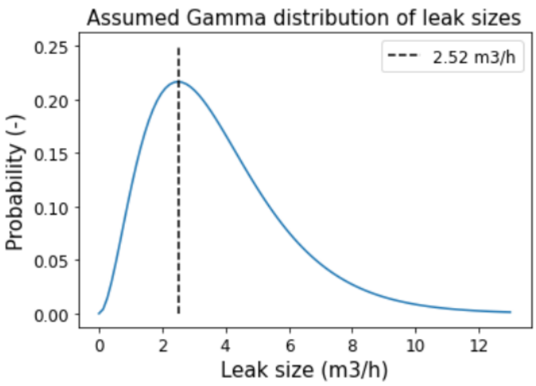

Figure 8.11: Assumed Gamma distribution of leak sizes (The parameter that were used: $\alpha = 3$ and scale factor = 1.25). This distribution is fictional and based on values found in Casillas Ponce et al. (2013)

The return interval ($T_{det,leak,size}$) of detecting a leak of a certain size determines for a large part in which time periods water losses can be prevented and how much money can be saved. This return interval is determined by (1) the probability that a leak of a certain size, or a larger size, emerges in the system (p_{size}; see Figure 8.11), (2) the probability that a leak of a certain size can be detected given the sensor placement ($p_{det,size}$; this follows from the coverage in chapter 7) and (3) the frequency at which new leaks emerge in the area (f_{leak}). This is shown in equation 8.1.

$$T_{det,leak,size} = \frac{1}{p_{size} \times p_{det,size} \times f_{leak}} \qquad \text{,with } f_{leak} = \frac{52}{2} \tag{8.1}$$

The proposed monitoring system with flow sensors and the monitoring system with pressure sensors and a pressure dependent DBM can detect leaks of 9 m3/h. The probability that a newly emerged leak is 9 m3/h or larger (p_{size}) is 0.025, following from the distribution in Figure 8.11. If four sensors are allocated, a monitoring system with flow sensors has a probability of detecting leaks of 0.25 (this equals the coverage, Figure 7.6), whereas the monitoring system with pressure sensors and a pressure dependent DBM has a probability of detecting leaks of 0.083 (Figure 7.14). The resulting $T_{det,leak,size}$ for a

[6]This value was chosen, since leaks that are interesting for a utility to detect were in the range of 2.52 to 10.8 m3/h in research of Casillas Ponce et al. (2013)

monitoring system with flow sensors is 6.2 years, against 18.5 years for a pressure dependent DBM system with pressure sensors.

These large return periods imply that the frequency of detecting leaks and thereby the water savings are very low, seeming to make the monitoring system an unfavorable investment. This is considerably impacted by the low probability at which leaks of 9 m3/h or higher occur (p_{size}). However, one should keep a few things in mind. The alarm values of the DBM systems were set to register the most extreme scenarios from 1000 random demand realizations. This assumption allowed us to pick the optimal location for the sensors. However, these very strict alarm values make that only a leak of 9 m3/h or higher could be detected. It could well be possible that, after the sensor has been installed, field testing and calibration to distinguish alarm values from leaks could result in alarm values that are more moderate. Imagine now that, by choosing more moderate alarm values after field calibration, leaks of 6 m3/h could be detected by the sensors with a similar coverage. This would increase the probability that a leak with this size emerges (p_{size}) to 0.143. Consequently, the return period of detecting leaks with a flow monitoring system would decrease to 1.1 years (3.2 years for the system with a pressure dependent DBM). These reduced return periods will show below that a monitoring system could be a financially beneficial investment *if* the system would be able to detect leaks of 6 m3/h and larger. Unfortunately, this can not be proven with the model that is constructed for this thesis, as the main purpose of this model was to optimally place sensors within the DMA. This could be proven by a pilot system or follow-up laboratory research.

Potential savings

The potential savings can be expressed by the monetary value of the water that is saved due to leak detection. In many NRW-strategies, the monetary value per cubic that is lost due to real leaks is given by the production costs of the water (Ziegler et al., 2011). These strategies do not use the sales price of water, since the water has not yet reached the customer. In Harare the sales price of water was pegged at 90 ZWL per cubic for the first five cubic going gradually up to 110 ZWL per cubic for consumption above 20 cubic per month. Usually, this sales price is set at the costs of production (Magedi MacDonald, 2020). Therefore, a monetary value of 0.28 USD/m3 (90 ZWL/m3) was used to express the value of the water that was saved. In this financial analysis, it is assumed that the leak is repaired immediately after it has been detected.

Figure 8.12 shows the payback periods of investing in a certain number of flow sensors, assuming that are able to detect leaks of 9 m3/h. The costs of the required flow and pressure meter at the DMA inlet is included in all scenarios. This shows that it can take more than 5 years to detect a leak with this system, but once the leak has been detected the savings start to increase quite fast, reaching a payback period of around seven years for the placement of five flow sensors. If it would be possible to decrease the detectable leak size to 6 m3/h, the payback period would reduce to around three years for the placement of both four and five flow sensors (Figure 8.13). If the initial investment would be funded by third parties such as the World Bank, profits can be made immediately after the leak has been found.

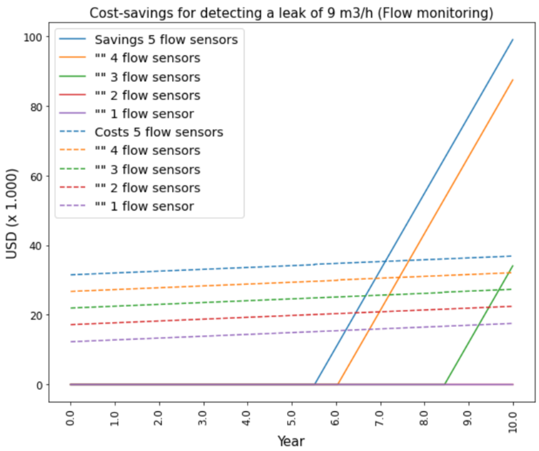

Figure 8.12: The costs and savings that result from detecting leaks of 9 m3/h with flow sensors.

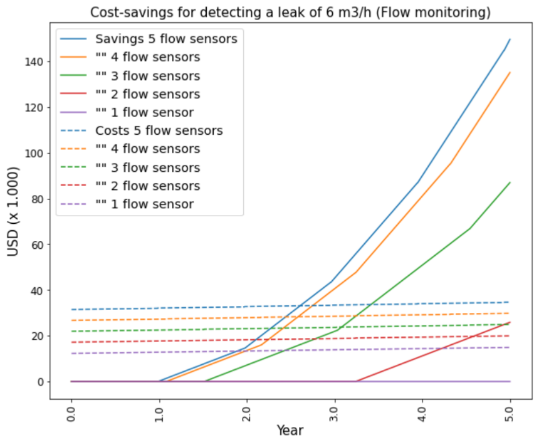

Figure 8.13: The costs and savings that result from detecting leaks of 6 m3/h with flow sensors.

The monitoring system with pressure sensors and a pressure dependent DBM was not able to detect leaks of 9 m3/h within 10 years and is therefore unlikely to be profitable. If its performance could be increased by enabling the system to detect leaks of 6 m3/h and larger, the system would have a payback period of around five years before it becomes profitable. The cohering cost-saving graphs are shown in appendix A.23.

Conclusion financial analysis

As explained in the introduction of this financial analysis, the outcome of this financial analysis should not be judged by its exact numbers. Therefore, more general conclusions shall be mentioned.

First of all, it is important to keep in mind that the costs for a monitoring system have a strong local component as product prices, transport fees and exchange rates differ geographically. Only little price differences were found between flow and pressure meters, although it should be taken into account that the provided financial data was limited. Furthermore, in order to make a proper estimation of the expected savings, it would be beneficial to have substantiated information of the rate at which leaks occur and some idea about the sizes of leaks. This could be implemented within the utility by assigning the leak repair team to register the leaks they repair (this already happens in Harare) and to take a moment before repairing the leak to estimate the leak volume (this does not happen yet). At last it can be concluded that significant savings begin as soon as a new leak is detected. Therefore, the payback time is largely dependent on the performance of the leak detection system. Since the flow monitoring system that was designed in this thesis yields better results in leak detection than the monitoring system with pressure sensors, it also has a lower payback time.

8.3 Business model

Finally, all the information in this chapter was used to construct a business model canvas. A business model canvas is a tool to help understand a business model a straightforward and structured way. It tries to capture the essence of how an organization captures, delivers and maintains the value of a certain product. In this case, it shows how a monitoring system with flow or pressure sensors can create value and would function in the already existing dynamics within a water utility and with potential partnerships. It thereby attempts to give a comprehensive answer to research question 5:

> **Q5:** What are the financial and social benefits of implementing a monitoring system?

The business model canvas, as constructed by Alexander Osterwalder and adapted by the TU Delft to make it more applicable to sustainable business cases (TU Delft, 2020a), was used for this purpose. The full canvas can be read in Figure 8.14. Disadvantages of using this specific canvas are that it is unable to show improvements of the system over time (such as its self-increasing performance, section 8.1) or the main barriers that block areas with IWS from using such a monitoring system (explained in the road map, section 8.1). Below some important or new features shall be shortly highlighted.

The business model canvas focuses around the value proposition, which forms the essence of why people would want to invest in your product. For the monitoring system, the value of the product is threefold and described as:

> *"A hydraulic monitoring system allows the water utility to* ***become aware of leaks*** *in their water system (1) and* ***gives direction*** *to the leaks' location (2),* ***even when the leak is not visible*** *at the surface* ***nor notified*** *by reports of pressure drops or flow deficiencies by local water users (3)."*

A new insight arose when taking a closer look at the customers of the product. These customers are likely to be the board of directors of a local water utility. They will need a strong proof of concept of the product before they make investments, since they carry a large responsibility for the water access of many inhabitants. Therefore, first a pilot project would be advised to substantiate the profitability and use of a hydraulic monitoring system by real data. The World Bank would be an interesting party to fund such a pilot project, since they regularly fund sustainable projects in developing countries, which often have IWS systems.

At last, the need for proper GIS data, the need for a regularly operated supply scheme and the strong cooperation with the leak repair team, show that personnel management will be a key activity that determines whether a monitoring system will be a success. Technology can be a very helpful tool to improve the livelihoods of many, but its success still strongly depends on human effort and cooperation.

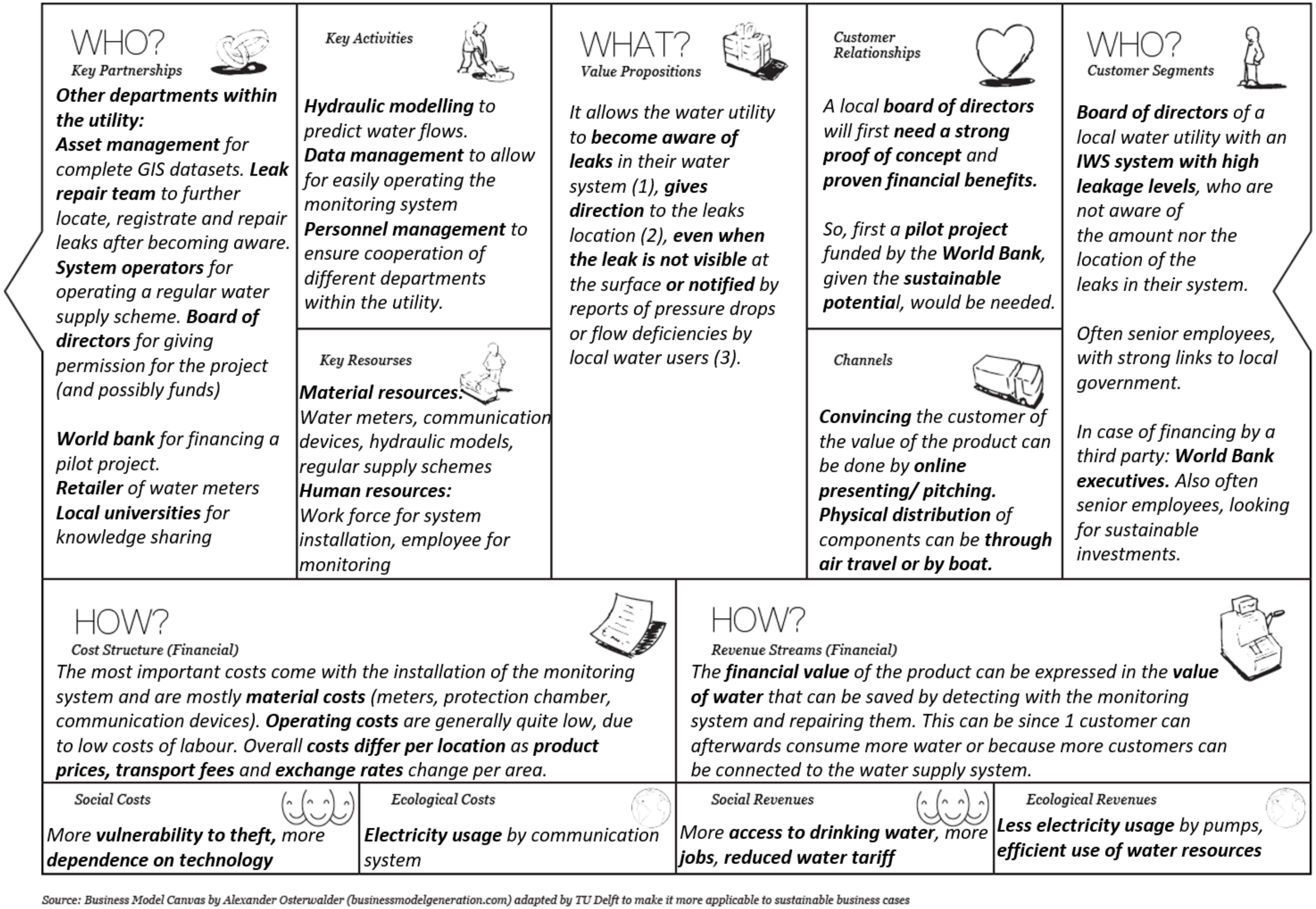

Figure 8.14: Sustainable business model canvas

9. Applicability IWS areas globally

In this chapter the applicability of the monitoring system for global cities with IWS conditions in their drinking water supply system will be considered.

9.1 Priorities of the water utility

Among all of the NRW volumes in the IWA water balance, NCWSC had most trouble in locating and estimating the amount of hidden leaks, according to its local NRW-manager Mugo (2020) at the time of the interview. Therefore, a monitoring system as developed in this thesis could be valuable tool for the utility, as their need for such a system is evident. An interview with a principal engineer in Harare (Shana, 2020), showed that the utility in Harare has trouble with both assessing the amount of visible and hidden leaks. Therefore, this utility might prioritize leak repair in general, above implementing a system which can detect hidden leaks. It is only up to the utility itself if NRW-strategies, and especially the detection of hidden leaks within its system, can be prioritized among all tasks that lie within their responsibility. Without the successful implementation having any priority, it is unlikely that the system is applicable, as its success depends on the cooperation of several departments within the utility.

9.2 Network characteristics

Section 7.3 showed that only a few network characteristics influence the ability of the monitoring system to detect leaks. When implementing a monitoring system with pressure meters, it is advised to place the pressure meters at locations where high pressures occur, giving a slight preference to cities with significant elevation differences. Furthermore, sensors for leak detection seem most effective at the branched parts of a network.

It should be noted that the monitoring system in this thesis is primarily designed for leak detection *within* a DMA, implying the requirement of a utility to have DMAs within their network. A second design option would be to use the monitoring system for leak detection in a water transport system, by applying a hydraulic model using nodes to represent DMAs (suggested in section 5.5). However, this would require the construction of DMAs as well, which has been a struggle so far in Nairobi (Mugo, 2020). There are also only a few DMA's in Harare (Shana, 2020). The earlier mentioned transition from areas in the cities' distribution scheme to DMAs seems a promising solution for this problem, which shall be recommended for further research. A third design option would be to have a monitoring system which has measuring points throughout the entire city, without these points being linked to DMAs. This increases the applicability of the monitoring system world-wide. However, it would hugely increase the computational time of the proposed algorithm if a design would be made for an entire city. Also the calibration method would need more than a single point to accurately calibrate the hydraulic model. Therefore, the sensor placement algorithm as described in this thesis would not work properly for the third option.

9.3 Implementation barriers

Practical barriers that were identified to prevent the transition from IWS systems to smart (hydraulic) IWS systems were: Irregular operational schemes, unknown demand patterns and incomplete GIS data (Figure 8.1). Irregular operational schemes can be linked to the resilience of a water utility, as it requires the utility to deliver water regularly, also at times when water resources are scarce. Unknown demand patterns could be clarified by hydraulic metering itself or by customer surveys. A complete GIS registration is a requirement for the proposed design of the monitoring system, which has so far hindered the design of a monitoring system in Nairobi. At last, financial resources are an often mentioned barrier that could prevent the monitoring system for being a success (Mugo, 2020)(Shana, 2020). So, the system seems more applicable for resilient utilities which operate a regular supply scheme, have an up-to-date GIS system and preferably some financial resources to make an investment. If these conditions are absent, the utility should at least be able to overcome these barriers to make the monitoring system applicable for their supply system. For financial resources, a third party like the World Bank could be approached for the investment.

9.4 Potential associated outcomes

There are many potential benefits of the monitoring system for utilities world-wide, most of which have been mentioned in section 8.1. The financial benefits mainly come from water savings and an increased revenue for the water utility. There are numerous other social benefits which could occur, such as jobs in smart water systems, more finances for investments in new drinking water infrastructure or drinking water resources, reduced pumping costs and a better competitive position against other sources of drinking water. Negative associated outcomes of the monitoring system might be events such as water spilling or an increased vulnerability of the supply system to theft. The effects which will occur depends on the managerial decisions within the local water utility.

9.5 Conclusions chapter 9

This chapter answers the last research question:

> **Q6:** To what extent can the outcomes of this research be used to construct monitoring systems for Water supply networks with IWS conditions around the globe?

The priorities of a local water utility, its network characteristics and the ability of the local utility to overcome barriers that prevent the implementation of a monitoring system, are all factors that determine its applicability in IWS systems around the globe. For example in Harare, the monitoring system might not be a priority, given that other circumstances such as leak repair in its water supply seem more urgent. To design a monitoring system for DMAs throughout the city, cities with IWS conditions will have to work on their digital registration, the regularity of supply and the construction of DMAs. If prioritized, properly installed and operated, the monitoring system could generate significant water savings (and more) and assist local utility with fulfilling their responsibility: supplying people with the basic need of drinking water.

10. Research limitations

This chapter discusses the limitations of this research.

A major short-coming of this thesis was that it was performed in the Netherlands, without having a local impression of how the water supply systems in Nairobi or Harare operate. Especially the assessment on NRW-measuring methods (section 4.2) and the business model (section 8.3), needed a well-provided image of the local situation. This image was very dependent on the limited conversations with local employees through online meetings. Especially the interview with Mugo (2020) was very important for the direction of this thesis. The international guidelines of the IWA, many adopted from Ziegler et al. (2011), were a very helpful tool to have a structured interview and minimize the subjectivity.

Furthermore, the analyses from the flow and pressure data of the DMA's in Harare showed interesting patterns and base demands, underlining the observation of De Marchis et al. (2010) that demands in IWS systems are indeed different. The reasons behind these patterns and base demands could unfortunately only be guessed by reasoning, as it was outside the scope of this thesis to proof direct relations.

The hydraulic model used pressure dependant outflow modelling, as suggested by Vairavamoorthy et al. (2007), but still a significant amount of uncertainty should be taken into account. Most importantly, the hydraulic model was calibrated with the measurements at the DMA entrance as the only calibration point. The challenge of modelling the possible flows *within* the DMA was solved by designing a new method for modelling demand realizations: the tap capacity method. Due to the lack of calibration points within the network, it was not possible to verify whether this method calculates flows within the network accurately. The results of the capacity method were promising at most, as the spread in its results coincided nicely with the spread in inflow data. This limitation of spatial demand and other data is a problem that many researchers have faced during demand allocation (Kanakoudis and Gonelas, 2014). Taking this new dimension of different demand realizations into account in constructing Boolean matrices, is a feature that was not found in other research (Perez et al., 2009)(Khorshidi et al., 2020). This does not guarantee an improved performance, as the performance largely depends on the quality of the underlying hydraulic model. The assumption of a constant flow rate from the tap is also prone to uncertainties, as this flow rate is known to be dependent on local pressures (Marchis and Milici, 2019) and can vary as people do not fully open their tap. The method is simplified a lot by modelling only a single action (opening the tap), whereas developed software like SIMDEUM models processes as well, which would for example take into account the favorite time of day when people would fill their storage (Blokker et al., 2017).

During the construction of the mean flow and pressure, data of all days with continuous supply throughout the year was incorporated. However, the yearly variation in flow into the DMA was not taken account for. Yearly variation can occur since people subtract more water for watering their gardens during the dry period, due to more leaks that result from system deterioration or due to more house connections being added to the distribution network. Furthermore, a clear distinction between the design of a flow monitoring system and a pressure monitoring system was made, by assuming that in Ashdown Park the flow in pipes is independent on the pressure. In this assumption, a few processes have been neglected. Below the nominal pressure, a situation which can not be excluded by only looking at averages, flow in the pipes will be dependent on the pressure (Wagner et al., 1988). Furthermore, the pressure determines leak flow rates (Marchis and Milici, 2019), which on their term influence flows in pipes as well. It is likely that the flow in pipes will therefore to some extent depend on the systems pressure. The clear distinction did, however, provide an overview in the differences in designing a monitoring system for IWS systems with flow sensors and a system with pressure sensors, whereas previous research has often solely focused on pressure sensors for leak detection in continuous supply systems (Perez et al., 2009)(Khorshidi et al., 2020).

To continue, only the situation of a single leak in the DMA was taken into account for the design of the monitoring system and the assessment of its performance, whereas multiple simultaneous leaks are likely to occur. Unfortunately, this thesis does not provide information how multiple simultaneous leaks would affect the hydraulics in the system and the design of the monitoring system. Furthermore, the ranges of "good detectability", provided in chapter 6.3, are based on a personal interpretation of the graphs. Other people can have a different approach or opinion here. Also, all pipes and nodes were incorporated as potential locations to place sensors. However, pipes or nodes which will not be accessible, for instance because they are located in private areas, should be excluded as potential sensor locations. The leak flow rate was chosen to be within a range that was found in literature (Moors, 2016)(Sophocleous et al., 2019)(Casillas Ponce et al., 2013) of which Casillas Ponce et al. (2013) at all was chosen to be most normative, because it has been defined from practical experiences in a utility. However, the range of leak sizes in literature show a large unpredictability, which is increased by the different networks characteristics that occur among utilities. Furthermore, opposite to the previous mentioned research that defined leak size by a flow rate, this thesis' research defined the leak size by area, as its flow rate is dependent on the pressure in the system (Marchis and Milici, 2019). At last, the translation from leaks being found to leaks that will be found once a day (section 4.5), is prone to a lot of uncertainty. This translation makes use of the previously mentioned "good detectability" range and it assumes some unknown relation between the detectability of leaks during sequential time steps.

Despite many practical issues that were mentioned in this research, a missing issue is whether a leak of a certain size will be visible at the surface. If the leak sizes for which the detection system is designed will be very likely to surface, than there is no need for a high-tech detection system, since people passing by are likely to notify the leak to the utility. This issue requires more knowledge about groundwater flows and porous media and was considered to be outside the scope of this thesis. Furthermore, the choice of entrepreneurial contexts which were described (section 8.1), was prone to subjectivity. Another author might have chosen different contexts, resulting in a different view of the situation. The financial analysis (section 8.2) is based on too many assumptions to make a solid financial plan. It therefore merely describes the factors that can determine the financial feasibility of the system and the potential profits. At last, especially world-wide, many more factors than mentioned in chapter 9 play a role to whether the monitoring system will be a success or not.

Overall it could be said that the research in this thesis is prone to quite some uncertainties. Uncertainties in the modelling of water supply systems can never be avoided, as it attempts to model human behaviour, which will always have a random component. Some of these uncertainties were common and known within the field of hydraulic modelling, such as the uncertainty in spatial demand allocation. Other uncertainties arose from attempting to develop new methods, such as the uncertainty of which flow rate would flow out of the tap. The uncertainty in this thesis can be best split in two types: 1) Uncertainties due to a lack of understanding of the local situation, for instance by missing out on very important contexts for entrepreneurship. 2) Uncertainties due to assumptions that underlie the analyses and hydraulic model, such as assuming a certain distribution of leak sizes. Both uncertainties can be reduced by continuing with further research into hydraulic monitoring systems, but to really understand the local situation, it is much advised to conduct this further research in the country of interest.

11. Overall conclusions and recommendations

The chapter shows the conclusions that can be drawn from this thesis' research and ends with recommendations for further research.

11.1 Overall conclusions

This thesis addresses the problem of the designing a monitoring system for the detection of hidden leaks in IWS areas. In doing so, it proposes new methods in the fields of hydraulic modelling in IWS areas and optimal sensor placement. Furthermore, it shows the business opportunities that could arise from such a system. Whereas the final sections of previous chapters show the answers to the research questions[1], this section summarizes and combines these answers to show the overall conclusions.

A hydraulic model was created for a DMA with IWS in Harare (Ashdown Park) as a case study, by using pressure dependant outflow modelling. The daily demand patterns found in several IWS systems in Harare showed a different pattern than for continuous supplied networks, showing less strong peaks. This can be due to a constant water demand for filling storage, leaks in the system or different consumer behaviour. Furthermore, a method was developed to calibrate the hydraulic model using flow and pressure measurements at the entrance of the DMA. This method showed that the complicated relations between demand and pressure in the model can be approached by using assuming linear relationships between the demands and inlet pressure on one side and the pressure at a specific node and flow at a specific pipe on the other. This first order approximation showed a total Normalized Root Mean Squared Error (NRMSE) between the output of the model and the measurements of only 0.019, when a nominal pressure of 12m was used. A higher nominal pressure (20m) increases the non-linearity of the previous mentioned relationship, but the method still achieved an NRMSE between the flow and pressure output and the cohering measurements of below 0.05. As well, it was noted that the inlet pressure fluctuates considerably over the week. This is mainly caused by the pressure in the main distribution system and therefore by the supply scheme of the utility. Since the pressure in Ashdown Park did not drop below its estimated nominal pressure, this fluctuation is unlikely to influence the total demand in the DMA.

With such a hydraulic model, an algorithm was developed to optimally place sensors within the DMA for leak detection and estimate the performance of the monitoring system. The proposed novel algorithm modelled different demand realizations by allocating demand among the nodes with a random weighted choice and a single tap capacity, which showed promising results for modelling flows that occur in Ashdown Park, especially at

[1]The research questions are answered in detail in the following sections: Q1: section 1.1, Q2: section 5.5, Q3: section 6.4, Q4: section 7.5, Q5: section 8.3, Q6: section 9.5

night. These results were promising, since the spread of the modelled inflow was well comparable to the spread of the inflow measurements. This tap capacity is especially suitable for IWS areas, since people in IWS areas usually only have one tap directly connected to the water supply system and water end-use devices are not directly connected to the network. Whereas other research has often focused on a single demand configuration (Perez et al., 2009)(Khorshidi et al., 2020), the proposed algorithm takes account for the influence of different demand realizations on the detection of leaks.

Furthermore, it was found that the weekly flow pattern in Ashdown Park on days with continuous supply was more constant than the weekly pressure pattern. Therefore, it can be concluded that the water use behaviour of inhabitants in Ashdown Park has been more constant than the supply behaviour of the water utility in the concerned area. This irregular water supply behaviour of the water utility significantly increases the difficulty of designing a monitoring system with pressure sensors. Since the demand of the people seems not limited by an insufficient pressure during days with supply, using a flow monitoring system to detect leaks showed a better performance (25% coverage of the DMA with four optimally allocated flow sensors) in Ashdown Park than using a monitoring system with pressure sensors (1% coverage with four optimally allocated pressure sensors). Making the monitoring system with pressure sensors dependent on the inlet pressure increased its performance (from 1% to 8.3%). Important factors that influence the performance of the monitoring system are the time of day, the leak size that occurs and the tap capacity at the household connections. Furthermore, the branched part of the system was generally more favorable to place sensors for leak detection. Nodes with high pressures were favorable when designing a monitoring system with pressure sensors. Additionally, the design of the monitoring system should start with devices that measure the flow and pressure at the DMA entrance, as this data is crucial for constructing and calibrating the hydraulic model.

The business model and financial analysis showed more practical challenges and opportunities that could arise from implementing the monitoring system. Main barriers that currently prevent many IWS systems from moving towards a system with smart meters are its irregular supply schemes, unknown demand patterns and incomplete GIS data. The existence of these barriers is understandable when taking into account the enormous historical growth and limited financial resources of Nairobi and Harare, compared to a city with continuous supply such as Amsterdam. The financial analysis showed that economic savings from a monitoring system can be made as soon as the leak is detected and repaired. Having an improved leak registration, that measures its frequency of occurrence and sizes, would improve the certainty of financial predictions. Important factors that determine the world-wide applicability of the monitoring system are the priorities of a local water utility, the technical performance of the monitoring system within the local water network and the ability of the local utility to overcome barriers that prevent the implementation of a monitoring system. If prioritized, properly installed and operated, the monitoring system could generate significant water savings and assist local utility with fulfilling their responsibility: supplying people with the basic need of drinking water.

11.2 Recommendations for further research

Further research should focus on intermittent supply conditions in large cities in sub-Saharan Africa as the situation to be and develop methods that specifically suit the situation. A very interesting case would be to see whether the areas that are mentioned in the water distribution scheme, that is used by utilities to show its consumers how the water is weekly divided among different areas, can be transformed to DMA's. This transition seems easily possible in theory, since the district areas are already formed to divide supply, but they are not metered yet. These DMA's can form a basis for the future development of hydraulic monitoring systems, as having DMA's is very beneficial for the design of such a system (explained in chapter 9.2).

In the field of hydraulic modelling, it would be interesting to see if the behaviour and demand patterns in IWS can be further researched. This could result in standard demand patterns which could be used or software which can accurately model different demand realizations. The demand behaviour which mainly occurs in IWS is from people filling their storage, which seems less difficult to implement in a model than when water is used for multiple purposes. If software like SIMDEUM can be designed for the continuous supply case, it seems that similar software could also be designed for IWS areas. A designer could expand the tap capacity method and add processes and actions, or could choose approach the problem from a different direction. Furthermore, it would be interested to see if a method can be developed that predicts demands based on mathematical relations and multiple calibration points, as attempted in this thesis. This would decrease the need for house to house surveys, but also adds the risk of making assumptions based on pure data instead of basing them on understanding the local situation.

Furthermore, it is definitely recommended to start a pilot project for a hydraulic monitoring system in an IWS area. If a pilot project would be developed, it would be recommended to draw a relation between the size of a leak, the soil conditions and whether it surfaces. This would give very valuable information about whether the monitoring system will be a helpful or unneeded tool. Another interesting feature of a pilot project would be whether the alarm values predicted by this research could be relaxed to find smaller leak sizes. At last, attempting a pilot project at a certain location would likely find more practical barriers and limitations than could be thought off in this thesis, therefore bringing more valuable information for the implementation of a hydraulic monitoring system. It would be a very good first step for moving towards smart IWS systems.

Bibliography

Al-Washali, T., Sharma, S., Al-Nozaily, F., Haidera, M., and Kennedy, M. (2019). Monitoring nonrevenue water performance in intermittent supply. *Water*.

Anderson, D. (2005). *Histories Of The Hanged: The Dirty War In Kenya And The End Of An Empire*.

Andey, S. P. and Kelkar, P. S. (2007). Performance of water distribution systems during intermittent versus continuous water supply. *Journal AWWA*, 99(8):99–106.

Blokker, E., Agudelo-Vera, C., Moerman, A., Thienen, P., and Pieterse-Quirijns, I. (2017). Review of applications of simdeum, a stochastic drinking water demand model with small temporal and spatial scale. *Drinking Water Engineering and Science Discussions*, 10:1–15.

Blokker, M. (2020). personal communication. Developer SIMDEUM. Date of contact: 19-11-2020.

Boyd, S. and Vandenberghe, L. (2018). *Introduction to Applied Linear Algebra*. Cambridge University Press.

Casillas Ponce, M. V., Garza Castañón, L. E., and Cayuela, V. P. (2013). Model-based leak detection and location in water distribution networks considering an extended-horizon analysis of pressure sensitivities. *Journal of Hydroinformatics*, 16(3):649–670.

City of Harare (2019). Harare bulk water block diagram.

Clement, A. and de Groot, R. (2020). Ontwerpcriteria drinkwater distributienetten. policy document Waternet.

Colombo, A. and Karney, B. (2005). Impacts of leaks on energy consumption in pumped systems with storage. *Journal of Water Resources Planning and Management*, 131:146–155.

Crowl, D. and Louvar, J. (2011). *Chemical Process Safety: Fundamentals with Applications*. Upper Saddle River, NJ: Prentice Hall, 3 edition.

de Groot, R. (2020). personal communication. Employee of Waternet. Regular contact moments during the course of this thesis (May 2020 - January 2021).

De Marchis, M., Fontanazza, C. M., Freni, G., Loggia, G. L., Napoli, E., and Notaro, V. (2010). A model of the filling process of an intermittent distribution network. *Urban Water Journal*, 7(6):321–333.

DeSilva, D., Moglia, M., Johnson, K., Davis, P., Gould, S., and Burn, S. (2005). Estimation of unavoidable leakage in a network through measurements in small zones. *Pipes Wagga Wagga*.

Dos Santos, S., Adams, E., Neville, G., Wada, Y., de Sherbinin, A., Mullin Bernhardt, E., and Adamo, S. (2017). Urban growth and water access in sub-saharan africa: Progress, challenges, and emerging research directions. *Science of The Total Environment*, 607-608:497 – 508.

Edmond, C. (2020). These are the top risks facing the world in 2020. *Part of the World Economic Forum Annual Meeting.*

Elala, D., Labhasetwar, P., and Tyrrel, S. (2011). Deterioration in water quality from supply chain to household and appropriate storage in the context of intermittent water supplies. *Water Science Technology: Water Supply*, 11:400.

Elger, D., Williams, B., Crowe, C., and Roberson, J. (2014). *Engineering Fluid Mechanics 10th edition.* WILEY.

Ernst Vink (2020). personal communication. Employee of Waternet, specialized in smart water meters. Date of contact: 09-12-2020.

Farley, M. (2001). Leakage management and control. *World Health Organization (WHO).*

Gurobi Optimization LLC (2020). *Gurobi Optimizer Reference Manual.* Gurobi Optimization, LLC.

Haestad Methods, Walski, T. M., Chase, D. V., Savic, D. A., Grayman, W., Beckwith, S., and Koelle, E. (2003). *Advanced Water Distribution Modeling and Management.* First edition edition.

Ilaya-Ayza, A. E., Campbell, E., Pérez-García, R., and Izquierdo, J. (2016). Network capacity assessment and increase in systems with intermittent water supply. *Water*, 8.

Inhabitant Budiriro (2020). personal communication. Date of contact: 04-12-2020.

Jacobs, H., Skibbe, Y., Booysen, M., and Makwiza, C. (2015). Correlating sound and flow rate at a tap. *Procedia Engineering*, 119:864 – 873. Computing and Control for the Water Industry (CCWI2015) Sharing the best practice in water management.

Jang, D. and Choi, G. (2017). Estimation of non-revenue water ratio using mra and ann in water distribution networks. *Water*, 10.

KAM (2020). Nairobi water and sewerage company seeks to raise kes 12 billion to expand sewerage coverage in the city. http://kam.co.ke/nairobi-water-and-sewerage-company-seeks-to-raise-kes-12-billion-to-expand-sewerage-coverage-in-the-city/. Published by Kenya Association of Manufacturers (KAM), accessed at 06-01-2021.

Kanakoudis, V. and Gonelas, K. (2014). Accurate water demand spatial allocation for water networks modelling using a new approach. *Urban Water Journal.*

Khorshidi, M. S., Nikoo, M. R., Taravatrooy, N., Sadegh, M., Al-Wardy, M., and Al-Rawas, G. A. (2020). Pressure sensor placement in water distribution networks for leak detection using a hybrid information-entropy approach. *Information Sciences*, 516:56 – 71.

Kober, E. and Gangl, G. (2009). New monitoring methodology for water distribution systems. RBS wave GmbH, Stuttgart, Germany.

Kumpel, E. and Nelson, K. L. (2013). Comparing microbial water quality in an intermittent and continuous piped water supply. *Water Research*, 47(14):5176 – 5188.

Kumpel, E., Woelfle-Erskine, C., Ray, I., and Nelson, K. L. (2017). Measuring household consumption and waste in unmetered, intermittent piped water systems. *Water Resources Research*, 53(1):302–315.

Kureva, T. and Moors, J. (2019). Results h2h-survey ashdown park. Technical report, VEI.

Lambert, A. and Hirner, W. (2000). Losses from water supply systems: Standard terminology and recommended performance measures. *International Water Association (IWA)*.

Lambert, A. O., Brown, T. G., Takizawa, M., and Weimer, D. (1999). A review of performance indicators for real losses from water supply systems. 48:227–237.

Magedi MacDonald (2020). personal communication. Employee of Harare Water. Date of contact: 30-10-2020.

Marchis, M. D. and Milici, B. (2019). Leakage estimation in water distribution network: Effect of the shape and size cracks. *Water Resource Management*, 33:1167–1183.

Mohapatra, S., Sargaonkar, A., and Labhasetwar, P. K. (2014). Distribution network assessment using epanet for intermittent and continuous water supply. *Water Resources Management*, 28:3745–3759.

Moors, J. (2016). Model-based leak localization in small water supply networks. Master's thesis, TU Delft.

Moors, J. (2020). personal communication. Employee of VEI. Regular contact moments during the course of this thesis (May 2020 - January 2021).

Morrison, J., Tooms, S., and Rogers, D. (2007). District metered areas guidance notes. *Internation Water Association (IWA)*.

MSc-Student University of Bulawayo (2020). personal communication. MSc. student Civil and Water Engineering, National University Of Science And Technology in Bulawayo. Date of contact: 03-12-2020.

Mugo, E. (2020). personal communication. Employee of NCWSC. Date of contact: 22-05-2020.

Muzenda, F., Masocha, M., and Misi, S. N. (2019). Groundwater quality assessment using a water quality index and gis: A case of ushewokunze settlement, harare, zimbabwe. *Physics and Chemistry of the Earth, Parts A/B/C*, 112:134 – 140. 18th WaterNet/WARFSA/GWPSA Symposium on Integrated Water Resources Development and Management: Innovative Technological Advances for Water Security in Eastern and Southern Africa - Part B.

Nation (2020). State takes over six city water systems. https://nation.africa/kenya/counties/nairobi/state-takes-over-six-city-water-systems-1762100. Accessed at 09-09-2020.

Ndegwa, L. M. (2016). The impact of non-revenue water on water utilities and their customers: a case study of nairobi city water and sewerage company. Bachelor thesis at the University of Nairobi.

NTESS (2019). *WNTR: Read the Docs*. Full name of the company: National Technology and Engineering Solutions of Sandia, LLC.

Nyirenda-Jere, T. and Biru, T. (2015). Internet development and internet governance in africa. *Internet Society*.

Oduro-Kwarteng, S., Nyarko, K., Odai, S., and Aboagye-Sarfo, P. (2009). Water conservation potential in educational institutions in developing countries: Case study of a university campus in ghana. *Urban Water Journal - URBAN WATER J*, 6:449–455.

Perez, R., Puig, V., Pascual, J., Peralta, A., Landeros, E., and Jordanas, L. (2009). Pressure sensor distribution for leak detection in barcelona water distribution network. *Water Supply*, 9:715–721.

Preston, S. J. and Sturm, R. (2002). Use of the infrastructure leakage index (ili) in malaysia. In *IWA Managing Leakage Conference*.

Republic of Zimbabwe (2017). Harare water network hydraulic modelling report (rev 2). Technical report.

Rossman, L. (2000). *EPANET 2 USERS MANUAL*. National Risk Management Research Laboratory.

Sakomoto, T., Lutaaya, M., and Abraham, E. (2020). Managing water quality in intermittent supply systems. the case of mukono town, uganda. *Water, 12(3)*.

Schutter, O. D. (2011). How not to think of land-grabbing: three critiques of large-scale investments in farmland. *The Journal of Peasant Studies*, 38(2):249–279.

Shana, V. (2020). personal communication. Principal engineer Harare Water Department. Regular contact moments during the course of this thesis (May 2020 - January 2021).

Soldevila, A., Blesa, J., Fernandez-Canti, R. M., Tornil-Sin, S., and Puig, V. (2019). Data-driven approach for leak localization in water distribution networks using pressure sensors and spatial interpolation. *Water*, 11.

Sophocleous, S., Savic, D., and Kapelan, Z. (2019). Leak localization in a real water distribution network based on search-space reduction. *Journal of Water Resources Planning and Management*, 145(7).

The New York Times (2019). Nairobi's gold-loving governor is arrested on corruption charges. https://www.nytimes.com/2019/12/06/world/africa/nairobi-governor-mike-mbuvi-sonko.html. Accessed at 09-09-2020.

TU Delft (2019). Network calculations. Lecture slides CIE5550 Pumping Stations and Transport Pipelines.

TU Delft (2020a). Business model canvas. Course document for CT3411-16 Waterworks in Practice.

TU Delft (2020b). Design of water distribution networks. Reading material CIE5550 Pumping Stations and Transport Pipelines.

UN-Habitat (2012). Reduction of illegal water. Principal Authors: UN-Habitat Lake Victoria Water and Sanitation Initiative team and National Water and Sewerage Corporation team.

Vairavamoorthy, K., Akinpelu, E., Lin, Z., and Ali, M. (2001). Design of sustainable water distribution systems in developing countries. *Water Development Research Unit, Faculty of the Built Environment South Bank University.*

Vairavamoorthy, K., Gorantiwar, S. D., and Mohan, S. (2007). Intermittent water supply under water scarcity situations. *Water International*, 32(1):121–132.

Van Lagen, G. (2020). A bayesian approach for active fault isolation with an application to leakage localization in water distribution networks. Master's thesis, Delft University of Technology.

Van Steen, J. (2020). Robustness assessment of leak detection and localization in water distribution networks under stochastic water demand. Master's thesis, Delft University of Technology.

Van Vossen-van den Berg, J. (2017). Overzicht en toepassing van lekopsporingstechnieken. KWR 2017.003.

VEI (2020). Nrw master class session 3. Lecture slides NRW Master Class Vitens Evides International (VEI).

Wagner, J., Shamir, U., and Marks, D. (1988). Water distribution reliability: Simulation methods. *Water Resources Planning Management*, 114:276–294.

Waternet (2020). Processed demand data waternet 2020.

World Bank (2020). Shared data of the world bank. https://data.worldbank.org/. Accessed at 07-12-2020.

Yassin, M. M., Amr, S. S. A., and Al-Najar, H. M. (2006). Assessment of microbiological water quality and its relation to human health in gaza governorate, gaza strip. *Public Health*, 120(12):1177 – 1187.

Ziegler, D., Sorg, F., Fallis, P., Hübschen, K., Happich, L., Baader, J., Trujillo, R., Mutz, D., Oertlé, E., Klingel, P., and Knobloch, A. (2011). Guidelines for water loss reduction. Technical report, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH, VAG Armaturen GmbH (VAG), Fachhochschule Nordwestschweiz (FHNW) and Karlsruhe Institute of Technology (KIT).

A. Appendix

A.1 Interview NRW-manager NCWSC (eng. Mugo)

An interview with Ephantus Mugo, the local NRW-manager of NCWSC, was held to identify which information is most valuable to monitor in Nairobi. The interview was structured according to the methodology mentioned in section 4.2. The results and conclusions from this interview are noted below. The results are structured according to the different volumes that are mentioned in the IWA water balance (Table A.1).

<table>
<tr><th>Level 0</th><th>Level 1</th><th>Level 2</th><th>Level 3</th><th>Type of water</th></tr>
<tr><td rowspan="10">System input volume Q_{in}</td><td rowspan="5">Authorised consumption Q_A</td><td rowspan="3">Billed authorised consumption Q_{BA}</td><td>Billed water exported</td><td rowspan="3">Revenue water</td></tr>
<tr><td>Billed metered consumption</td></tr>
<tr><td>Billed unmetered consumption</td></tr>
<tr><td rowspan="2">Unbilled authorised consumption Q_{UA}</td><td>Unbilled metered consumption</td><td rowspan="7">Non-revenue water</td></tr>
<tr><td>Unbilled unmetered consumption</td></tr>
<tr><td rowspan="5">Water losses Q_L</td><td rowspan="2">Apparent losses Q_{AL}</td><td>Unauthorised consumption</td></tr>
<tr><td>Customer meter inaccuracies and data handling errors</td></tr>
<tr><td rowspan="3">Real losses Q_{RL}</td><td>Leakage on transmission and distribution mains</td></tr>
<tr><td>Leakage and overflow at storage tanks</td></tr>
<tr><td>Leakage on service connections up to point of customer meter</td></tr>
</table>

Table A.1: IWA water balance with column headings

System input volume (level 0)

Nairobi is divided in seven regions. They are currently constructing 53 DMA's. The NRW-team considers the 53 DMA's as primary DMA's and the seven regions as secondary DMA's. They are busy with installing regional bulk meters to measure the inflow into all the DMA's. The water flow is measured with electromagnetic flow meters when it is leaving the drinking water production site, which is about 25km outside the city. From here the water flows to the reservoirs. The flow going from the reservoirs into the distribution system is measured with ultrasonic flow meters (Mugo, 2020).

Billed authorised consumption and unbilled authorised consumption (level 2)

There were no volumes of exported billed water found during the course of this thesis, but this should be well-known within NCWSC. It is expected to be low, since they already face water shortage within the county. The metered consumption is determined every month by the meter readers. They take a picture of the new meter reading and send it to the company, where it is registered. Customers whose meters are broken are billed based on an estimation from meter readings in the past. Both billed metered and billed unmetered consumption are therefore highly dependant on flow measurements at household level. The goal of NCWSC is to have no broken meters in the field.

The fire hydrants in Nairobi are used for fire fighting and to fill water lorries. Water lorries are used to transport water to places within the country that have limited access to water and to sell it there. The volume that is taken out of the fire hydrant is measured by in-pipe flow meters. The volume that is used by the lorries is registered and added to the billed metered consumption. The remaining volume that is subtracted from the fire hydrant is used for fire fighting and forms the unbilled metered consumption. Therefore, unbilled metered consumption depends on flow measurement within the fire hydrants. No examples of unbilled unmetered consumption could be identified (Mugo, 2020).

Apparent losses

Meter inaccuracy is a problem in intermittent supply areas. Usually a mechanical pressure meter needs a certain pressure to measure volumes flowing through. When this pressure is not available the meter does not register. In intermittent supply areas periods with water often come with low pressures, since everyone opens their taps when there is water. Therefore, there is a lot of under registration. The customer meters are tested on a bench and usually only when there are customer complaints. The NRW-team of NCWSC used to have a portable flow meter, but it broke down. They are currently looking into the possibility to purchase new flow meters.

A few years ago a new system was introduced that minimizes data handling errors. In this new system, meter readers need to make pictures of their readings. Nevertheless, there will always remain some human error in assessing and calculating the data.

Usually, around 5 to 12% of the customers have illegal connections. A recent field study in Nairobi showed that 15% of the customers did not have any meters. This resulted from a house-to-house study, which is performed every now and then in Nairobi and takes about four months to finish (Mugo, 2020).[1]

Real losses

The visible leaks in the transmission and distribution mains are repaired by a repair team. However, the volumes of water lost are not measured during repair. The NRW-team was thinking about implementing this by using graduated containers to measure this volume. According to the NRW-manager, nightmares of the drinking water supply network are mainly the hidden leaks. The pipes in the centre of Nairobi are sometimes almost 100 years old and are perforated quite a lot. The amount of hidden leaks in this part should be considerable. Using a minimum night flow analysis to estimate the volume of hidden

[1] A paradoxical situation occurs if the utility start measuring unauthorised consumption. If the utility would measure these volumes, but take no action, it would be labelled as "authorised unbilled consumption". If the utility would say this consumption is unauthorised and take action, these volumes would be reduced by their actions.

leaks is difficult, since people are filling the storage tanks during the night due to the intermittency of supply. This makes it difficult to estimate a minimum night consumption. Some people even have underground water tanks and use their own pump to pump the water up to their houses. It is therefore a challenge to create awareness of hidden leaks. Continuous pressure, flow or noise monitoring is not yet a proven practice in intermittent supply systems. When there is awareness of a hidden leak, this is usually through customer complaining about having no water when they should receive water according to the schedule. If this happens, NCWSC has a team that can detect leakages. They use ground microphones to locate the leak or use signalling through metallic pipes. However, both methods require the supply network to be under pressure, which is not always the case during intermittent supply. Background leakages are difficult to estimate, since often the pressure in the system is very low or unknown.

NCWSC is entirely responsible for the drinking water infrastructure up to the customer meters. Therefore, they go to the client and repair the pipes when customers complain about leakages. They also repair leaks after the meter, since the customer has already paid for the water but it is still lost. When these repairs take place, the volume lost due to these leakages is usually not measured.

Leakages in reservoirs have also been an area of concern. Currently, there are quite some losses due to overflow in the reservoirs. Therefore, the NRW-team is thinking about installing level gauges at the reservoirs to measure the overflow. The leakage within the reservoir can be repaired during the maintenance schedule. This happens once every few years, but there is no regular schedule. Usually, the reservoirs are cleaned if there is sludge in the reservoir. This can be seen when the water reaches a very low level, which happens regularly. If a leak is spotted during this cleaning process, it is repaired. When leaks in reservoirs are repaired, they are also not measured Mugo (2020).

Conclusion on important information in Nairobi

As far as authorized consumption is considered, it can be concluded that flow meters at customer level and within fire hydrants are essential to measure volumes. Furthermore clear regulations about which volumes to bill and which volumes not to bill will make measuring or estimating the authorized consumption more easy.

For the assessment of apparent losses, it is quite well known within the utility which methods to use. However, a few recommendations for improvements can be made. In the case of meter inaccuracy, the estimation of the inaccuracy can be improved by grouping meters by a certain size, age and brand. This way, the found inaccuracy can be more presentable for the meters within the area than when only broken meters are tested. When a portable flow meter would be purchased, this will also help in determining the meter inaccuracy. Getting data from a survey is a methodology that has been previously applied in Nairobi and can therefore be considered as suitable. Most of the data that is required to make these methods work is already available.

The real losses are a challenge for NCWSC. Visible leaks are reported by customers and repaired at different locations in the system (transmission and distribution mains, up to the customer meter and at the reservoirs). However, no measurements or estimations about the lost volumes are performed. Hidden leaks are only repaired when awareness is raised due to customer complaints. A big question for NCWSC is how to become aware of the leakages before there are complaints, since the regular practice of minimum night flow does not work. After they are aware of the leak, there is a repair team which will further locate and repair the leak. So, it can be concluded that the largest need in terms of assessing real losses would be to create awareness of hidden leaks. Furthermore, a general idea about the pressure in the area is needed to estimate the background losses and it is advised to do estimations on the volumes lost when repairing visible leaks.

As can be concluded, the information that is lacking the most in the system seems to be the awareness of hidden leaks. Therefore, the main goal of the monitoring system will be to allow the utility to become aware of hidden leaks in the system, before customer complaints occur. Other desirable information is flow data at the customer meter and at fire hydrants to estimate authorized consumption and pressure data to estimate background leakage and get insight into meter inaccuracies that occur due to low pressures.

A.2 Data dependency trees

These data dependency were used as a guiding theme during the interview with the NRW-manager during May 2020 (Mugo, 2020).

Authorized consumption

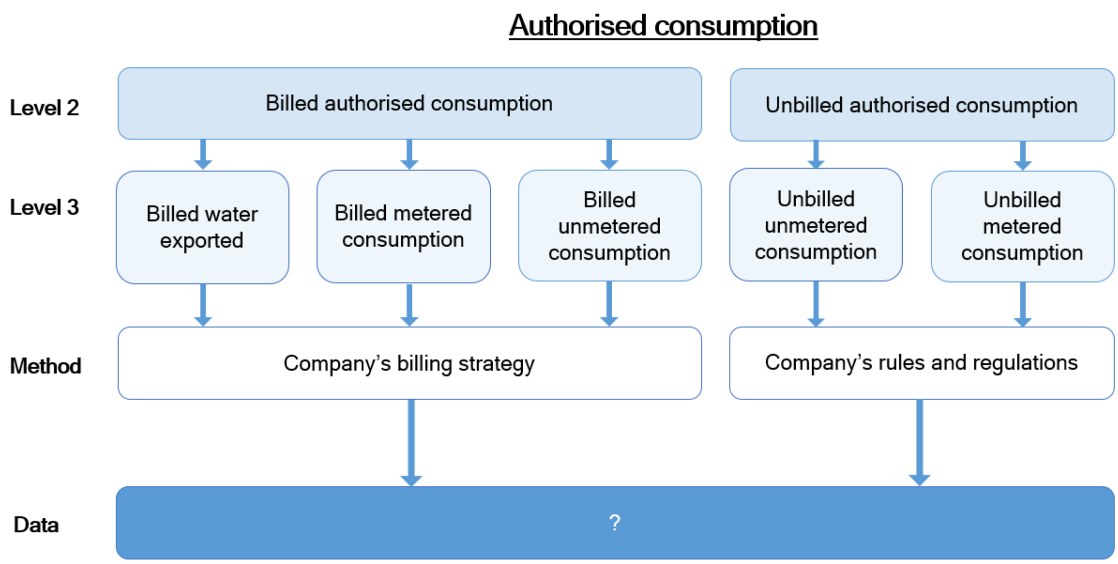

Figure A.1: Authorized consumption

Apparent losses

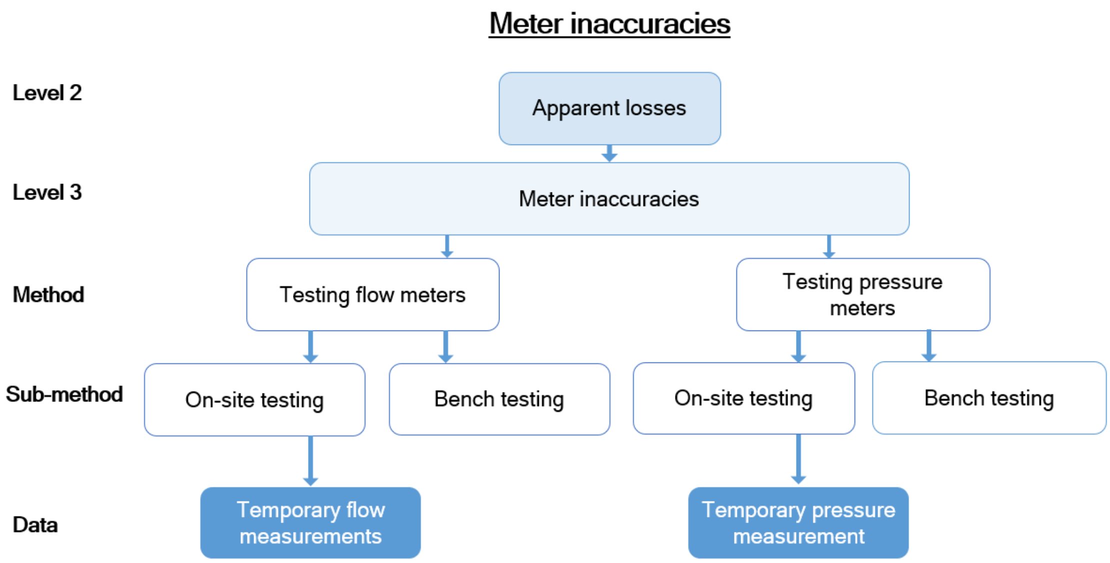

Figure A.2: Meter inaccuracies

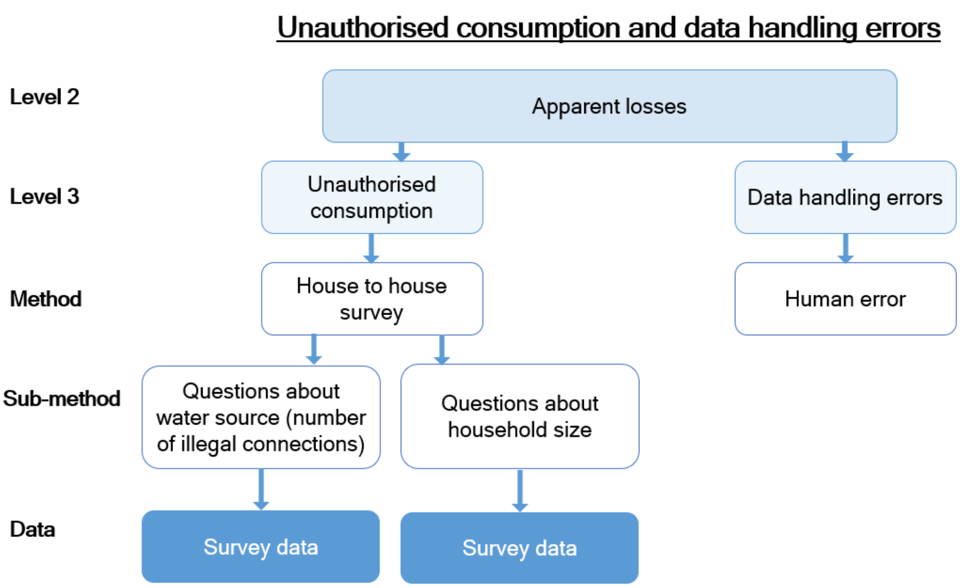

Figure A.3: Unauthorized consumption and data handling errors

Real losses

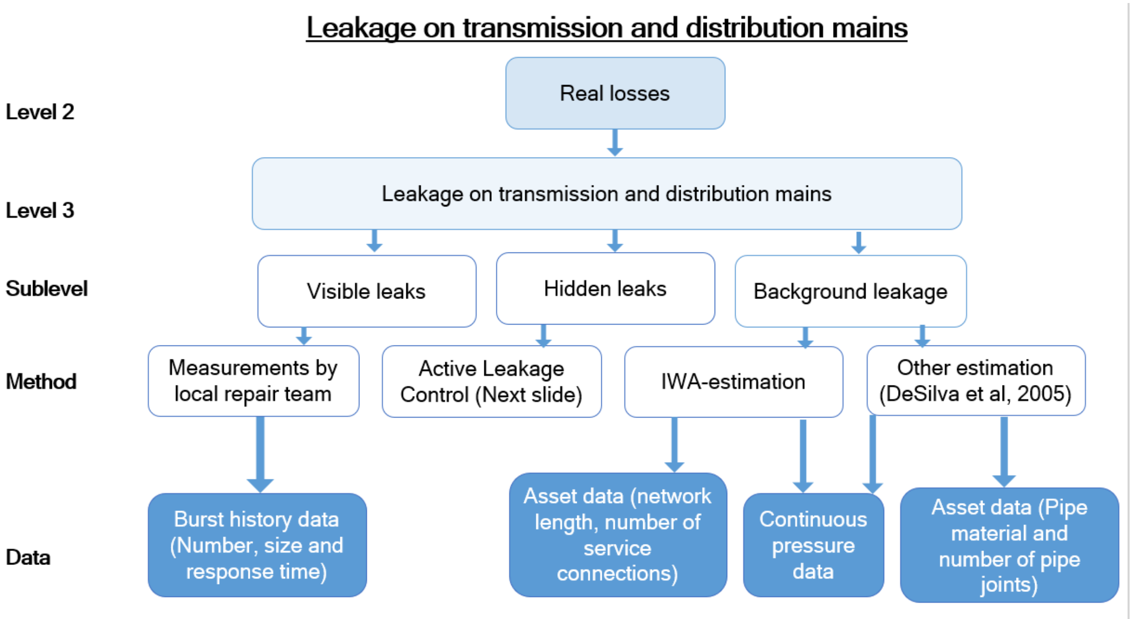

Figure A.4: Leakages at transmission and distribution mains (visible and background leakages)

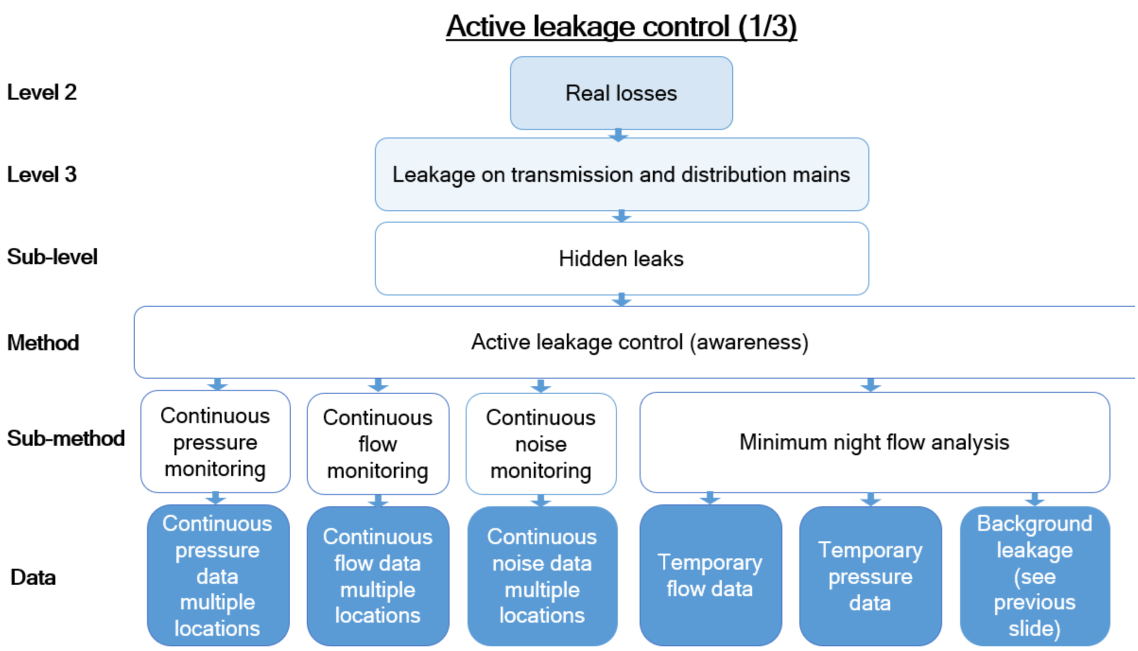

Figure A.5: Active leakage control (1/3)

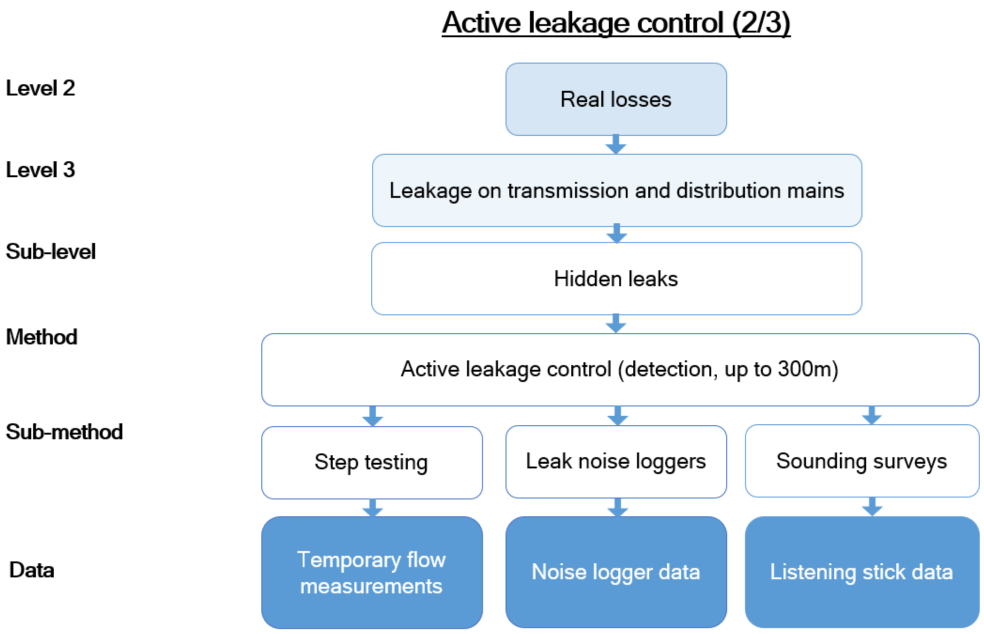

Figure A.6: Active leakage control (2/3)

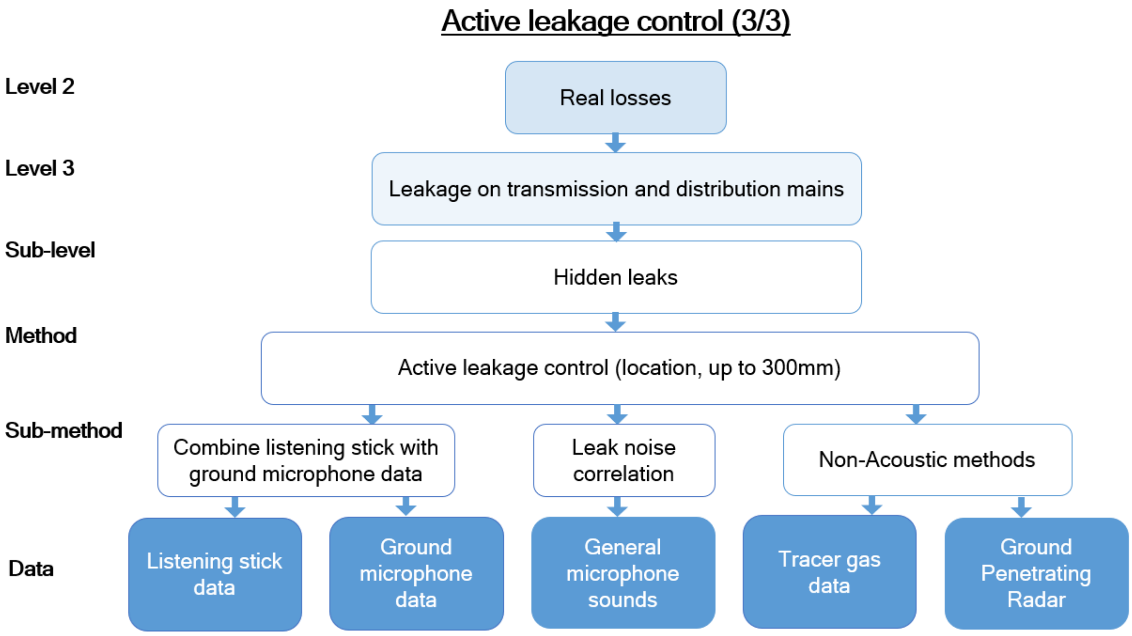

Figure A.7: Active leakage control (3/3)

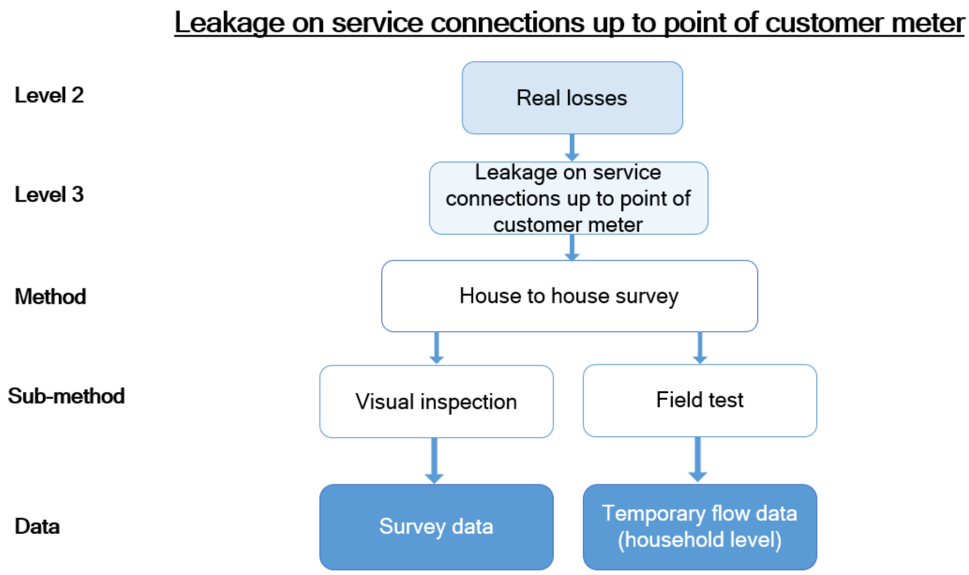

Figure A.8: Leakages on service connections up to point of customer meter

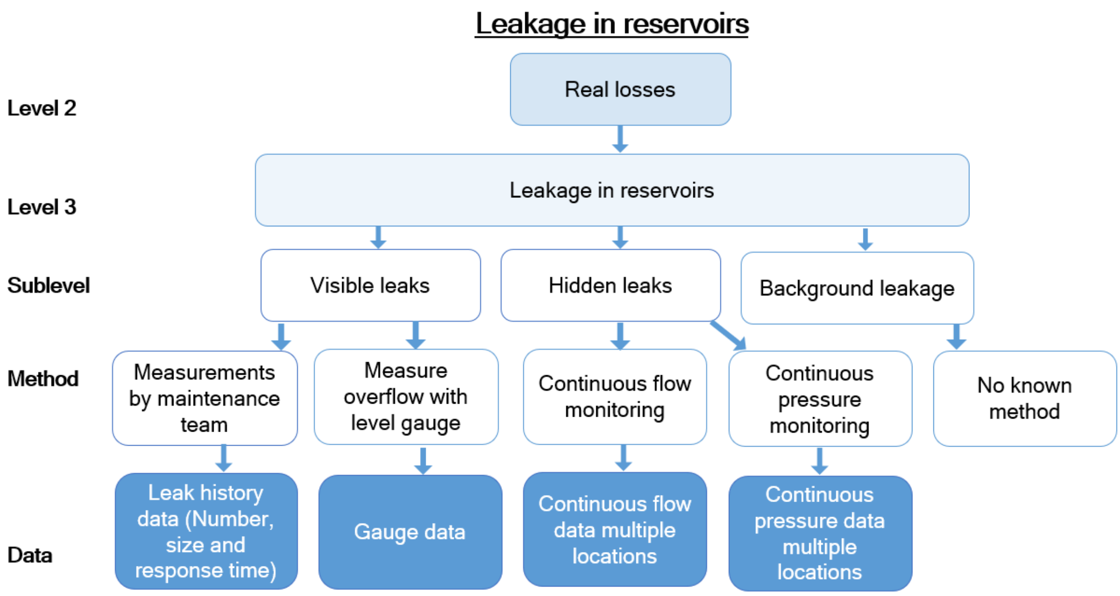

Figure A.9: Leakage and overflow at storage tanks

A.3 Report of thesis planning

This report has been made as part of the course Research Skills 1 (CIE5431) at TU Delft.

This document contains the work I have done regarding the planning of my thesis project in the last months. It starts with some background information which gives an idea of my personal perspective on the subject. Later in the process, I have chosen to switch the location of part of my study from Nairobi (Kenya) to Harare (Zimbabwe) due to several circumstances. I will elaborate upon how I came to this decision and how this switch has influenced my planning and final product.

Background information

My first personal interests in solving drinking water problems in Africa started to grow after my high school. I took a gap year (between high school and university) and lived for six months in Malindi, Kenya. I lived with a Kenyan family and worked for a local orphanage. This period has taught me a lot about the Kenyan environment and the habits of the Kenyan people. Later, during my minor, I have also lived in The Gambia for three months. Whilst some living conditions and especially religion differ a lot, I also noticed a lot of similarities. Especially the fact that people adapt a lifestyle that is more day-to-day than compared to the Netherlands. This is usually the case because people are more occupied by basic activities that they need to do in order to cover their daily needs.

It has been a wish for me for a long time to get involved in a project in my field of expertise, which is drinking water, that improves the living conditions in Kenya. Therefore, I approached Wereldwaternet (from now on referred to as WW) already in December, whilst I had planned to begin my thesis in May. During the period from December to May, we had several discussions about what thesis topic would be of most use for Nairobi. We had discussed the following three topics:

1. The first topic was to locate drinking water pipelines in Nairobi, since the GIS-file they had was not yet complete. After some discussion with geo-engineers at the TU Delft, the best way of locating pipelines would be to use a ground penetrating radar or to dig trenches at strategic locations in order to find the pipelines. The first of the two options was considered to be too expensive and the latter case was difficult since its scientific innovativeness was questionable and it absolutely required my personal presence in Nairobi (at this part of the discussions, Covid-19 had just reached the Netherlands).
2. The second topic would be to do research on the implementation of a certain software in Nairobi. This was not chosen since the contract with the software developer was not yet signed and this topic did not have my personal preference.
3. The third option would be to look more in general to the high Non-Revenue-Water percentage ($> 40\%$) and see how we could decrease this number. The main problem about the high percentage is that distinct water balances, constructed in different years, showed quite some difference in determining where the "lost water" was going to. The goal of this topic would be to develop a monitoring system that can use pressure and flow sensors to retrieve important information from the drinking water

system, in order to estimate the causes for the high NRW-percentage more accurately. This topic has eventually been chosen.

Build-up of thesis

My thesis has been divided into four parts:

- Part 1: Theoretical study. The main objective of this part of my thesis was to find out which information is important to monitor in Nairobi. Furthermore, its objective was to find out which DMA would be interesting to use as a case study for designing a monitoring network.
- Part 2: Modelling case study. The focus of this part is to design an optimal monitoring system for a single DMA in Nairobi.
- Part 3: Financial and social study. In this part of the study, I will investigate the financial and social benefits of implementing a monitoring network. It is a mandatory part of my annotation "Entrepreneurship".
- Part 4: World-wide applicability. In this part I will describe which factors are important when implementing a similar monitoring system in other places in the world.

Phase 1: Determining important information

During the first two months I have been working on determining which information is important to monitor in Nairobi. I have been through literature about the causes of Non-Revenue Water and constructed an interview structure which I could use for my interview with Nairobi Water (the operating water utility in Nairobi). At the end of this period and after some emails back and forth I managed to plan a meeting with the head of the NRW-department, engineer Mugo. The meeting was very valuable, and he could explain me a lot about the local problems for the water utility.

From this interview, I could conclude that Nairobi Water had a lot of struggles in dealing with hidden leaks in their system. These are leaks in the pipes that result in significant water loss but are not visible at the surface. Furthermore, the DMA which could be best use for a modelling study was the Northern Region of Nairobi. This is a part of Nairobi at which inflow and outflow are regularly monitored. With this information I could draw the conclusions that I needed for my first phase and move on with the next phase of my thesis: the modelling study. However, in the meanwhile there were some developments in Nairobi that complicated taking this next step.

New developments in Nairobi

After phase 1 of my study, I needed some physical data (GIS/ EPANET) to construct a hydraulic model. Previously, I had already noticed that engineer Mugo could take two or three weeks before he answered his emails. So, whilst waiting for the physical data from engineer Mugo, I set out a request on an online shared forum of WW and partners. I asked if anyone had flow or pressure data from an intermittent supply system somewhere in the world which I could analyse. I figured that having experience with analysing such a system would give me a head-start when I would receive information from Nairobi. I got a reply from an employee of VEI (a partner of WW, both operating under WaterworX which is a program of the ministry of Foreign Affairs), which had flow and pressure data from three intermittent supply systems in Harare, Zimbabwe. This data had been monitored for a few years and stored on an online platform. Furthermore, they had a complete GIS and EPANET file of the DMA's. I started analysing this data and, as expected, it gave me quite some valuable insight about what happens hydraulically inside an intermittent supply system.

At this point it is useful to explain a bit about the structure within the cooperation of WW and Nairobi Water. At WW, we have one colleague who is permanently stationed in Nairobi for two years. During the pandemic he came back to the Netherlands. This person is responsible for introductions from both sides and gives updates when there are changes in the local situation. After waiting for the reply of eng. Mugo for a few weeks, I received news from this college that things had changed in Nairobi. The national government represented by the Athi Water Works Development Agency (AWWDA) had stepped in to take over the management of the main water facilities of Nairobi Water. This development came after the government had committed to spend Kshs 18 billion to upgrade the city's water and sewer services (Nation, 2020). This development cannot be viewed separately from other developments in the city. The governor of Nairobi has been arrested on charges of corruption last year (The New York Times, 2019). Afterwards, the national government has created a new administrative organisation which stepped in to take over the responsibilities of the governor. This administrative organisation, the Nairobi Metropolitan Services, is the driving force behind Athi Water, which are taking over a lot of the daily management of Nairobi Water. This change of management had resulted in several internal changes of personnel, including a new head of the NRW-department. This new person in charge first had to orient itself on the new job, then we had to plan an introduction and afterwards I could request the needed GIS/ EPANET data. So, I had to wait another month before receiving the data. I also had planned a holiday during this period, reducing the effective waiting time to a few weeks.

Phase 2: The modelling study

For the modelling study it was important to find a software that could model intermittent supply behaviour. Therefore, whilst waiting on data from Nairobi, I experimented with a python package that could run pressure dependant simulations. For these simulations I used the data from Harare. I got experience in fitting the model to measurement data, running these pressure dependant simulations and drawing conclusions from the results. At the end of this period, I reported the progress of my planning study at Nairobi to my supervisor, using Figure A.10. I found a suitable software for my simulations, had requested the GIS data but not yet received it and was therefore unable to calibrate my Nairobi model and take further steps into my research.

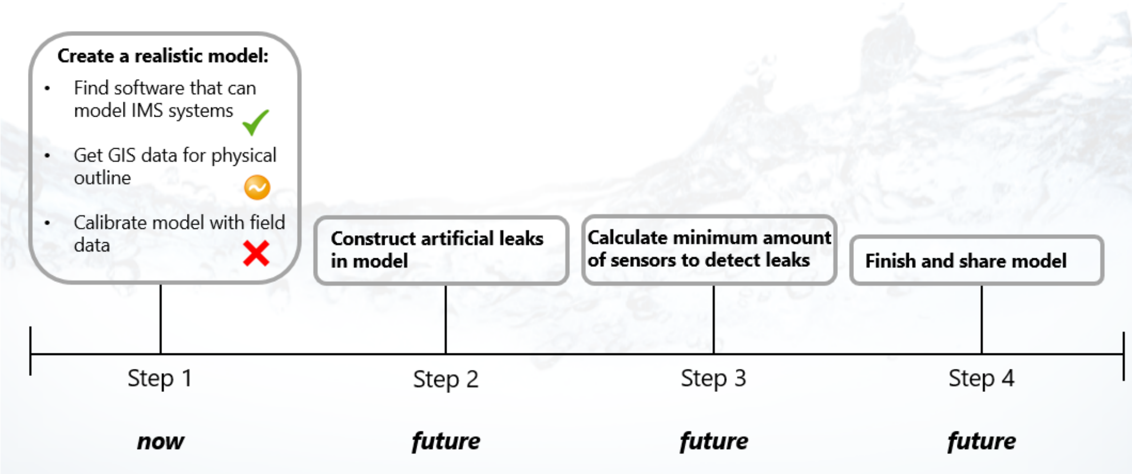

Figure A.10: Progress in the modelling study for Nairobi in July

At the end of July, I received the required GIS-data from Nairobi. However, its quality was very disappointing. A lot of the pipes seemed to be not documented in GIS, which had been mentioned previously whilst discussing a suitable topic with WW. However, I had hopes that the quality of the GIS data would still be well enough to construct a simplified model. When I had received the documentation of the bulk meters and pipes, there was a lot of contradictory information about the lengths and diameter of the pipes as well. I showed my conclusions and the files to my supervisor and we discussed whether I could still make a representable model from the data. The conclusion was that I could not.

At this point I considered two options. The first option was to get into contact with Nairobi and see if I could validate which of the contradictory information was true. However, I noticed that getting into contact with Nairobi Water through mail was time-consuming and there were many contradictions to be validated. Also, there was no perspective for me to visit Nairobi anywhere in the future. Furthermore, the lock-down in Nairobi had resulted in an increase of police violence in the city, so validating the data that I needed for my thesis, on top of covering their daily needs, would probably be low on the priorities list for many people.

The second option would be to build a monitoring system for a DMA in Harare, since I already obtained all the needed physical data of the system. Making this shift would speed up the modelling process tremendously, since I had already had a working model of the hydraulics of the DMA. This would allow me to go to the next step of my modelling study immediately (see Figure A.11). Therefore, I decided to make this switch in consultation with my supervisor.

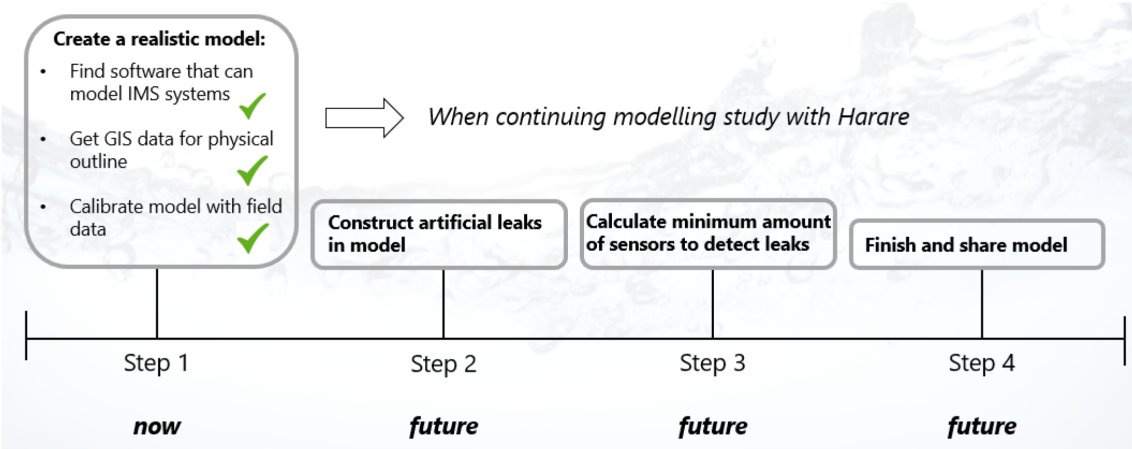

Figure A.11: Progress in modelling study when switching from a Nairobi to a Harare model

Evaluation of the switch

Switching the location of my modelling study will have both negative and positive influences on the outcome of my thesis.

To begin with the main downside, this choice will result that the final product (the monitoring system) of my thesis probably cannot be directly applied in Nairobi. It will be a suitable solution once they have completed their GIS-administration, which could take a while. Since my personal objective was to construct a solution for the drinking water problems in Nairobi, it deviates a bit from this objective.

However, it is in many ways the most suitable option. Firstly, one of the goals of my thesis was already to investigate the most important factors for developing a monitoring system, so that it can give direction how to develop such a system in other places in the world. This would be described in phase 4 and now has the extra goal to bridge any missing links between the problem description of phase 1 in Nairobi and the constructed model and its conclusions in phase 2 in Harare. So, the work of my thesis can still give directions to Nairobi how to develop their monitoring system. The financial and social benefits of such a system, could be an incentive to use a part of the investment of the government to start placing more sensors on strategic locations.

Furthermore, when keeping the modelling study in Nairobi, I would have had no clue when to receive GIS-data of sufficient quality to construct a proper hydraulic model. Therefore, it could take up way more time than initially planned. Looking back to my planning after the switch: 1) I first made a problem analysis for Nairobi, 2) received data from Harare soon afterwards, 3) made a hydraulic model and 4) will start developing the monitoring system from this point on. Since I intended to take all these steps from the beginning, only using Nairobi as a test case all the way, this switch has allowed that my initial planning was kept mostly intact during the change in circumstances.

A.4 Calibration method

This chapter describes how the hydraulic model was calibrated, using the measurements at the inlet of the DMA for calibration.

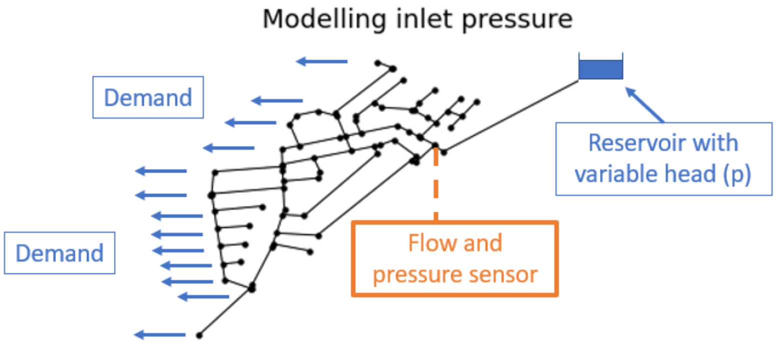

Figure A.12: The model can use input values (in blue) to simulate the flow and pressure at the location of the sensor (in orange).

The two main parameters in the model that can be changed for the simulations are the nodal demand and the reservoir head, as can be seen in Figure A.12. The model goes through an iterative process of solving mass and energy balances, until the remaining mass and energy deficits of the balances approach 0. As a solution, the model can calculate the flows that occur in pipes (noted by vector $\mathbf{q_t}$) and the pressure that occurs at nodes (noted by vector $\mathbf{h_t}$) that result from the demands in the nodes (noted by vector $\mathbf{d_t}$) and the pressure in the reservoirs and storage of the model (in this case noted by $\mathrm{h0_t}$, since only one reservoir is used). Hence, equation A.1 shows the model with its output ($\mathbf{q_t}$ and $\mathbf{h_t}$) on the left side and its input ($\mathbf{d_t}$ and $\mathrm{h0_t}$) on the right side.

The resulting pressure at any node i ($h_{i,t}$) and flow at any pipe j ($q_{j,t}$) at time t can be read from the indices of vectors $\mathbf{h_t}$ and $\mathbf{q_t}$. The devices that measure flow and pressure at the DMA inlet (Figure A.12) were estimated to be closest to pipe P73 and node N24. Therefore, the field measurements (measured flow noted by $\mathrm{q_t^*}$ and measured pressure noted by $\mathrm{h_t^*}$) should be compared with the computed flow at pipe P73 ($\mathrm{q_{j=P73,t}}$) and the computed flow at node N24 ($\mathrm{h_{i=N24,t}}$). Unfortunately the model's computations, using the initial demand settings in EPANET ($\mathbf{d_{init}}$[2]) and the average measured pressure at the reservoir ($\mathrm{h0_{init}}$), did not coincide with the measured values (equation A.2). Therefore, the model needed to be calibrated.

$$f(\mathbf{q_t}, \mathbf{h_t}; \mathbf{d_t}, \mathrm{h0_t}) = 0 \tag{A.1}$$

$$f(\mathbf{q_t}, \mathbf{h_t}; \mathbf{d_{init}}, \mathrm{h0_{init}}) = 0 \implies \mathrm{q_{j=P73}} \neq \mathrm{q_t^*}, \quad \mathrm{h_{i=N24}} \neq \mathrm{h_t^*} \tag{A.2}$$

[2]In the EPANET-file received from VEI there was already a demand pattern inserted to change the demands over time. However, using this demand pattern, the output of the model did not approach the measurements at any time. Therefore, the calibration method will "override" the inserted demand pattern and only the initial base demands $\mathbf{d_{init}}$ are taken into account at this stage.

The purpose of the calibration is to find certain factors ($\mathrm{D_t}$ and $\mathrm{H0_t}$) whereby the demands and reservoir pressure can be multiplied, so that the model's output for flow at P73 and pressure at N24 approaches the measurements at all times (equation A.3).

$$f(\mathbf{q_t}, \mathbf{h_t}; \mathbf{d_{init}} \times \mathrm{D_t}, \mathrm{h0_{init}} \times \mathrm{H0_t}) = 0 \implies \mathrm{q_{j=P73,t}} = \mathrm{q_t^*}, \quad \mathrm{h_{i=N24,t}} = \mathrm{h_t^*} \tag{A.3}$$

In order to find $\mathrm{D_t}$ and $\mathrm{H0_t}$, it is assumed that there exists a linear relationship between all demands, the pressure at the reservoir and the model's output of flow and pressure at the DMA inlet. This relationship is shown in equation A.4. A and B are matrices with parameters which define this relationship, whose values are unknown. Their sizes depend on the amount of nodes with demand and the amount of reservoirs.

$$\mathrm{A} \times \mathbf{d_{init}} + \mathrm{B} \times \mathrm{h0_{init}} = \begin{bmatrix} \mathrm{h_{i=N24}} \\ \mathrm{q_{j=P73}} \end{bmatrix} \tag{A.4}$$

Now, the effect of changing the demand (whilst keeping the reservoir pressure equal) can be examined. The amount of demand that is added to a node should be in relation to the demand that was initially used for that node. This is done by introducing a factor δ_d that allows the added demand to be a factor of the initial demand. Therefore, the vector that stores the added demands ($\mathbf{\Delta d}$) is given by $[\delta_d \times \mathrm{d_{init,1}}, \delta_d \times \mathrm{d_{init,2}}, ..., \delta_d \times \mathrm{d_{init,n}}]$ and can be written as $\delta_d \times \mathbf{d_{init}}$. In the example below, $\mathbf{\Delta d}$ was set to equal the initial demand (so $\mathbf{\Delta d} = \mathbf{d_{init}}$ and $\delta_d = 1$), or in other words: the demand is doubled. The influence that this change in demand has on the pressure and flow near the sensor's location is shown in equation A.5.

$$\begin{gathered} \mathrm{A} \times (\mathbf{d_{init}} + \mathbf{\Delta d}) + \mathrm{B} \times \mathrm{h0} = \begin{bmatrix} 0.999 \times \mathrm{h_{i=N24}} \\ 1.988 \times \mathrm{q_{j=P73}} \end{bmatrix} \\ \mathbf{\Delta d} = \delta_d \times \mathbf{d_{init}}, \quad \text{with } \delta_d = 1 \end{gathered} \tag{A.5}$$

The perturbation caused by adding $\mathbf{\Delta d}$ can be written explicitly, as shown in equation A.6.

$$\mathrm{A} \times (\mathbf{d_{init}} + \mathbf{\Delta d}) + \mathrm{B} \times \mathrm{h0_{init}} = \begin{bmatrix} \mathrm{h_{i=N24}} \\ \mathrm{q_{j=P73}} \end{bmatrix} + \begin{bmatrix} -0.001 \times \mathrm{h_{i=N24}} \\ 0.988 \times \mathrm{q_{j=P73}} \end{bmatrix} \tag{A.6}$$

Therefore, the sole impact of $\mathbf{\Delta d}$ can be expressed (equation A.7).

$$\mathrm{A} \times \mathbf{\Delta d} = \begin{bmatrix} -0.001 \times \mathrm{h_{i=N24}} \\ 0.988 \times \mathrm{q_{j=P73}} \end{bmatrix} \tag{A.7}$$

The same procedure can be followed for the pressure that is added to the system by the reservoir. Keeping the demands equal and by adding a certain $\Delta\mathrm{h0}$ (determined by δ_{h0}), the perturbation caused by this $\Delta\mathrm{h0}$ can be expressed. Equation A.8 shows the perturbation caused by $\Delta\mathrm{h0} = \mathrm{h0_{init}}$ (so the inlet pressure has been doubled and $\delta_{h0} = 1$).

$$\begin{gathered} \mathrm{B} \times \Delta\mathrm{h0} = \begin{bmatrix} 1.000 \times \mathrm{h_{i=N24}} \\ 0.000 \times \mathrm{q_{j=P73}} \end{bmatrix} \\ \Delta\mathrm{h0} = \delta_{h0} \times \mathrm{h0_{init}}, \quad \text{with } \delta_{h0} = 1 \end{gathered} \tag{A.8}$$

The difference between the measured flow and pressure and the output of the model at the sensor's location, after using the initial demand and inlet pressure as input values, can also be written as a perturbation. The magnitude of this perturbation is time-dependent, since it depends on the values of the measurements which change over time. Equation A.9 shows how this is done.

$$\begin{bmatrix} \mathrm{h_t^*} \\ \mathrm{q_t^*} \end{bmatrix} = \begin{bmatrix} \mathrm{h_{i=N24}} \\ \mathrm{q_{j=P73}} \end{bmatrix} + \begin{bmatrix} (\mathrm{h_t^*}/\mathrm{h_{i=N24}} - 1) * \mathrm{h_{i=N24}} \\ (\mathrm{q_t^*}/\mathrm{q_{j=P73}} - 1) * \mathrm{q_{j=P73}} \end{bmatrix} \tag{A.9}$$

Now, the target is to tweak $\boldsymbol{\Delta}\mathbf{d}$ and $\Delta \mathrm{h0}$ in such a way that they result in the desired perturbation as defined in equation A.9. The problem can be written by introducing parameters $x1_t$ and $x2_t$ that determine the magnitude of perturbations Δd_t and $\Delta \mathrm{h0_t}$, as shown in equation A.10.

$$\mathrm{A} \times \boldsymbol{\Delta}\mathbf{d} \times x1_t + \mathrm{B} \times \Delta \mathrm{h0} \times x2_t = \begin{bmatrix} (\mathrm{h_t^*}/\mathrm{h_{i=N24}} - 1) * \mathrm{h_{i=N24}} \\ (\mathrm{q_t^*}/\mathrm{q_{j=P73}} - 1) * \mathrm{q_{j=P73}} \end{bmatrix} \tag{A.10}$$

Combining equation A.7, A.8 and A.10, the problem can be rewritten to a linear problem with two equations and two unknowns ($x1_t$ and $x2_t$). This is shown in equation A.11 for $\boldsymbol{\Delta}\mathbf{d} = \mathbf{d_{init}}$ and $\Delta \mathrm{h0} = \mathrm{h0_{init}}$. The right-hand side solution changes as the measured values ($\mathrm{h_t^*}$ and $\mathrm{q_t^*}$) change over time. Since the number of equations equals the number of unknowns, these equations can be solved for $x1_t$ and $x2_t$.

$$\begin{bmatrix} -0.001 \\ 0.988 \end{bmatrix} x1_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} x2_t = \begin{bmatrix} (\mathrm{h_t^*}/\mathrm{h_{i=N24}} - 1) \\ (\mathrm{q_t^*}/\mathrm{q_{j=P73}} - 1) \end{bmatrix} \tag{A.11}$$

Substituting equation A.4 and A.10 into equation A.9, gives:

$$\mathrm{A} \times (\mathbf{d_{init}} + \boldsymbol{\Delta}\mathbf{d} \times x1_t) + \mathrm{B} \times (\mathrm{h0_{init}} + \Delta \mathrm{h0} \times x2_t) = \begin{bmatrix} \mathrm{h_t^*} \\ \mathrm{q_t^*} \end{bmatrix} \tag{A.12}$$

With the solution for $x1_t$ and $x2_t$, the earlier mentioned factors $\mathrm{D_t}$ and $\mathrm{H0_t}$ can be calculated. Equations A.3 and A.12 are combined to result in equations A.13 and A.14, which express $\mathrm{D_t}$ and $\mathrm{H0_t}$ as a function that is determined by δ_d and δ_{h0} and the solutions $x1_t$ and $x2_t$. These latter solutions are determined by the measurements at the sensor and the initial demands and inlet pressure. The equation for output of the model at the sensor's location for the calibrated model is given in equation A.15.

$$\begin{gathered} \mathrm{D_t} \times \mathbf{d_{init}} = \mathbf{d_{init}} + \boldsymbol{\Delta}\mathbf{d} \times x1_t \\ \mathrm{D_t} = \frac{\mathbf{d_{init}}}{\mathbf{d_{init}}} + \frac{\delta_d \times \mathbf{d_{init}}}{\mathbf{d_{init}}} \times x1_t = 1 + \delta_d \times x1_t \end{gathered} \tag{A.13}$$

$$\begin{gathered} \mathrm{H0_t} \times \mathrm{h0_{init}} = \mathrm{h0_{init}} + \Delta \mathrm{h0} \times x2_t \\ \mathrm{H0_t} = \frac{\mathrm{h0_{init}}}{\mathrm{h0_{init}}} + \frac{\delta_{h0} \times \mathrm{h0_{init}}}{\mathrm{h0_{init}}} \times x2_t = 1 + \delta_{h0} \times x2_t \end{gathered} \tag{A.14}$$

$$\begin{gathered} f(\mathbf{q_t}, \mathbf{h_t}; \mathbf{d_{init}} \times (1 + \delta_d \times x1_t), \mathrm{h0_{init}} \times (1 + \delta_{h0} \times x2_t)) = 0 \implies \\ \mathrm{q_{j=P73,t}} \approx \mathrm{q_t^*}, \quad \mathrm{h_{i=N24,t}} \approx \mathrm{h_t^*} \end{gathered} \tag{A.15}$$

In Figure A.13, you can see how the output of the model at the location of the sensor compared to the flow and pressure data that were measured on Mondays, for using $\delta_D = 1$ and $\delta_p = 1$. It can be seen that the results of the model fit the measurements quite well.

At higher flows, the fit is less precise. This is likely due to the assumption of linearity, whilst the relationships are not linear in the calculations that are included in the modelling software. The influence of choices for $\delta_D = 1$, $\delta_p = 1$ are shown in section 5.3.

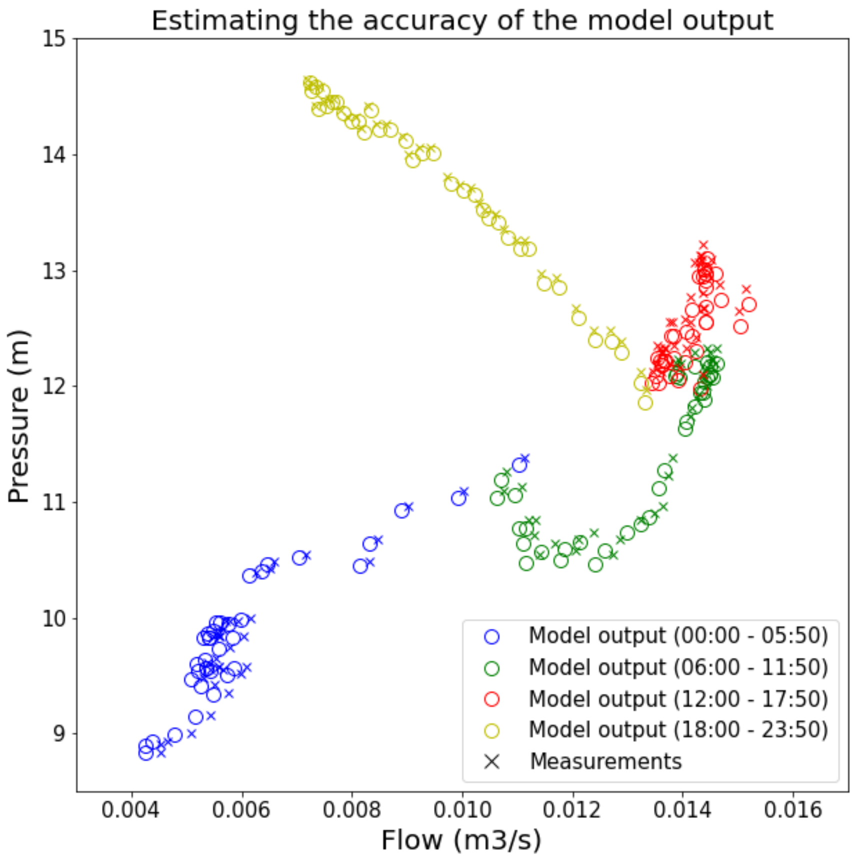

Figure A.13: Representation of how well the model output coincides with the measured data points. The dots represent the model output and the x's represent the sensors measurements. The different colors represent different hours of the day.

A.5 Daily demand patterns

Ashdown Park

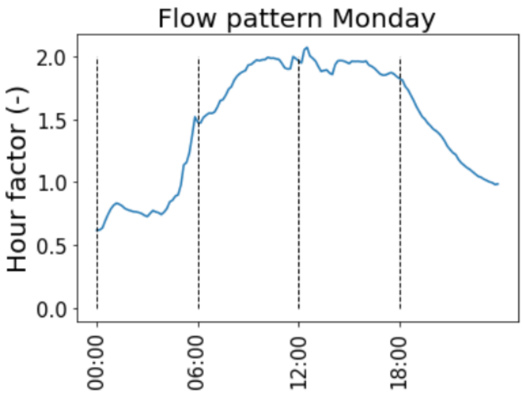

Figure A.14: Hourly factors Monday Ashdown Park

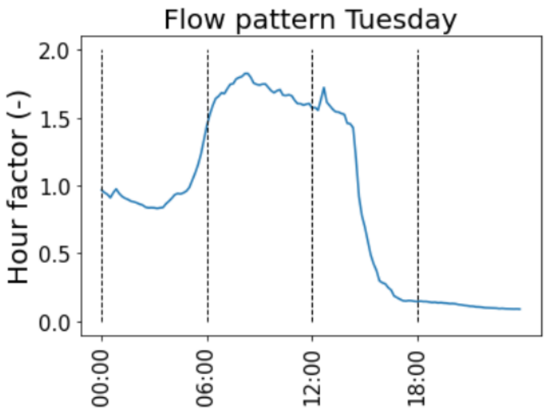

Figure A.15: Hourly factors Tuesday Ashdown Park

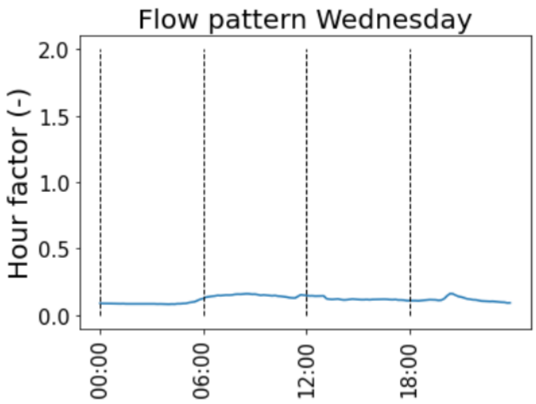

Figure A.16: Hourly factors Wednesday Ashdown Park

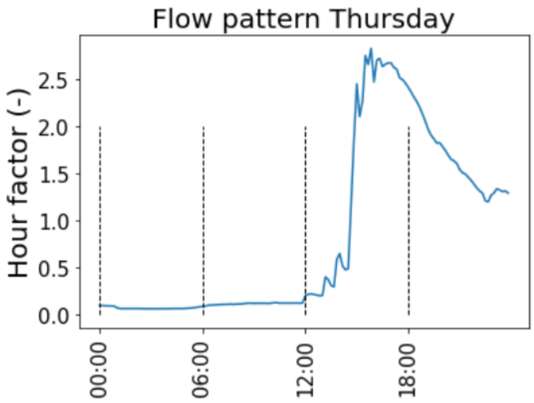

Figure A.17: Hourly factors Thursday Ashdown Park

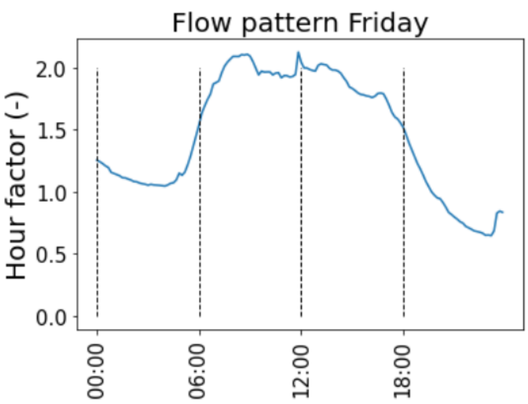

Figure A.18: Hourly factors Friday Ashdown Park

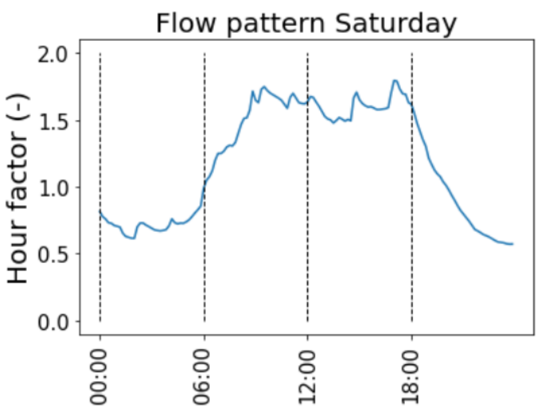

Figure A.19: Hourly factors Saturday Ashdown Park

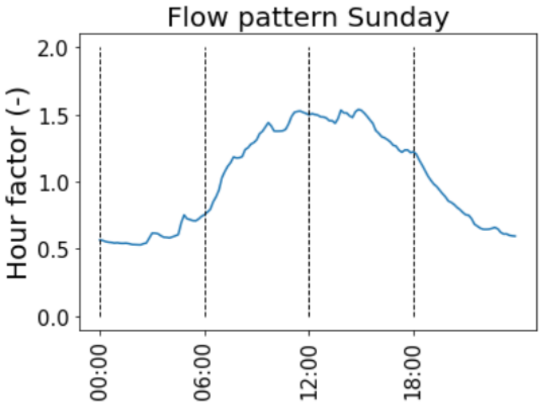

Figure A.20: Hourly factors Sunday Ashdown Park

Marimba Park

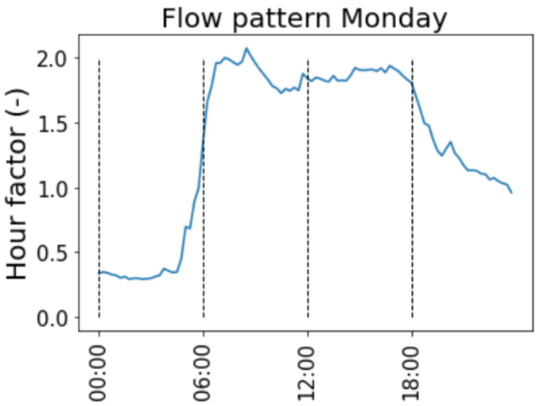

Figure A.21: Hourly factors Monday Marimba Park

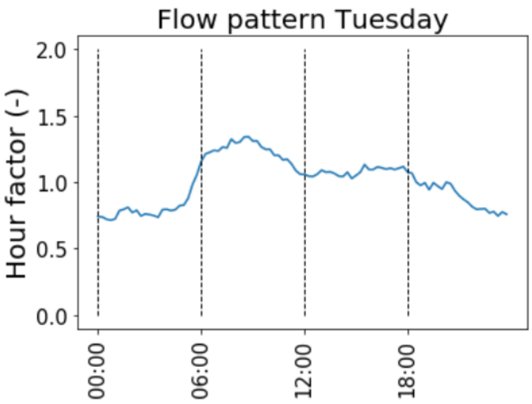

Figure A.22: Hourly factors Tuesday Marimba Park

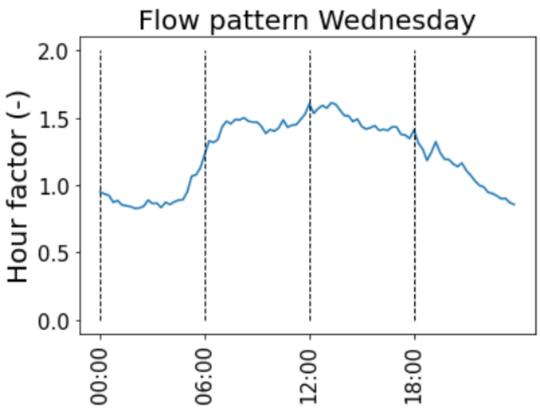

Figure A.23: Hourly factors Wednesday Marimba Park

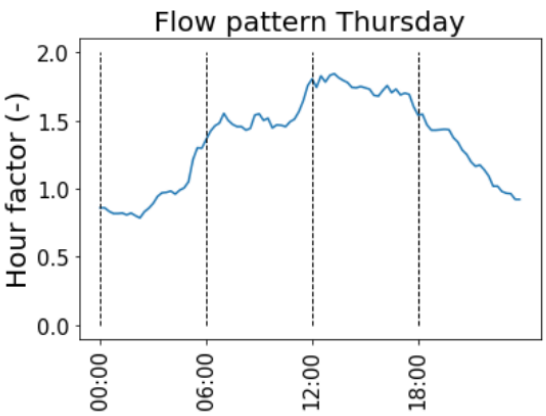

Figure A.24: Hourly factors Thursday Marimba Park

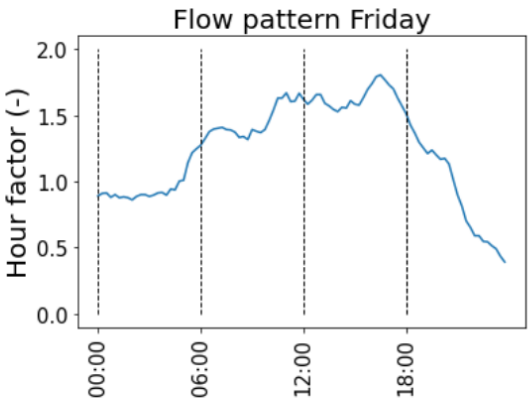

Figure A.25: Hourly factors Friday Marimba Park

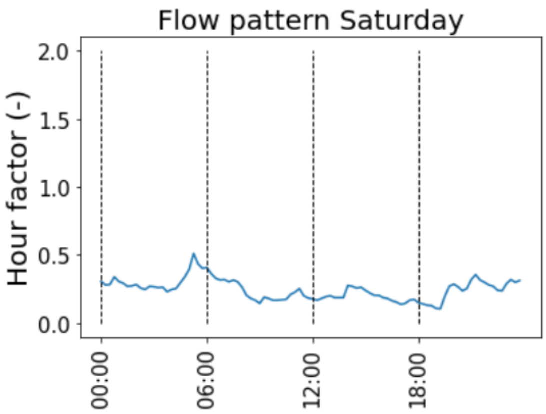

Figure A.26: Hourly factors Saturday Marimba Park

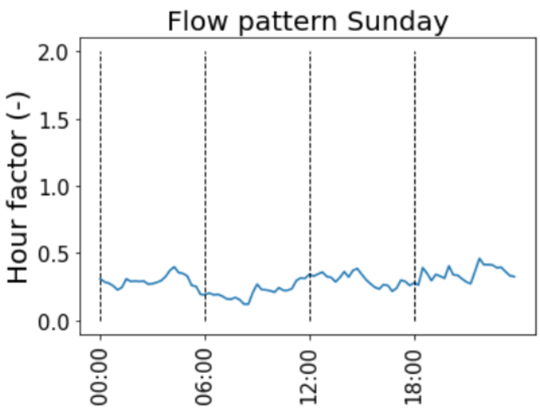

Figure A.27: Hourly factors Sunday Marimba Park

A.6 Pump configuration

A large part of Harare's drinking water supply comes from the Morton Jaffray (MJ) water treatment plant. This treatment plant also produces water for Ashdown Park. Water from MJ is pumped to the Warren Control (WC) centre, from which it is directed towards other pumping stations. The main part of the drinking water from WC is pumped towards Alexandra Park pumping station (mostly called "Alex") or towards Letombo pumping station, before which the water is directed towards DMA's or smaller pumping stations. Ashdown park is connected to the transport line from WC towards Alex. A schematic overview of this configuration is given in Figure A.28.

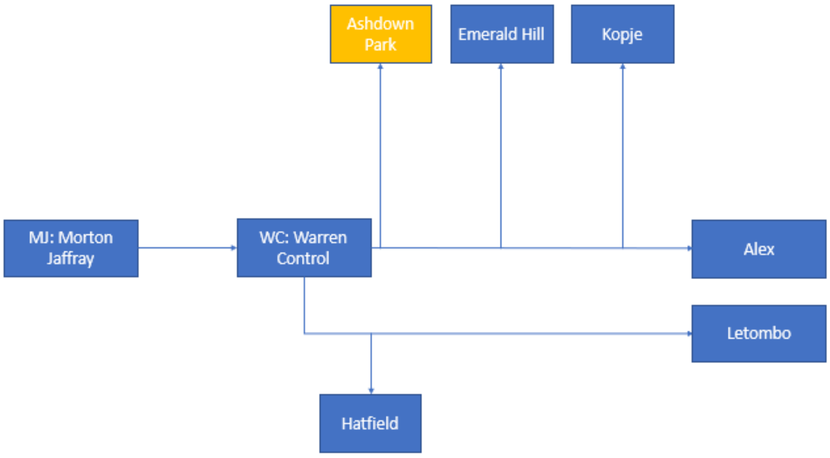

Figure A.28: Schematic pump configuration from Morton Jaffray

At Waternet, in Amsterdam, variable speed pumps (VSP's) are used to add pressure the system. The rotations per minute (rpm) of these pumps are constantly adjusted to retain a constant pressure at different locations throughout the system (de Groot, 2020). Harare does not use VSP's, resulting in pumps that are either "ON" or "OFF" (Moors, 2020). As a consequence, the number of pumps that pump water from WC to Alex should influence the pressure in Ashdown Park.

To see whether this indeed occurs, pump data from 01-10-2019 until 19-09-2020 was analyzed. This data-set contained information how many pumps were used to pump water from WC towards Alex (max: 3) and how many pumps were pumping towards Letombo (max: 3). Although the data-range does not completely coincide with the data range from the measurements (08-06-2019 until 08-06-2020), the average weekly pattern allows for reasonable comparisons. For example, the pressure pattern on Friday (Figure A.29) can be compared to the average pump configuration on Friday at WC (Figure A.30). The pressure pattern shows a comparable gradient to the average number of operational pumps from WC to Alex, meaning that if less pumps are operational the pressure in Ashdown Park decreases and vice versa.

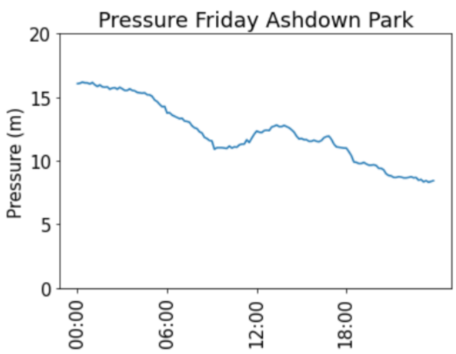

Figure A.29: Pressure pattern on Friday at Ashdown Park.

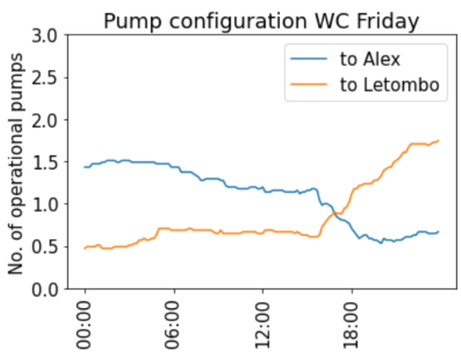

Figure A.30: Average of operational pumps on Friday from WC.

The pressure pattern and average number of operational pumps towards Alex have a similar pattern on other days with supply (see Figure A.31 - A.36). This substantiates the statements that the pump configuration at WC influences the pressure in Ashdown Park.

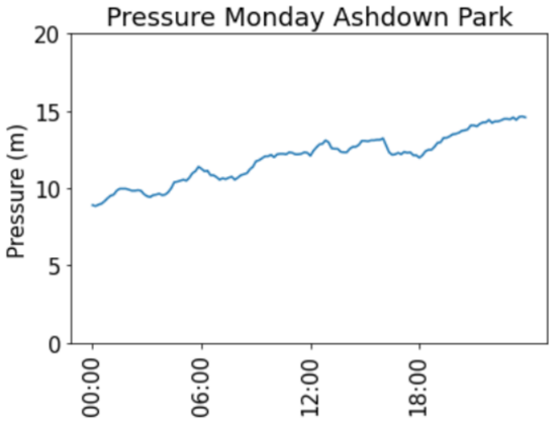

Figure A.31: Pressure pattern Monday Ashdown Park.

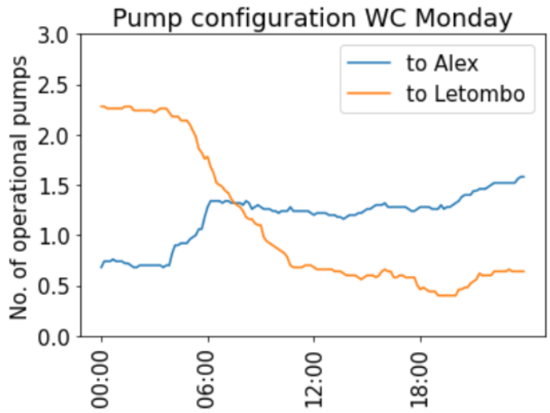

Figure A.32: Pump configuration Monday Ashdown Park.

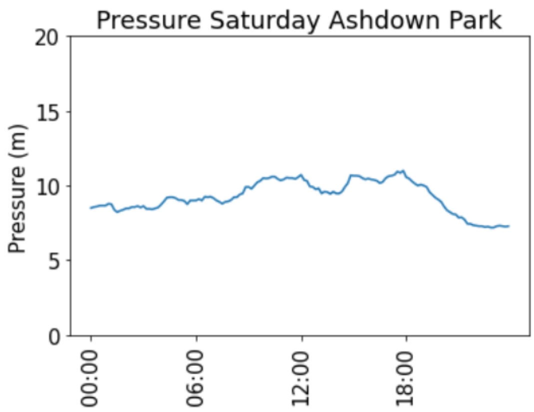

Figure A.33: Pressure pattern Saturday Ashdown Park.

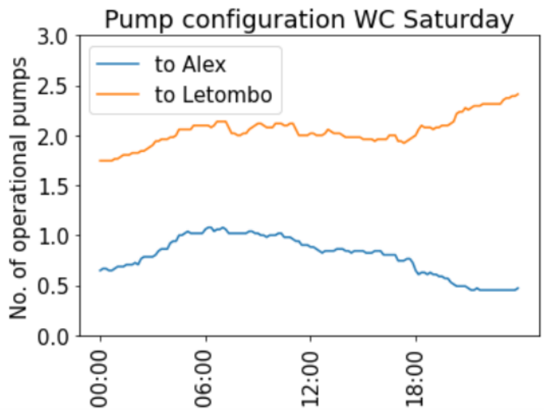

Figure A.34: Pump configuration Saturday Ashdown Park.

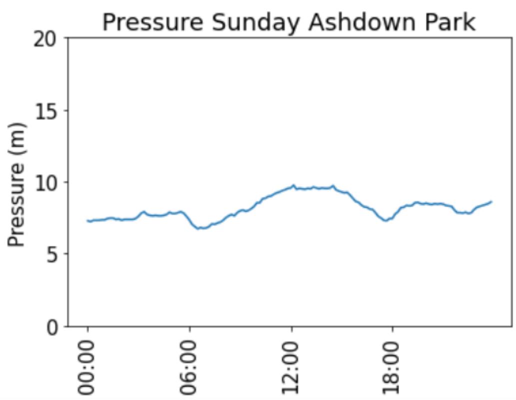

Figure A.35: Pressure pattern Sunday Ashdown Park.

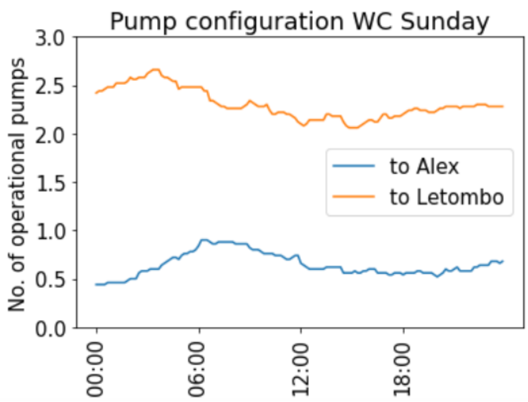

Figure A.36: Pump configuration Sunday Ashdown Park.

The pressure of the days with intermittent supply can not be compared to the pump configuration, since the pressure drops to 0 when the DMA is closed off. This can be seen for example at Tuesdays (Figure A.37 and A.38).

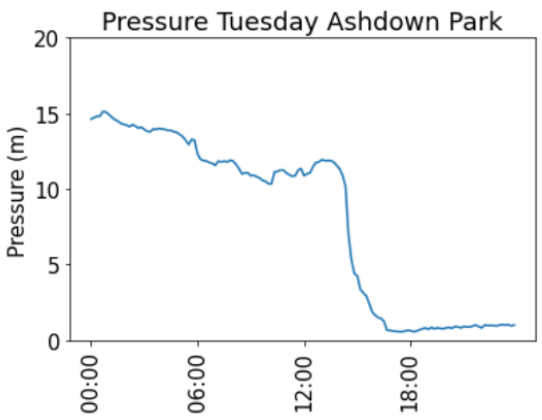

Figure A.37: Pressure pattern Tuesday Ashdown Park.

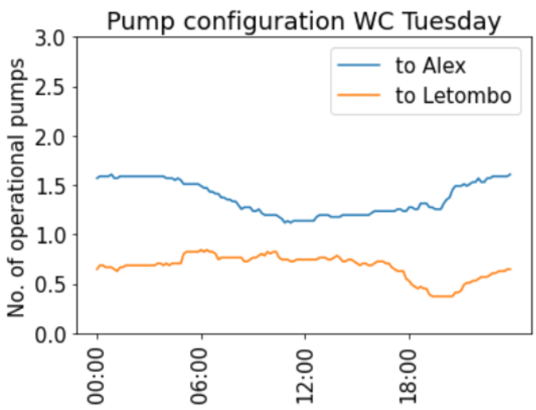

Figure A.38: Pump configuration Tuesday Ashdown Park.

A.7 Roughness coefficients Darcy-Weisbach and Hazen-Williams

The initial roughness value in the received EPANET-file was 0.6×10^{-3} (feet). This value indicates that roughness of the pipe is in the range of the roughness of galvanized iron pipes (0.5×10^{-3} feet) and cast iron pipes (0.85×10^{-3} feet), see Table A.2. Therefore, the adjusted Hazen-Williams roughness coefficient had to be between 120 (-) (galvanized iron) and 130 (-) (cast iron). The chosen roughness coefficient for all the pipes in the model was 125 (-).

Material	*Hazen-Williams C (unitless)*	*Darcy-Weisbach ε (feet x 10^{-3})*	*Manning's n (unitless)*
Cast Iron	130 – 140	0.85	0.012 - 0.015
Concrete or Concrete Lined	120 – 140	1.0 - 10	0.012 - 0.017
Galvanized Iron	120	0.5	0.015 - 0.017
Plastic	140 – 150	0.005	0.011 - 0.015
Steel	140 – 150	0.15	0.015 - 0.017
Vitrified Clay	110		0.013 - 0.015

Table A.2: Table with roughness coefficients (Rossman, 2000).

A.8 Sensitivity analysis

For the sensitivity analysis, multiple values for δ_D and δ_p were used in the method described in chapter **??** at different nominal pressure. The resulting NRMSE for the flow and the pressure were used as a indicator of how well the method performed. The values for δ_D and δ_p which resulted in the lowest NRMSE's were used to estimate the demand and flow factors. When a NRMSE value was below 0.05, the model output is regarded as comparable to the measurement values. The values which were chosen during the different iterations can be found in the tables below. Usually the initial approach was to change one parameter and keep the other parameter the same, to see the effect that changing a single parameter has on the NRMSE's.

Nominal pressure: 12m

First, δ_D was kept at a value of 1 and δ_p was changed. This showed that the optimal value of δ_p would be around 1.

δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
1	0.8	0.0064	0.10	0.11
1	0.9	0.0073	0.047	0.054
1	1	0.013	0.0070	0.020
1	1.1	0.0217	0.058	0.080
1	1.2	0.034	0.11	0.14

Table A.3: Iterations for the sensitivity analysis at a nominal pressure of 12m, whilst keeping ΔD equal.

Afterwards, δ_p was kept equal at 1 and δ_D was changed. This showed that changing δ_D was of minor influence. The lowest NRMSE values were found when δ_p would be set to 1 and δ_D would be 0.1. This can be seen in Table A.4.

δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
0.1	**1**	**0.012**	**0.0070**	**0.019**
1	1	0.013	0.0070	0.020
2	1	0.013	0.0070	0.020
3	1	0.013	0.0070	0.020

Table A.4: Iterations for the sensitivity analysis at a nominal pressure of 12m, whilst keeping ΔP equal.

Nominal pressure: 20m

First, δ_D was kept at a value of 1 and δ_p was changed. This showed that the optimal value of δ_p is expected to be around 0.9.

δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
1	0.1	0.012	0.46	0.47
1	0.5	0.081	0.23	0.31
1	0.8	0.14	0.062	0.20
1	0.9	0.16	0.041	0.20
1	1	0.18	0.081	0.26
1	1.1	0.21	0.13	0.34
1	2	0.47	0.66	1.13

Table A.5: Iterations for the sensitivity analysis at a nominal pressure of 20m, whilst keeping ΔD equal at 1.

Next, δ_p was kept at a value of 0.9 and δ_D was changed. This showed that quite high values of δ_D needed to be chosen in order to achieve low NRMSE's. The most promising combination would be a δ_D of 17 and a δ_p of 0.9. This combination resulted that both RMSE's are below 0.05, so they are comparable to the measurements. However, the total NRMSE was still above 0.05, making the model output not fully satisfactory. Therefore, another round of iterations was made whilst keeping δ_p at 1.

δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
0.1	0.9	0.51	0.64	1.15
1	0.9	0.16	0.041	0.20
2	0.9	0.15	0.029	0.18
5	0.9	0.13	0.038	0.17
10	0.9	0.089	0.042	0.13
15	0.9	0.047	0.043	0.090
16	0.9	0.042	0.043	0.085
17	0.9	0.041	0.043	0.084
18	0.9	0.044	0.043	0.087
20	0.9	0.060	0.043	0.10

Table A.6: Iterations for the sensitivity analysis at a nominal pressure of 20m, whilst keeping ΔP equal at 0.9.

The results of the iterations whilst keeping δ_p at 1 is shown below in Table A.7. The best combination was a δ_D of 18 and a δ_p of 1, resulting in both NRMSE's below 0.05 and a total NRMSE that was lower than the previously found combination. The total NRMSE is still above 0.05, indicating that a first order approximation of the hydraulic equations is not entirely suitable with high nominal pressures. Still, since both individual NRMSE's are below 0.05, this model output was regarded as sufficient for the purposes of this thesis.

δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
10	1.0	0.11	0.014	0.12
15	1.0	0.063	0.012	0.075
17	1.0	0.051	0.012	0.063
18	**1.0**	**0.048**	**0.012**	**0.060**
19	1.0	0.049	0.012	0.061
20	1.0	0.053	0.012	0.065
21	1.0	0.060	0.012	0.072
25	1.0	0.11	0.012	0.12

Table A.7: Iterations for the sensitivity analysis at a nominal pressure of 20m, whilst keeping ΔP equal at 1.

Nominal pressure: 1m

First, δ_D was kept at a value of 1 and δ_p was changed. This showed that the optimal value of δ_p would be around 1.

δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
1	0.1	0.0090	0.47	0.11
1	0.9	0.0090	0.047	0.054
1	1	0.0090	0.0072	0.020
1	1.1	0.0090	0.058	0.080
1	2	0.022	0.53	0.14

Table A.8: Iterations for the sensitivity analysis at a nominal pressure of 1m, whilst keeping ΔD equal at 1.

Afterwards, δ_p was kept equal at 1 and δ_D was changed. This showed that changing δ_D had almost no influence on $NMRSE_{pressure}$, but it did on $NMRSE_{flow}$. The lowest NRMSE values were found when δ_p would be set to 1 and δ_D would be 13. This can be seen in Table A.9.

δ_D	δ_p	$NRMSE_{flow}$	$NRMSE_{pressure}$	Total NRMSE
0.1	1	0.0091	0.0072	0.016
1	1	0.0090	0.0072	0.016
2	1	0.0089	0.0072	0.016
5	1	0.0081	0.0072	0.015
10	1	0.0061	0.0071	0.013
11	1	0.0058	0.0071	0.013
12	1	0.0054	0.0071	0.013
13	**1**	**0.0051**	**0.0071**	**0.012**
14	1	0.0089	0.0072	0.016
15	1	0.012	0.0072	0.019

Table A.9: Iterations for the sensitivity analysis at a nominal pressure of 1m, whilst keeping ΔP equal at 1.

A.9 Calculations nominal pressure Ashdown Park

The length of the pipe connection the houses to the network can be estimated at 10m, with 20mm diameter. The pressure loss over the water meter ($h_{loss,meter}$) is usually around 0.5m and the highest point that can be reached is the inlet of the storage tanks, which can be 4-5m above ground level (Moors, 2020). The roughness of the pipe can be between 0mm (plastic) and 0.046mm (wrought iron) (Elger et al., 2014). The formulae that can be used to calculate the pressure loss from the network to the storage tank can be expressed as:

$$H_{loss} = \Delta z + h_{loss,friction} + h_{loss,meter} \tag{A.16}$$

$$\Delta z = z_{storage} - z_0 \tag{A.17}$$

$$h_{loss,friction} = K\frac{(\frac{Q}{A})^2}{2g}, \quad A = \frac{1}{4}\pi D^2 \tag{A.18}$$

$$K = f\frac{L}{D} + k_{bends} + k_{inlet} \tag{A.19}$$

Filling in formula A.17 gives us $\Delta z = 5m$. Using the highest roughness value of 0.046mm and the Darcy-Weisbach formula, f was found to be 0.044 (-). The calculation of f is explained in more detail in appendix A.10. A normal loss factor resulted from a bend in the pipe is 0.2 (Elger et al., 2014). Given that there are usually a quite some bends up to the storage tank, k_{bends} is roughly estimated as 1. A normal value for k_{inlet} is 0.5 (Elger et al., 2014).

To calculate frictional losses, it is needed to determine a certain maximum household flow. During this thesis a field experiment in Ashdown Park was conducted. During this field experiment, two 20L jerrycans were filled directly from the water distribution system. The first jerrycan took 4 minutes to fill, so the flowrate was 0.3 m3/h, and the second jerrycan took 4.5 minutes to fill (flow rate: 0.27 m3/h).[3] These flowrates give some indication of the rates that can be expected to come from the tap, but it is still likely that higher values can occur. Therefore, a safety factor of 1.5 was used to get to a design flow rate of 0.45 m3/h, which shall be used for the head loss calculations. [4] With a flow of 0.45 m3/h, the resulting $h_{loss,friction}$ is 0.057m, using equation A.18. This makes the total head loss over the connection at home roughly 5.57m. Taking some uncertainties into account, a nominal pressure in the range of 6-8m could be expected for Ashdown Park. The systems pressure should be above this value to fulfill people's demand.

A.10 Friction loss house connection

In order to calculate the friction loss with the Darcy-Weisbach formula, the friction factor f need to be calculated. First, the Reynolds is calculated with the formula:

$$Re_D = \frac{vD}{\nu} \tag{A.20}$$

In this equation v is the velocity (0.22 m/s), D is the diameter (0.02 m) and ν is the kinematic viscosity of water ($1.14 * 10^{-6}\ m^2/s$). This results in a Reynold's number of around 3800 (-). The formula which can be used to calculate the cohering friction factor f is (Elger et al., 2014):

$$f = \frac{0.25}{[log_{10}(\frac{k_s}{3.7D} + \frac{5.74}{Re_D^{0.9}})]^2} \tag{A.21}$$

As mentioned in chapter 5.3, the roughness (k_s that is used is $0.046 * 10^{-3}$ m. Filling in equation A.21 results in a friction factor f of 0.044 (-).

A.11 Extensive consideration of calculating the mean and spread from the historical measurements

In order to construct the alarm values, it is important to analyze which flow and pressure values can be expected in the "normal scenario". It is therefore very helpful to determine the expected flow and pressure at a certain time. This is done by analyzing the data-set as described in Figure **??** and retrieving a mean value and spread of the flow and pressure at a certain time. Furthermore, it is important to keep in mind that this mean value and spread can only be expected at the entrance of the DMA, since this is the location where the measurements where conducted. The expected alarm value *per pipe* or *per node* will result from simulations of the model.

The mean

[3] Only a small number of these experiments could be conducted, since I was not able to travel to Harare myself and had to depend on the goodwill of other people for conducting experiments.

[4] Another way of assessing the maximum household flow, would be to simply use 1 m3/h as the maximum flow, as this is a value that is often used for determining head losses at household connections in the Netherlands (Clement and de Groot, 2020).

The average flows and pressures during the different days are shown in Figure A.39 and A.40. Since the daily patterns at different weekdays can differ significantly in IMS systems , it is important to look at the weekly pattern and take into account the differences between these weekdays. By doing so, the choice emerges to construct a mean value in two different ways. One could determine a separate mean (per time step) for each weekday (Figure A.39 and A.40) or one could construct one mean value (per time step), which holds for every day with continuous supply. Below, the consequences and differences between the two different ways are explained.

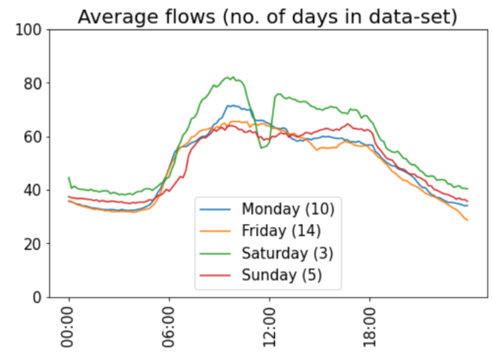

Figure A.39: Average flows of normative days.

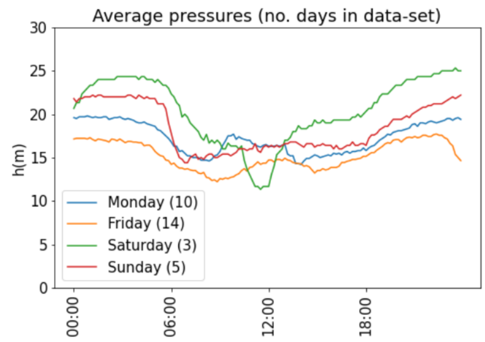

Figure A.40: Average pressures of normative days.

First of all, the mean value for each weekday can be calculated separately. The advantage of using separate means will be that the means includes difference in hydraulic behaviour on the separate weekdays and therefore describes the flow and pressure quite accurately. A downside of using separate means is that they are constructed from small data-sets. The mean value for the pressure on Saturday at 00:00 for instance would be based on only three measuring points (see Figure **??**), making the value not very reliable. Using separate means would also eventually result in separate alarm values for the different days, slightly decreasing the ease of operating the Dynamical Bandwidth Monitor.

The other option would be to use one mean value for the flow and pressure, which holds for every day at a specific time. The advantage is that it a relatively large dataset can be used to calculate the mean. Using this mean will also result in alarm values that can be used for every day with continuous supply, making the Dynamical Bandwidth Monitor easy to operate. The disadvantage of using this mean is that it does not include the hydraulic behaviour of the system during different weekdays.

The difference in using the two above described options can be understood when looking into the flow factors (or pressure factors). The flow factor is defined by the measured flow, divided by the mean flow. Using a different mean will result in different flow factors. The data-set that is used to construct the alarm values of flows consists of 4608 measurements (32 days with 144 time-intervals). So, constructing flow factors with the two different types of means result in two different data-sets of flow factors, each with 4608 entries. The similarity of the two data-sets can be expressed by calculating the Root-Mean-Square error (RMSE) between the two Boyd and Vandenberghe (2018). If the RMSE is low, say below 0.05, the data-set can be considered as similar and there is not a lot of difference whether a single mean or distinct means are used to calculate the flow factors. If the RMSE is above 0.05, the difference between the two data-sets can be considered significant. A visualisation of comparing these different ways of constructing flow factors is given in Figure A.41.

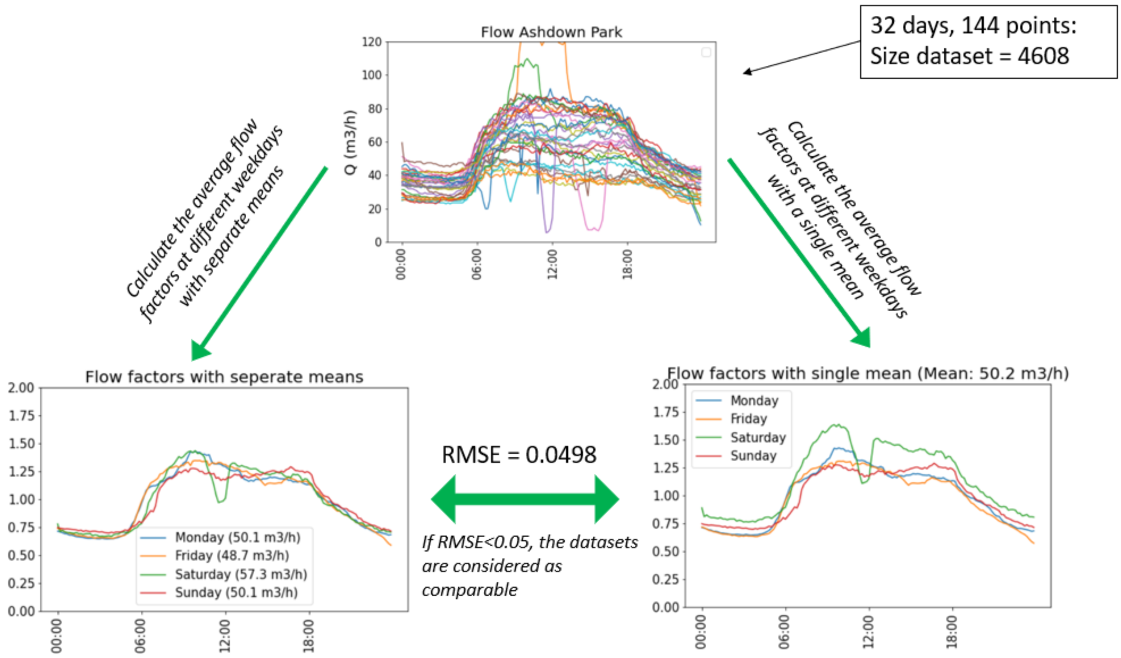

Figure A.41: Flow factors calculated with different means.

The above described comparison was both made for flow factors and for pressure factors. In case of the flow factors, the RMSE was 0.0498. This indicates that there is not a significant difference between using a single mean or separate means for each weekday. Therefore, a single mean was used to construct a baseline for the flow values. This gives the added benefits of using a relatively large data-set to determine the mean and use the same alarm values for each day with continuous supply.

When doing the same comparison for pressure factors, the resulting RMSE was 0.101. This indicates that there is a significant difference between using separate means for constructing the pressure baseline and using a single mean. In other words, the pressure during the different weekdays deviates significantly from the mean pressure of all days with continuous supply. So, a single value for the mean pressure can not be used for the baseline and a mean for every weekday has to be constructed separately. This has the disadvantage that smaller data-set has to be used to determine the mean and the resulting alarm values will differ for each weekday.

The fact that the flows on the considered days show more similarities than the pressure on the respective days, should be explained by looking at the driving force behind the quantity. Although the quantities have some correlation, chapter 5.3 has shown that in Ashdown park the quantities do not influence each other a lot. Therefore, it can be stated that the inlet pressure is driven by the pressure in the main transport system and the flow is driven by the water consumption of consumers. Apparently, the behaviour of consumers has a more similar pattern at different days with continuous supply than the pressure in the main transport system.

The spread

The spread is of data is in this thesis described as the standard deviation (σ) of the flow or pressure factors. It is feasible to use a single mean value for the flow during days with continuous pressure, which holds for every day at a specific time. Therefore, the flow at any day and at a specific time ($Q_{wd,time}$) can be described by equation A.22, in which Q_{avg} is the average weekly flow, Qf_{time} is the flow factor of that specific time and x_{random} is a random draw from a normal distribution with mean 1 and the standard deviation of all flow factors at a specific time ($\sigma_{Qf,time}$).

$$Q_{time} = Q_{avg} * Qf_{time} * x_{random} \quad , x_{random} \sim N(1, \sigma^2_{Qf,time}) \tag{A.22}$$

$$Q_{avg,time} = Q_{avg} * Qf_{time} \tag{A.23}$$

The pressure at the various weekdays differs too much to use a single mean. Therefore, a separate mean is used for each weekday. This sequentially results in different pressure factors during different weekdays. Therefore, the pressure at any weekday and at a specific time ($h_{wd,time}$) can be described by equation A.24, in which $h_{wd,avg}$ is the average pressure of that weekday and time, $hf_{wd,time}$ is the pressure factor of that specific weekday and time and x_{random} is a random draw from a normal distribution with mean 1 and the standard deviation of all pressure factors of that weekday at a specific time ($\sigma_{hf,wd,time}$).

$$h_{wd,time} = h_{wd,avg} * hf_{wd,time} * x_{random} \quad , x_{random} \sim N(1, \sigma^2_{hf,wd,time}) \tag{A.24}$$

$$h_{wd,avg,time} = h_{wd,avg} * hf_{wd,time} \tag{A.25}$$

Again, it should be noted that $Q_{wd,time}$ and $h_{weekday,time}$ describe the flow and pressure that can occur at the DMA inlet.

A.12 Detailed explanation modelling demand configurations according to method 1.

Let us consider the average flow that enters the DMA at a specific time, say 10:00. In areas with water supply with strong daily variations, which is common in IMS, it is even better to consider a flow at a specific time on a specific weekday, so say 10:00 on Monday. However, let us just assume 10:00 at any day for now. Sometimes at 10:00 the incoming flow might be above the average and on some days, it might be below the average. The extent to which the flow value deviates from the average value can be described by calculating the mean and standard deviation from a historical dataset of measurements. This mean (the average flow value, $Q_{avg,10:00}$) and standard deviation ($\sigma_{Q;10:00}$) can then be used to construct a normal distribution. The flow at any day at 10:00 can be described by a random draw from this normal distribution, as described in equation A.26.

$$Q_{10:00} \sim N(Q_{avg,10:00}, \sigma_{Q;10:00}) \tag{A.26}$$

Dividing the mean value and the standard deviation by the mean itself, allows us to describe the flow at 10:00 as a random draw from another normal distribution with mean 1. This results in equation A.27.

$$Q_{10:00} = Q_{avg,10:00} * Q_f, \quad Q_f \sim N(1, \frac{\sigma_{Q;10:00}}{Q_{avg,10:00}}) \tag{A.27}$$

Next, it is assumed that all the water that enters the DMA flows to the nodes (eq. A.28, no leakage) and that the demand at the nodes has the same variability as the inflow at the DMA inlet ($D_f = Q_f$). Let the demand at node i at 10:00 be given by $D_{i,10:00}$ and let there by n nodes in the network. The demand at any node i can be described by the equation A.30 (which results from eq. A.29).

$$\sum_{i=1}^{n} D_{i,avg,10:00} = Q_{avg,10:00} \tag{A.28}$$

$$\sum_{i=1}^{n} D_{i,10:00} = \sum_{i=1}^{n} (D_{i,avg,10:00} * D_f), \quad D_f \sim N(1, \frac{\sigma_{Q;10:00}}{Q_{avg,10:00}}) \tag{A.29}$$

$$D_{i,10:00} = D_{i,avg,10:00} * D_f, \quad D_f \sim N(1, \frac{\sigma_{Q;10:00}}{Q_{avg,10:00}}) \tag{A.30}$$

By running several simulations and letting the demand at every node being determined by a draw from the normal distribution, deviations from the average demand allocation can be included in the model (Figure A.42). By a shift in spatial allocation of demand, the flow magnitudes change. When the largest and lowest flows that occur in pipes are saved, alarm values for the Dynamic Bandwidth Monitors can be constructed.

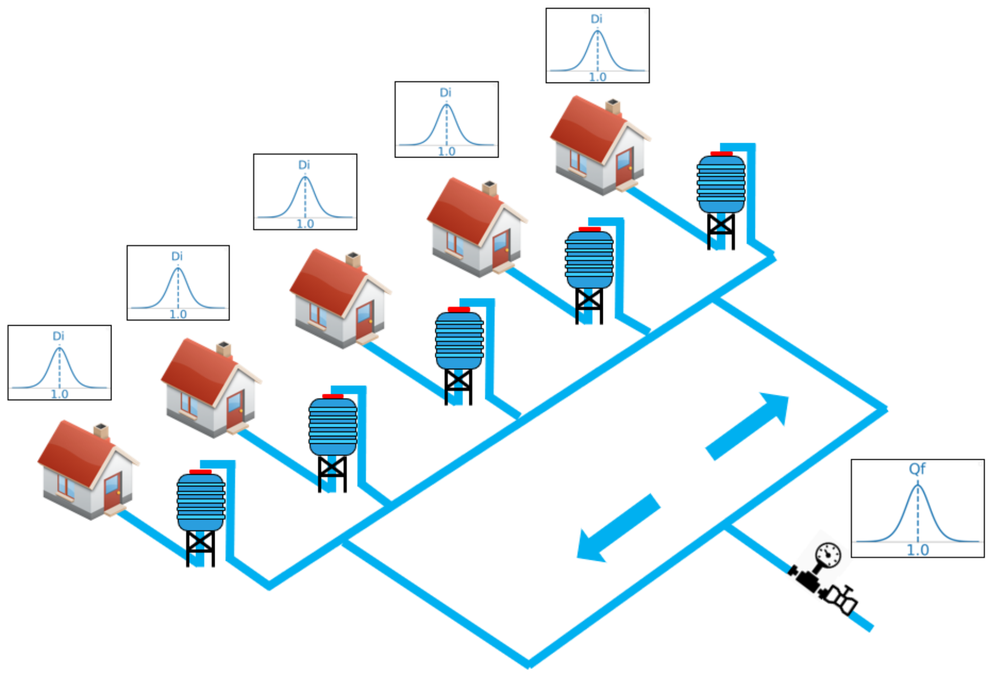

Figure A.42: Modelling the demand as random draws from a normal distribution.

A large advantage of using normal distributions to model different spatial demand configurations is that the variety in demands at different moments can be taken into account. A disadvantage, however, is that it is unsuitable to model a single household. Whilst a household has its taps either “open” or “closed”, the normal distribution method will not be able to model closed taps, since it can not assign 0 demand to a node. It will always assign a certain demand to a node at any time. Therefore, the method does not model the configurations within a distribution system which results in the most extreme flows. If Figure A.42 were to be a transport system for water and the houses were to be DMA’s, this method would be well suitable, since a DMA always has a certain demand and can deviate from its mean.

A.13 Spread in demand realizations for pressure alarm values

Figure A.43 shows the spread that occurs due to modelling different demand realizations in order to construct pressure alarm values at every time step. It compares the spread with the 95-percentile values of flows that are expected to occur on Friday (chapter **??** explains why Friday is used) and the 5-percentile values. The figure shows that this way of modelling different demand realizations produces the best results between 00:00-06:30 and 18:30-23:50.

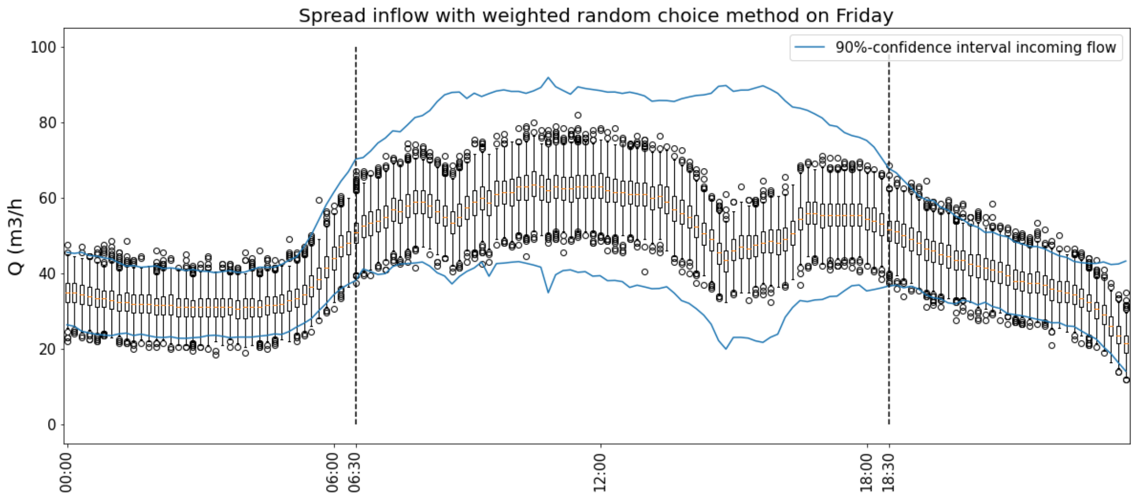

Figure A.43: The spread in demand realizations compared to the modelled inflow for the construction of a monitoring network with pressure sensors.

A.14 Number of experiments for the monitoring system with pressure sensors

When doing the analysis as described in chapter 6.3 for the design which used pressure meters, similar results were obtained. The percentage of false alarm seemed quite unpredictable when modelling only 10 different experiments for the 3D-matrix, probably because it does not take into account a significant part of the possible demand realizations. The best and most steady results were obtained when using 1000 experiments to define the alarm values and 100 experiments to construct the 3D-matrix. The results of this analysis for pressure meters can be found in Figure A.44.

Percentage of false alarms with a different number of experiments for constructing **pressure** alarm values and the Boolean 3D-matrix

Calculated by simulating leaks of 0.1mm and calculating the percentage of 1's in the 3D-matrix

Time	06:00			12:00			22:00		
Number of experiments	10	100	1000	10	100	1000	10	100	1000
10	9.2%	0.26%	1.6%	5.6%	2.7%	0%	7.9%	1.6%	0%
100	7.6%	1.1%	0.09%	15.7%	2.0%	0.03%	6.5%	1.4%	0%

Number of experiments in 3D-matrix (rows) / Number of experiments for constructing alarm values (columns)

Figure A.44: Percentage of false alarms for the design with pressure meters and a different number of iterations.

A.15 Comparing the accuracy and precision of the detectability at different times of the day

The detectability results for using a network with flow sensors and a network with pressure sensors shall be compared. First, a closer look is taken at the similarity in the detectability. For this, a second degree polynomial was fitted through both plots. These polynomials can be seen in Table A.10. This plot shows as well the ranges that were set for good detectability. It shows that the average results of the detectability for both sensors are quite similar in the time ranges were the leaks are best detected. In the region were both models produce less accurate output (during the middle of the day) the similarity between the two networks decreases. As said, the night times have the best detectability. This can be explained by noticing that low flows occur during the night (Figure A.46). Therefore the leak volume becomes relatively large compared to the flows in the network, increasing its effect on flow magnitudes and pressure.

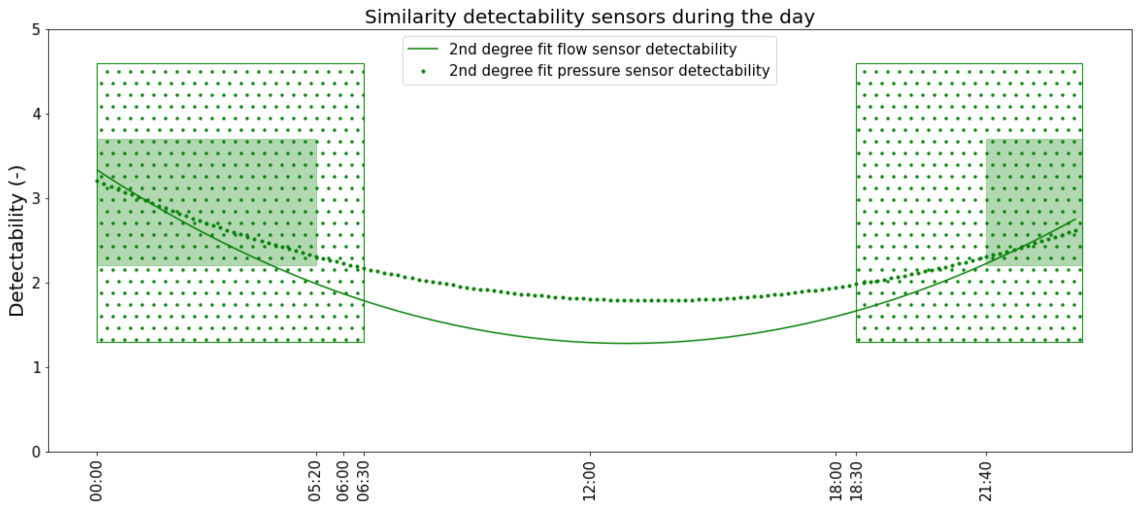

Table A.10: Detectability of leaks with pressure sensors at different times of the day.

However, the precision of the two methods shows a significant difference. The precision, visualized by the difference between the measuring points and the polynomial fit in Figure A.45, is significantly lower for the network with pressure sensors than for the network with flow sensors. This low precision adds uncertainty to whether the network will be able to detect leaks.

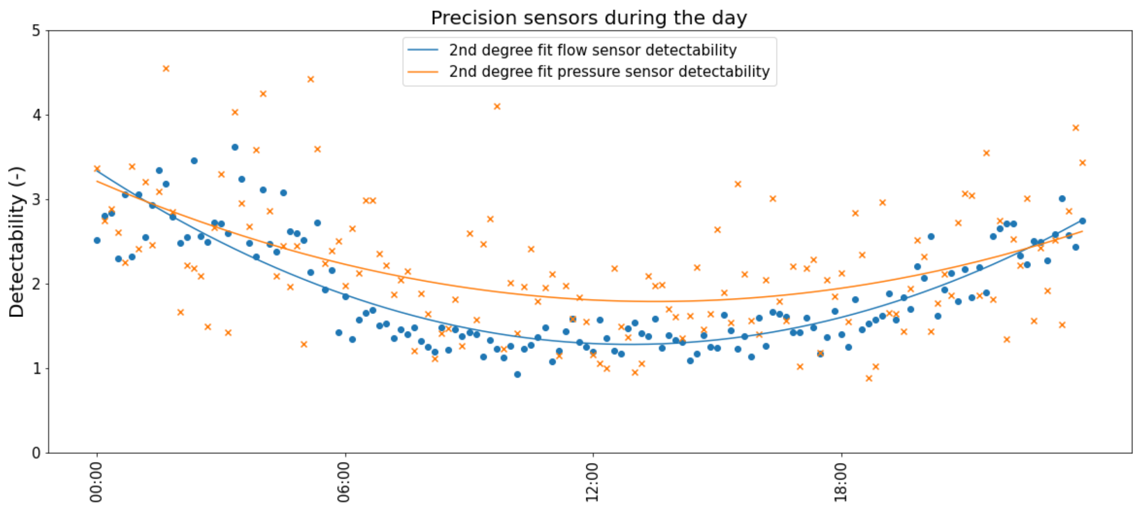

Figure A.45: Detectability of leaks with pressure sensors at different times of the day.

As described in previous chapters, the detectability of leaks is determined by first modelling a range of expected values in scenarios *without* leaks and than compared to results of scenarios *with* leaks. The main difference between these scenarios is logically the leak. For the network with flow sensors, the leak volumes are slightly higher, since the modelled pressure at the inlet of Ashdown Park is higher (Figure A.47). The flow magnitude is similar for calculating the allocation of flow and pressure sensors (Figure A.46).

A larger leak volume in the flow sensor network does not solely explain the difference in precision between the two networks. The difference between pressure in subsequent time steps is also larger in the pressure network, since this only uses data from Fridays This results in a less smooth the daily pressure curve, as can be seen in Figure A.47. This could also add to the precision. At last a significant part of the precision is introduced by calculating the detectability at scenario's with random allocated demand. A model output which is sensitive to the extreme scenarios within this randomness of allocation will produce less precise results than a model which is less sensitive to this random allocation. Flow magnitude and direction within the pipes is directly influenced by a new demand allocation and so sensitive to every new scenario. The pressures at the nodes are more indirectly influenced by a new demand allocation, as they are the result of an equilibrium of the energy distribution over the entire system. The scenario's that result in large pressure drops in some parts of the system might therefore be more rare, but have a larger impact on the detectability. The sensitivity of the model to these extremes might explain the decreased precision in using pressure meters, but it underlying causes can not be determined with certainty in this thesis.

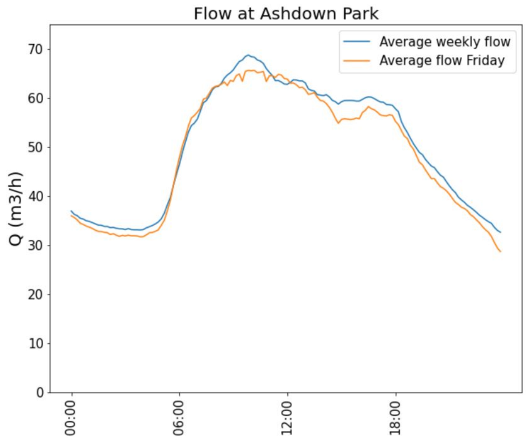

Figure A.46: Flows at Ashdown Park for pressure and flow sensor calculations.

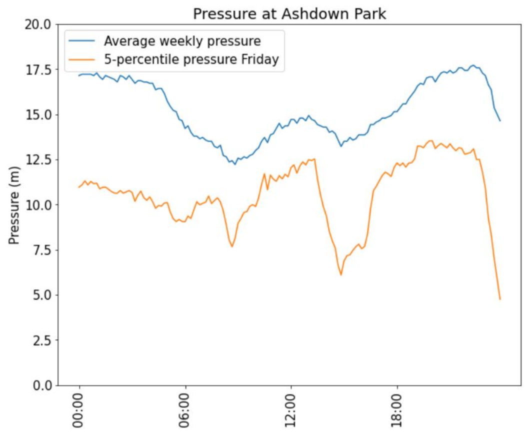

Figure A.47: Pressure at Ashdown Park for pressure and flow sensor calculations.

A.16 Calculating the number of possible combinations

The number of possible combinations (C) of samples (r) from a number of objects (n) can be calculated by applying the following formula to calculate the binomial coefficient:

$$C(r,n) = \binom{n}{r} = \frac{n!}{r!(n-r)!} \tag{A.31}$$

Using this formula it can be calculated that 71 samples of 424 objects have $8.2 * 10^{81}$ possible combinations.

A.17 Percentages of leaks found when applying a different number of sensors

Monitoring network with flow sensors

Figure A.48, A.49, A.50, A.51 and A.52 show the percentage of leaks which can be found in which pipes under the different scenarios for optimally allocating one, two, three, four and five flow sensors in Ashdown Park.

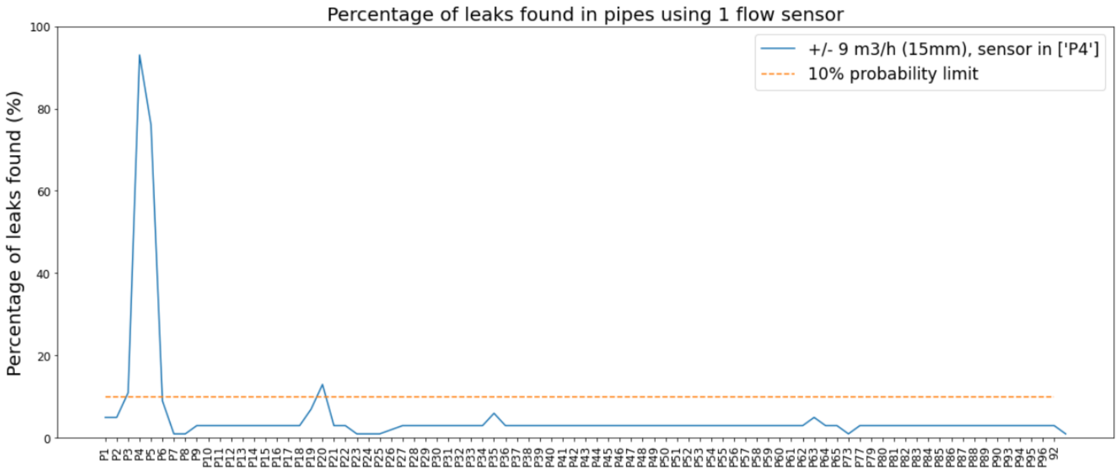

Figure A.48: Percentage of leaks found using one flow sensor.

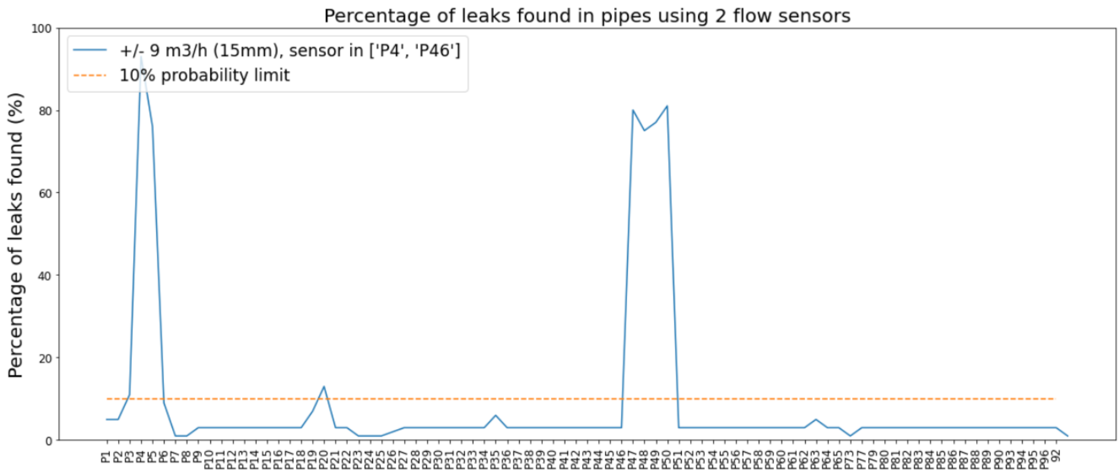

Figure A.49: Percentage of leaks found using two flow sensors.

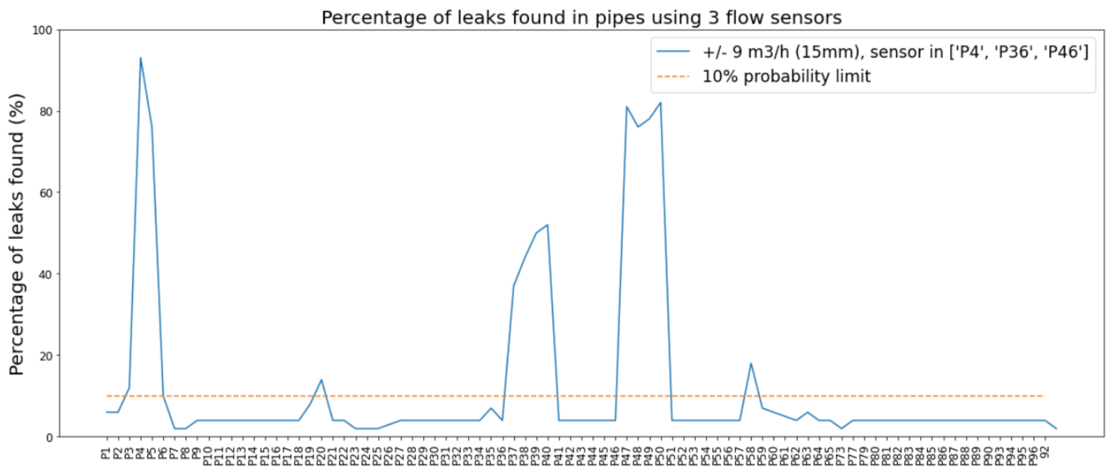

Figure A.50: Percentage of leaks found using three flow sensors.

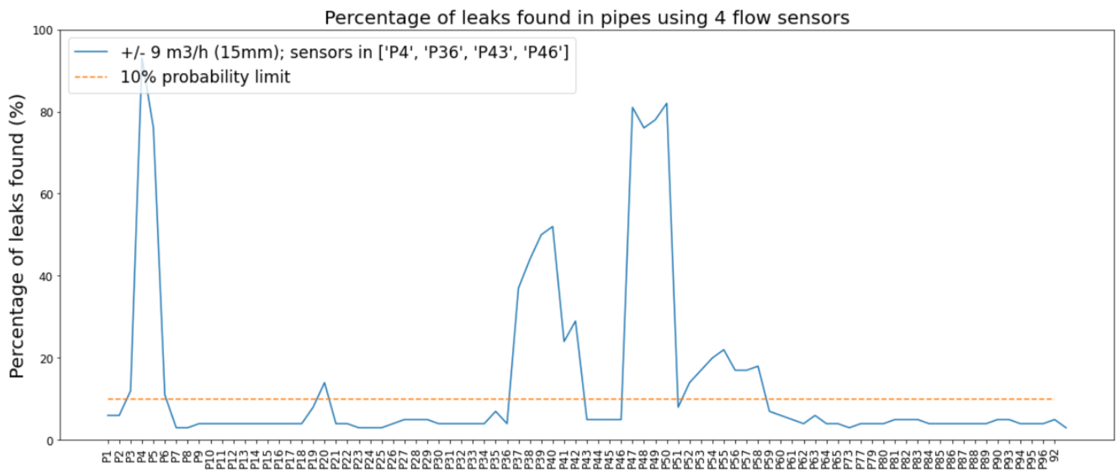

Figure A.51: Percentage of leaks found using four flow sensors.

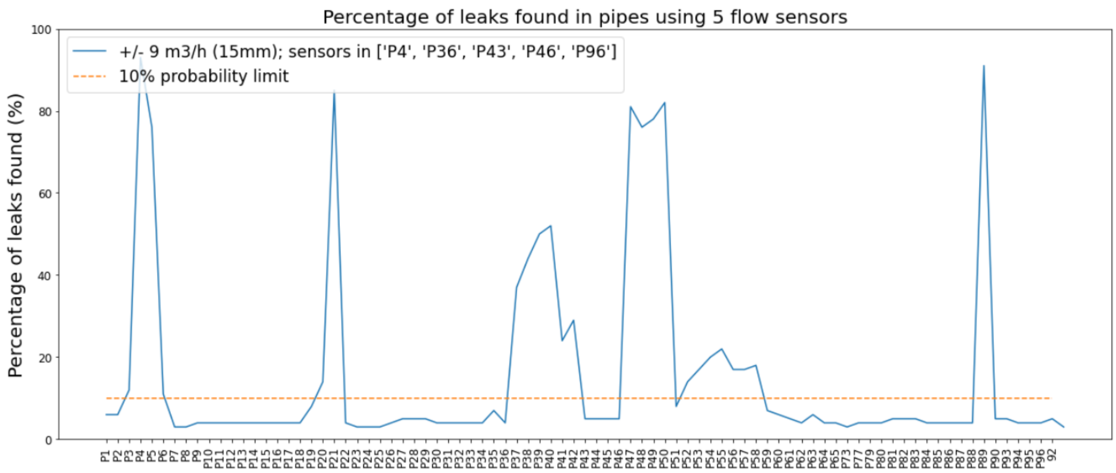

Figure A.52: Percentage of leaks found using five flow sensors.

Monitoring network with pressure sensors

Figure A.53, A.54, A.55, A.56 and A.57 show the percentage of leaks which can be found in which pipes under the different scenarios for optimally allocating one, two, three, four and five pressure sensors in Ashdown Park.

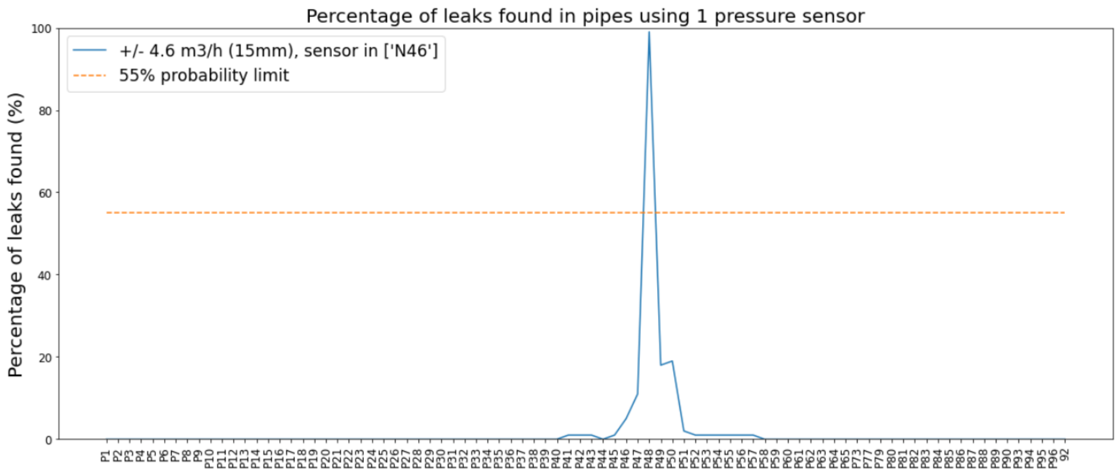

Figure A.53: Percentage of leaks found using one pressure sensor.

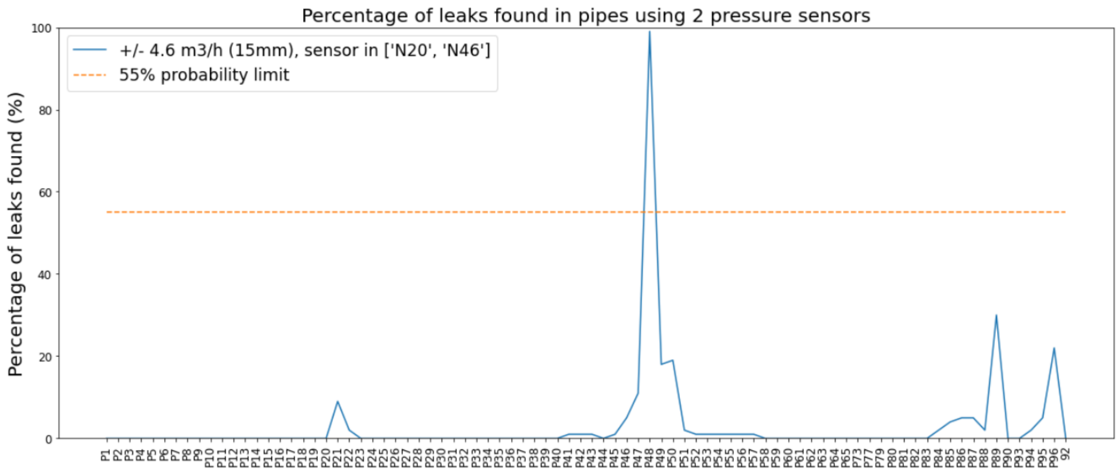

Figure A.54: Percentage of leaks found using two pressure sensors.

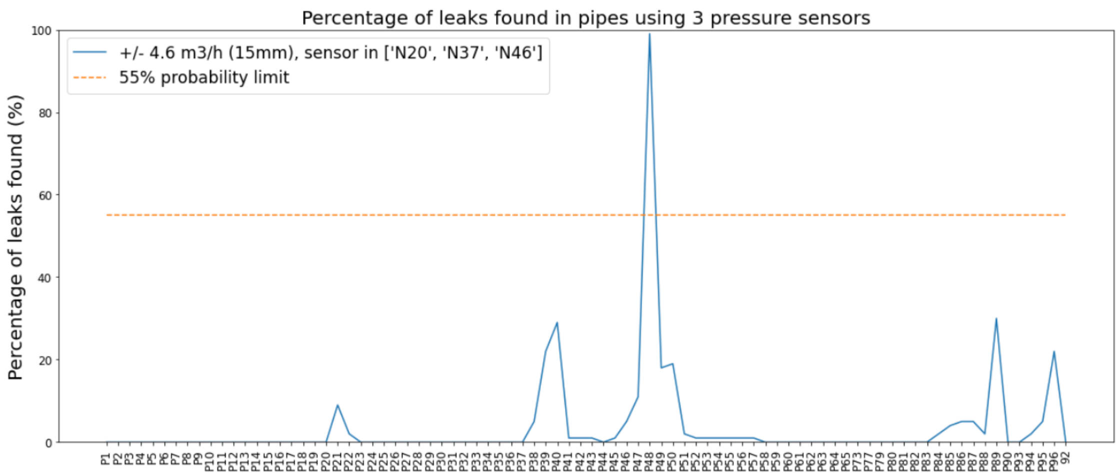

Figure A.55: Percentage of leaks found using three pressure sensors.

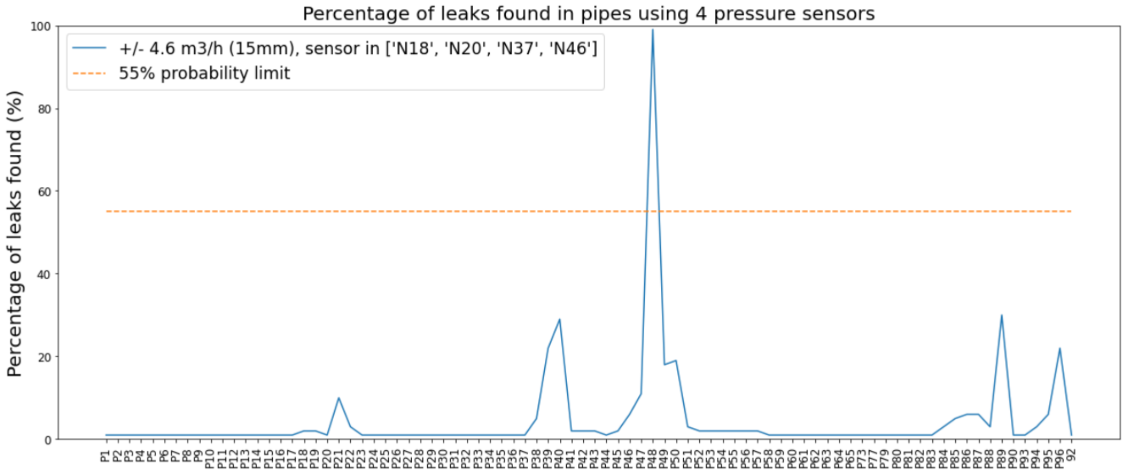

Figure A.56: Percentage of leaks found using four pressure sensors.

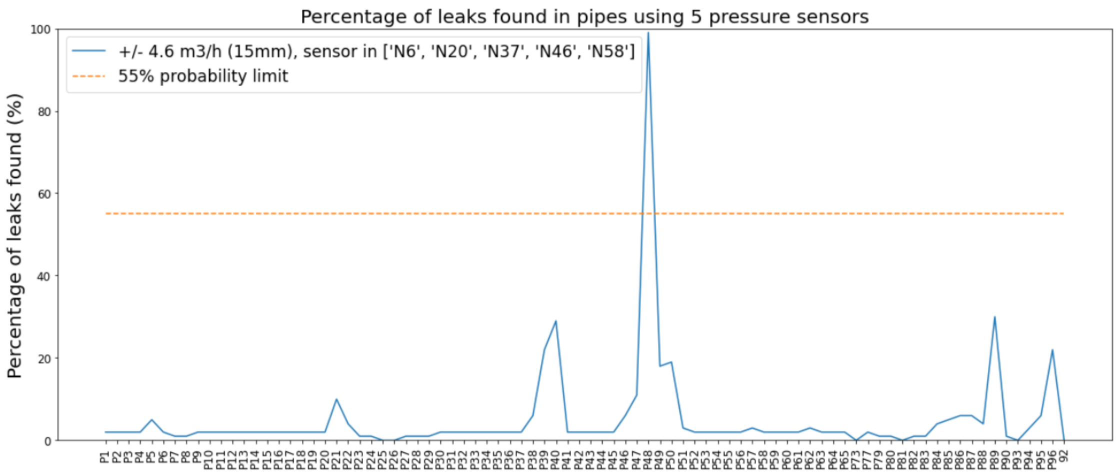

Figure A.57: Percentage of leaks found using five pressure sensors.

Monitoring network with pressure sensors and a pressure dependent DBM system

Figure A.58, A.59, A.60, A.61 and A.62 show the percentage of leaks which can be found in which pipes under the different scenarios for optimally allocating one, two, three, four and five pressure sensors with a pressure dependent DBM system in Ashdown PArk.

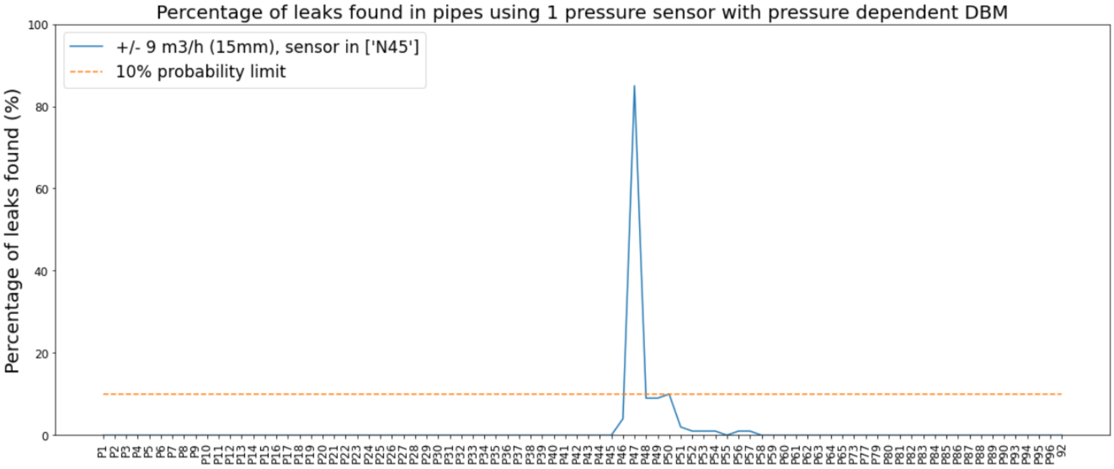

Figure A.58: Percentage of leaks found using one pressure sensor in a pressure dependent DBM system.

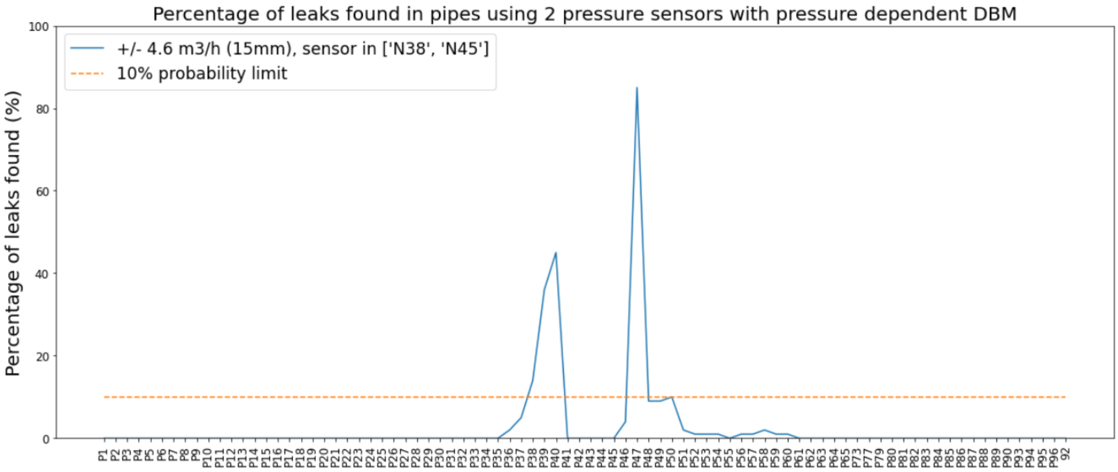

Figure A.59: Percentage of leaks found using two pressure sensors in a pressure dependent DBM system.

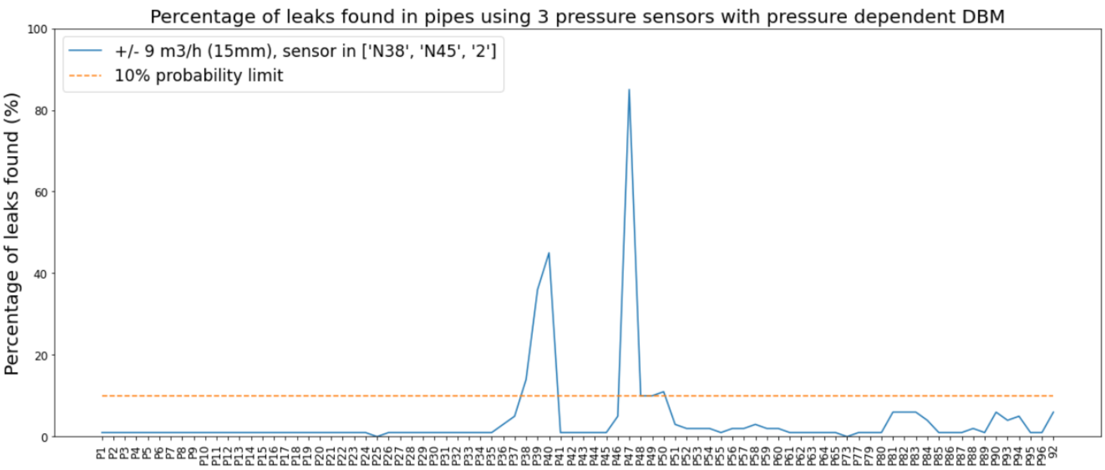

Figure A.60: Percentage of leaks found using three pressure sensors in a pressure dependent DBM system.

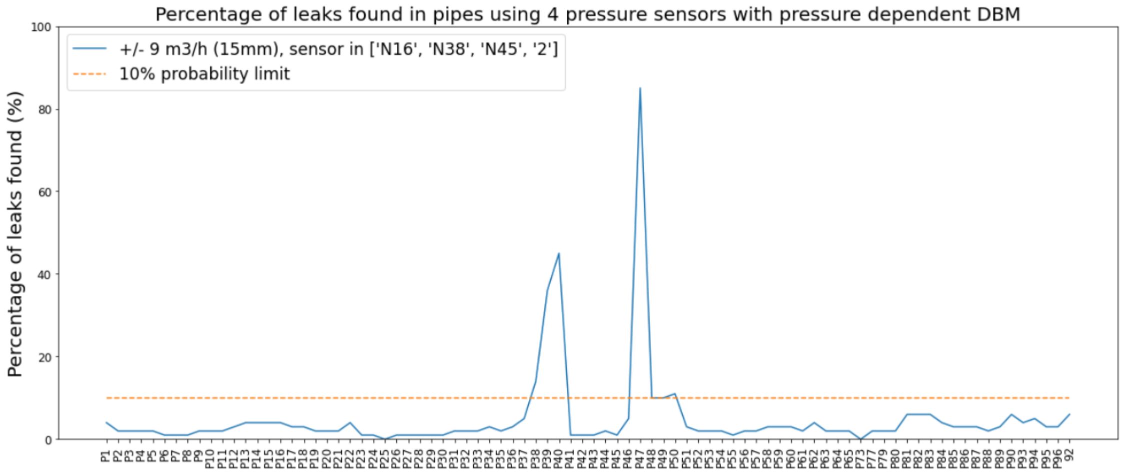

Figure A.61: Percentage of leaks found using four pressure sensors in a pressure dependent DBM system.

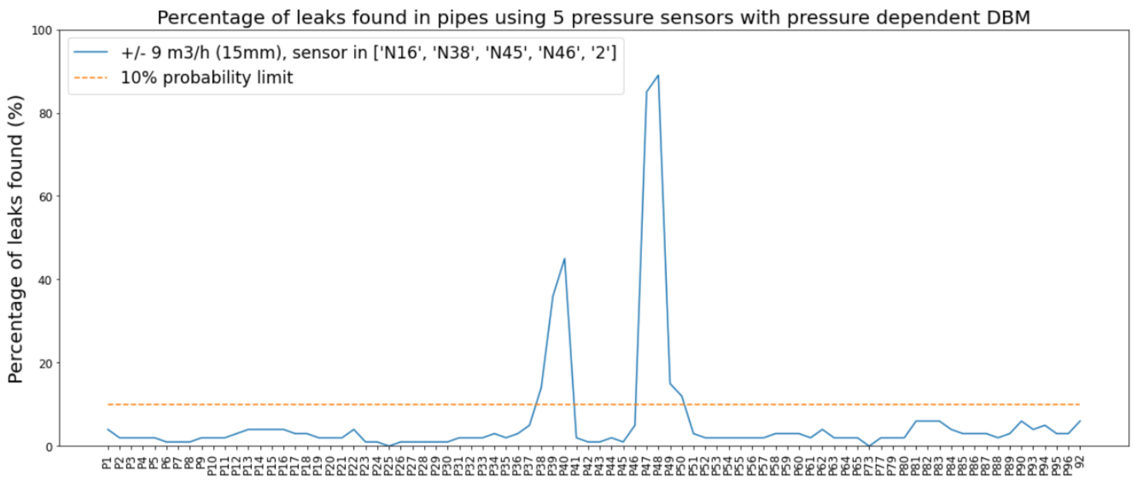

Figure A.62: Percentage of leaks found using five pressure sensors in a pressure dependent DBM system.

A.18 Detailed explanation for leak detection once a day

Earlier results show us, for each sensor allocation and leak size, a certain chance of a sensor finding a leak in a pipe. However, this is probability can not be translated one to one to the daily chances of finding a leak. For example, it is very likely that if the results above show a 1% chance of finding a leak, that this is a false alarm. This leak will than not be found once every 1000 minutes or once every 100 days, it will never be found. It is therefore important to translate the above probabilities to whether a leak can be found within a certain time frame. For this thesis, it is assumed that a leak should be found at least once a day, since it would otherwise be too likely to be a false alarm. This translation to a probability of once a day involves quite some stochastic assumptions and is therefore merely an estimation of what can be expected in reality.

Let us explain this translation by introducing a fictive example. For instance, a flow sensor in pipe 1 has a 60% chance of "detecting" a leak of 15mm in pipe 4 at 22:30. However, this percentage represents the possibility of detecting a leak at a single time step, making the detection vulnerable to reporting false leaks. By using a certain leak detection time, which is a time frame during which the leak detection system has to report a leak constantly, extreme demand deviations can be distinguished from actual leak detection.

In a different research Van Steen (2020), which modelled a DMA of 2825 households in the Netherlands, a leak detection time of 25 minutes was proposed to distinguish a demand deviation from a actual leak. The same leak detection time will be used in this thesis. Since the algorithm in this thesis distinguishes time steps of 10 minutes, a leak should be able to detected for 3 time-steps in a row.

So, let us continue with the example of the beginning of this section, which evaluated the chance of finding a leak at 22:30. Let consider the second time step (22:40). If the detectability matrix would not change, the changes of finding a leak would again be 60%, if the two sequential time steps would be uncorrelated. However, it is very likely that the flow or pressure measurements at two sequential time steps are correlated. So, if a measurement exceeds the alarm value at t=t, it is likely that it will exceed the alarm value at t=t+1 too.

A measure for correlation which is often used statistics is the Pearson correlation coefficient (r), which compares the strength and direction of two variables. This Pearson correlation can be calculated from the measurement data at the DMA inlet for the most important parameters. So, for leak detection with flow meters this would imply calculating the correlation coefficient between the average flow at the DMA inlet at t=t and t=t+1, written as $r_{flow,t+1}$ (and for the second time step between t=0 and t=2, written as $r_{flow,t+2}$). For the leak detection with pressure meters, these coefficients were calculated for the 5-percentile of the inlet pressure at same the time steps. Figure A.63 and A.64 show the zoomed in graphs of the observed variables, which shows that the graphs of t=t are very similar to the graphs of t=t+1. The computed correlation coefficients for the time steps ranged between 0.90 and 1.00 (which was actually 0.998, but rounded), as can be seen in Table A.11 and A.12. These high correlation coefficients show that there is a very strong correlation at the subsequent time steps for flow and pressure values at the DMA inlet.

Average weekly flow values	t = t+1	t = t + 2
t = t	$r_{flow,t+1}$ = 1.00	$r_{flow,t+2}$ = 0.99

Table A.11: Pearson correlation coefficient when comparing the weekly flow values of t=t to t=1 and t=t to t=t+2.

5-percentile inlet pressure values Friday	t = t+1	t = t + 2
t = t	$r_{pres,t+1}$ = 0.97	$r_{pres,t+2}$ = 0.90

Table A.12: Pearson correlation coefficient when comparing the 5-percentile pressures on Friday of t=t to t=1 and t=t to t=t+2.

So, if the detectability matrix in two sequential time steps is the same, the change that the second matrix measures a leak as well could be 60% (time steps are uncorrelated) or 100% (time steps are the same, so completely related) or somewhere in between. A different demand realization does not affect the measurements at the DMA entrance much, since all the flow passes through the inlet. The sensor which is added for the DBM how-

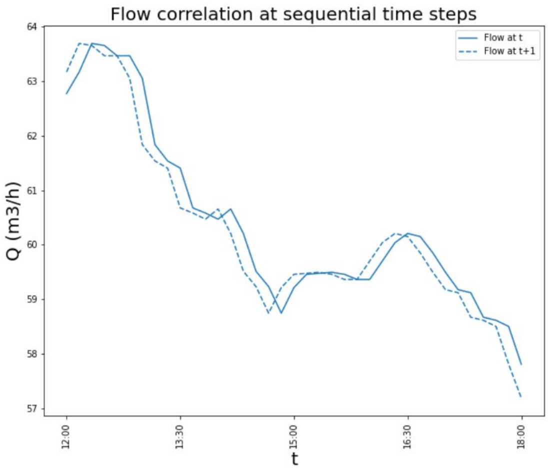

Figure A.63: Similarity between flow at t=t and t=t+1.

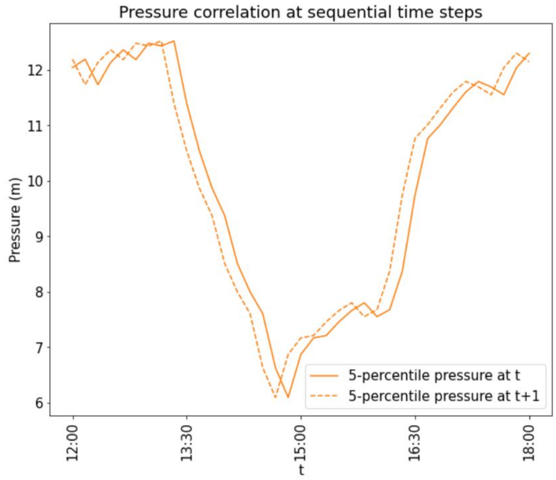

Figure A.64: Similarity between the 5-percentile pressure values on Friday for t=t and t=t+1.

ever is not located at the DMA entrance, making the flow and pressure measurements at the DBM sensor more sensitive to different demand realizations. Therefore, it is unlikely that the flow and pressure at the DBM sensor at t=t and t=t+1 is as strongly correlated as the measured flow and pressure at the DMA inlet. It is not completely uncorrelated either. Therefore, it is chosen to give a sequential time step the average probability of the correlated and uncorrelated option. This is shown in equations A.32, A.33 and A.34.

$$P_{sens:1,leak:4}(t) = 0.60 \tag{A.32}$$

$$P_{sens:1,leak:4}(t:t+1) = 0.60 \times \frac{r_{flow,t+1} + 0.60}{2} = 0.48 \tag{A.33}$$

$$P_{sens:1,leak:4}(t:t+2) = 0.60 \times \frac{r_{flow,t+1} + 0.60}{2} \times \frac{r_{flow,t+2} + 0.60}{2} = 0.38 \tag{A.34}$$

So far, the most of these principles can apply to both flow and pressure sensors. To clearly differentiate between the two, let us continue explain separately for flow and pressure sensors how it will be calculated whether a leak can be found once a day.

Flow sensors

So, with the formulae above it is possible to calculate the possibility that a leak is found in three consecutive time steps. As a next step, it is required to look deeper into the assumption that the detectability matrix in two sequential time steps is the same. Figure 4.9 has shown us that a monitoring system with flow meters detects the most leaks weekly between 21:40 and 05:20. So, let us assume that the Boolean 3D-matrix within this time frame remains similar. This leaves us with a time frame that consists of 46 steps (of 10 minutes). Of this consistent time steps, 44 distinct possibilities of three consecutive time steps can be made (t=0,1,2; t=1,2,3; etc.). The change that three consecutive time steps record a leak is 0.38 (equation A.34), so on average a leak is measured in three consecutive time steps 17 (=44*0.38) times a day. Now, it is possible to calculate a probability limit (p_{limit}), which would result in a leak that is being found once a day. This probability limit has to satisfy equation A.35. It was found that the probability limit for detecting leaks once a day with flow sensors was 0.079. Due to the high uncertainties that arose in estimating this limit probability, it is rounded up to 0.1 (10%). So, leaks which have a higher probability than 10% of being detected by flow sensors can be detected once a day.

$$p_{limit} \times \frac{r_{flow,t+1} + p_{limit}}{2} \times \frac{r_{flow,t+2} + p_{limit}}{2} \times 44 \geq 1 \tag{A.35}$$

The amount of leaks that can be found using 4 flow sensors, together with the probability limit (above which the leaks can be detected once a day) can be found in Figure A.65. By optimally allocating 4 flow sensors, leaks in 25% of the pipes in the DMA can be detected once a day.

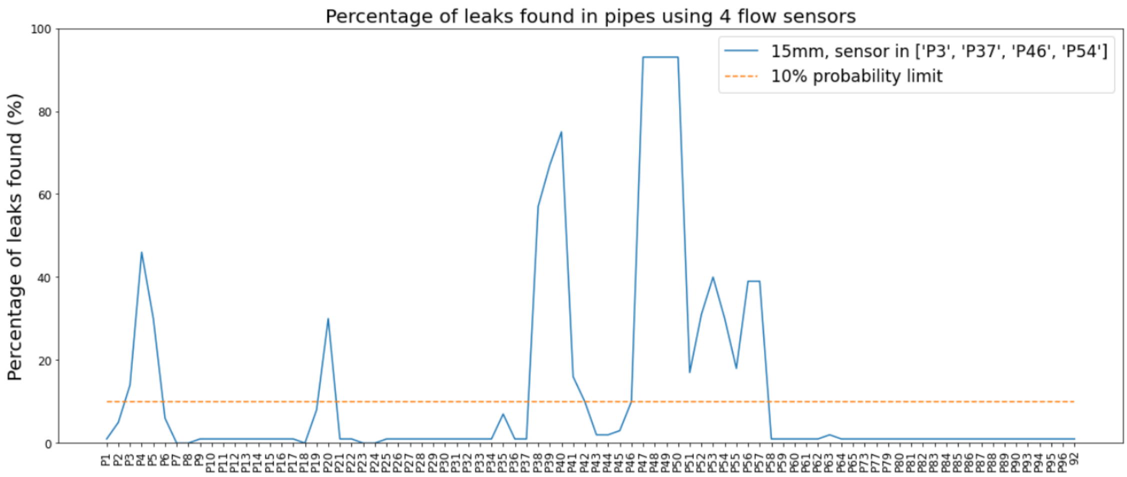

Figure A.65: Leaks that can be detected using 4 flow sensors and the probability limit.

Pressure sensors

Although the approach for calculating the chance of a leak occurring once a day is similar with using pressure sensors, there is one crucial difference. Whereas the flow sensor monitoring network is based on average weekly flow and pressure values, the pressure monitoring network uses the 5-percentile values for pressure at the inlet of the DMA to construct the artificial DBM's. Adding the chance of the occurence of the 5-percentile pressure to the chance of finding a leak in three sequential time steps decreases the chance of finding a leak significantly (from 60% to 1.8 %, shown in equation A.36).

$$P_{sens:1,leak:4}(t:t+2) = 0.05 \times 0.60 \times \frac{r_{pres,t+1} + 0.60}{2} \times \frac{r_{pres,t+2} + 0.60}{2} = 0.018 \tag{A.36}$$

Although the precision was lower, it was assumed that leaks were well detectable with pressure sensors on Friday between 00:00 and 06:30 and between 18:30 and 23:50. This results in 72 time steps, from which 68 distinct possibilities can be made of three consecutive time steps (t=0,1,2; t=1,2,3; etc.). Calculating the probability limit again, this time using equation A.37, resulted in a probability limit of 0.54, rounded to 0.55. This is a high probability which is hard to obtain.

$$p_{limit} \times 0.05 \times \frac{r_{pres,t+1} + p_{limit}}{2} \times \frac{r_{pres,t+2} + p_{limit}}{2} \times 68 \geq 1 \tag{A.37}$$

The amount of leaks that can be found using 4 pressure sensors, together with the probability limit (above which the leaks can be detected once a day) can be found in Figure A.65. By optimally allocating 4 pressure sensors, leaks in only 1% of the pipes in the DMA can be detected once a day. From this result it can be concluded that a monitoring system with pressure sensors as designed above is not useful. Therefore, the possibility of using a pressure dependant DBM will be discussed in the next chapter.

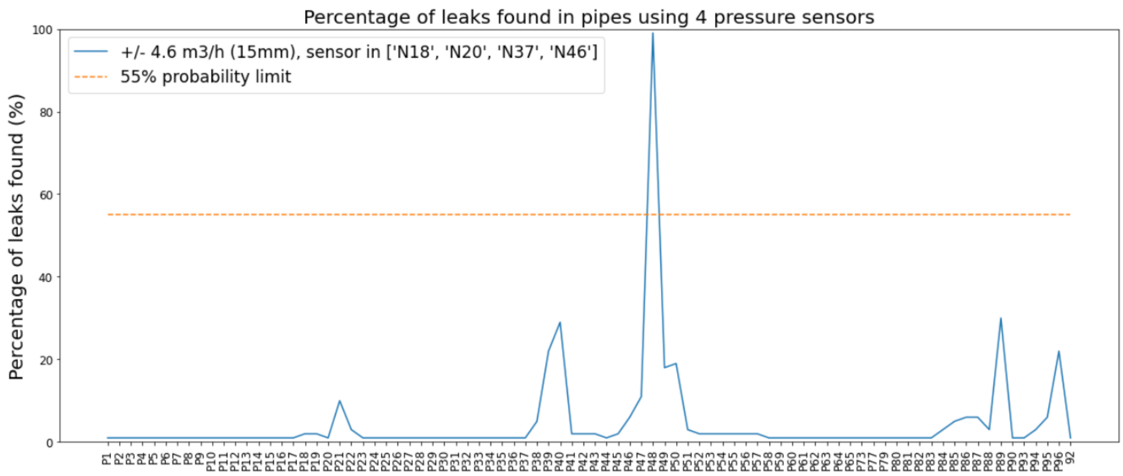

Figure A.66: Leaks that can be detected using 4 pressure sensors and the probability limit.

A.19 Figures of important factors for sensor allocation

Important factors when allocating flow sensors

Figure A.67 shows the elevations of all pipes. The chosen pipes have an average elevation, so the elevation will probably have a limited effect on the optimal location for flow sensors.

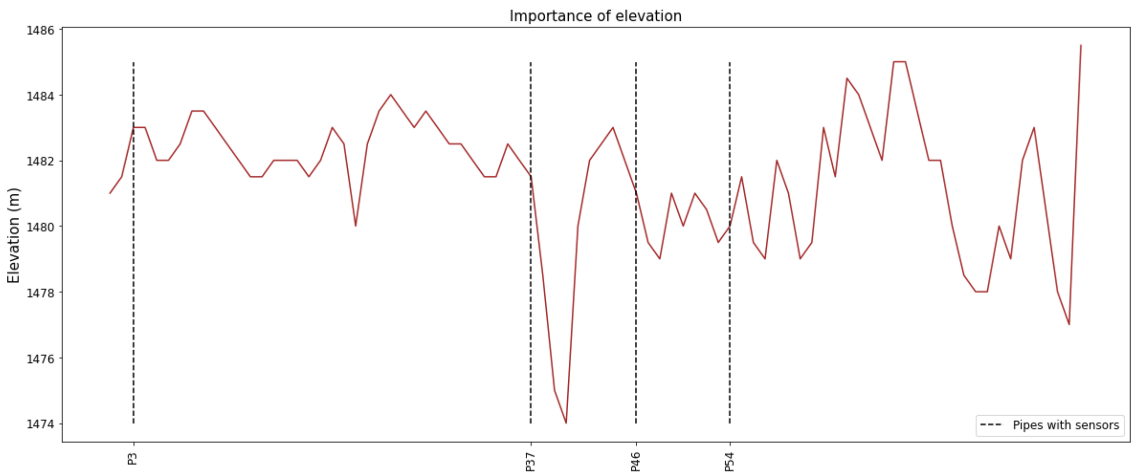

Figure A.67: Influence elevation on optimal allocation flow sensors.

Figure A.68 shows the diameters of all pipes. The chosen pipes have an average diameter, so the diameter will probably have a limited effect on the optimal location for flow sensors as well.

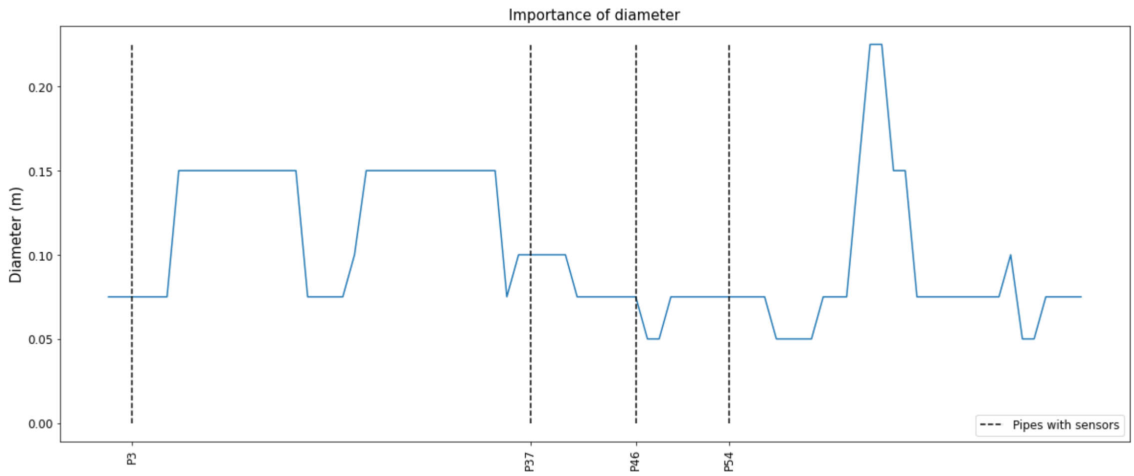

Figure A.68: Influence diameter on optimal allocation flow sensors.

Figure A.69 shows the average flows in all pipes. The chosen pipes have a relatively low flow, but these low flows are not very unique compared to other pipes. The average flow in the pipes will therefore probably have a limited effect on the optimal location for flow sensors as well.

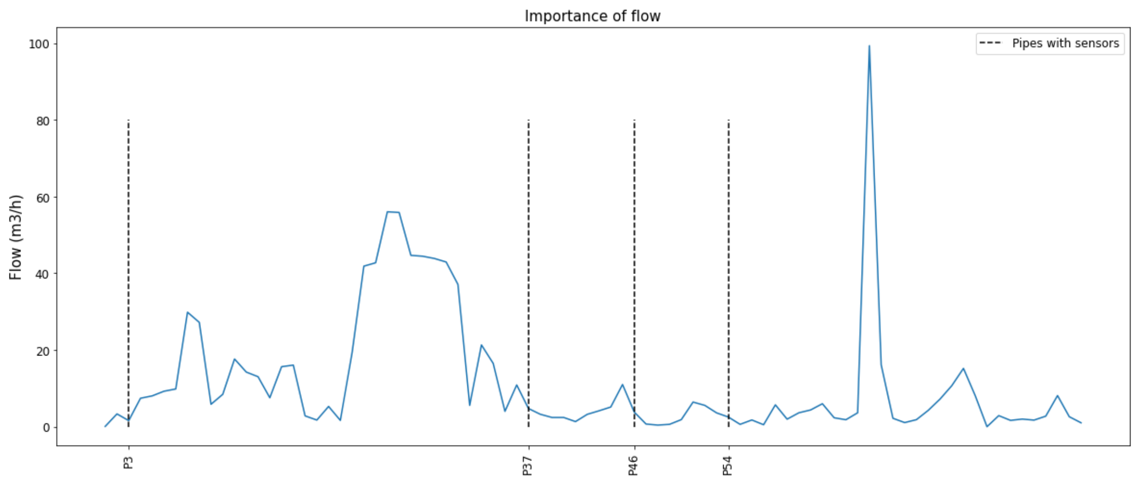

Figure A.69: Influence average flow on optimal allocation flow sensors.

Important factors when allocating pressure sensors

For determining the important factors for allocating pressure sensors, the characteristics of only two nodes were compared to the other nodes in the network. These two nodes resulted of optimally allocating pressure sensors with a pressure dependent DBM system. More than two nodes had little influence on the coverage of the network. Therefore, only two nodes were taken into account.

Figure A.70 shows the elevations of all nodes. In the case of optimally allocating pressure sensors (with a DBM system), the chosen nodes have very low elevations. Therefore, it seems advisable to place pressure sensors for a monitoring system on locations with low elevations.

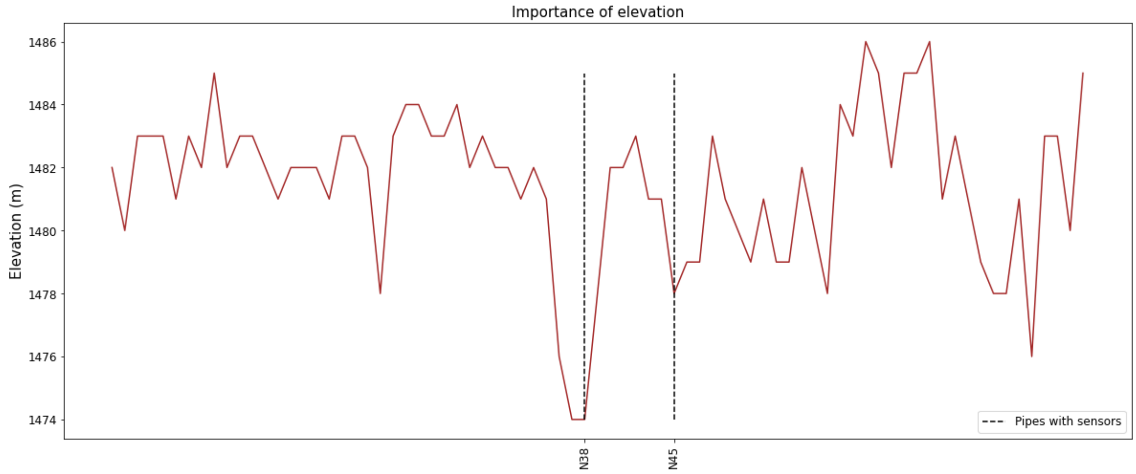

Figure A.70: Influence elevation on optimal allocation pressure sensors with a pressure dependent DBM.

Figure A.70 shows the pressures at all nodes. This pressure is influenced by the elevations as well, since low elevations usually have higher pressures. Therefore, the importance of pressure should be seen in line with the importance of elevation, as nodes where high pressures occur seem very suitable for allocating sensors for a pressure monitoring system.

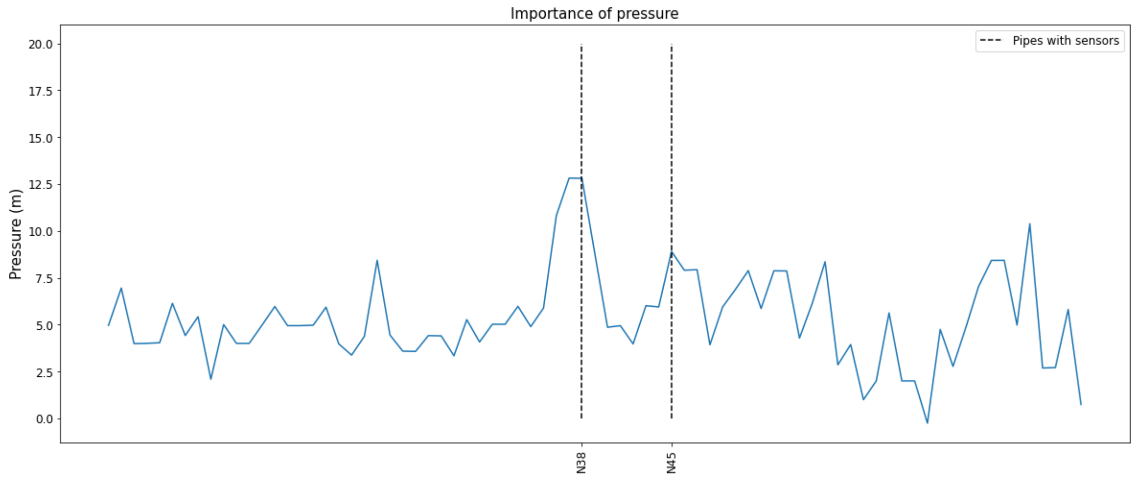

Figure A.71: Influence pressure on optimal allocation pressure sensors with a pressure dependent DBM.

A.20 Details interviews Harare

Inhabitant Budiriro

This interview was held through a WhatsApp-call. Therefore the answers are not always the exact things said, but as remembered and written down during and after the call.

Q: Where do you live in Harare?

In Budiriro

Q: How would you compare an area such as Budiriro to an area like Ashdown Park?

Budiriro can be described like a township/ghetto. It is a high-density area. Ashdown Park is more of a low-density area, where white people used to live.

Q: What is your opinion about the drinking water services of the municipality?

Not very positive. The water services are not like the old days anymore. This is due to the bad economical situation and corruption. The first 10 years after independence, we had clean water from the tap. After these 10 years the situation got worse.

Q: How would you describe the quality of the drinking water?

The water is not clean enough for drinking. If you fill the bathtub and let it rest, you can see the dirt in the water settling after a few hours. The water can be used for washing or for flushing the toilet, but not for drinking.

Q: What sources for drinking water do you use?

Almost everyone in Budiriro fetches drinking water from a bore hole. You pay a small fee to the owner of the bore hole. They charge 1 USD/month to fill 3 buckets per day. It is cheaper than water from the utility. People place buckets and pans under their roof as well when it rains. This is free clean water.

Q: How many days do you receive water and how do you store water?

Usually there is no water a few days per week. Sometimes there is water during the night, but not during the day. We store water by filling the bathtub. People with money can buy a big storage tank to store water. Sometimes they pump their own groundwater to the storage tank.

Q: Are there other things that are good to know regarding the drinking water situation in Harare?

Harare faces a lot of problems with the sewer system as well. Wastewater at some points can flow into the street and cause illness. Especially the area Kambuzuma has trouble with this issue.

MSc-student university of Bulawayo

This interview was held through a WhatsApp messages. The questions are therefore somehow elaborate at times.

Q: For a financial analysis I am trying to compare different options of the people to supply themselves with drinking water. At this moment I know that a m3 of water from the piped system costs ZWL 90, but that the quality is too bad for drinking water (at least in Budiriro). People can also collect water from a constructed wellin the neighbourhood, being able to take 3 buckets of drinking water for 1 USD/month. Would you agree with this information in Harare? And if people were to buy a bottle of water in the store, how much would they need to pay?

Yss.....but for drinking water it depends on the number of people....1 bucket of water costs about 7.50 local currency and the price for a water bottle in the shop its going for 30 to 50 dollars local currency thats 500ml bottle. From a Budiriro person

Q: Furthermore, I was curious whether people will notify the water utility when they see a leak in the drinking water distribution system. Or that they do not trust the municipality in repairing the leak and therefore do not take the effort of notifying the municipality of the leak.

They do notify the municipality, whether it quickly response or not that is the decision of the council but people do report.

Q: And another final question: if people have storage at home, do they have a system that automatically fills their storage? Or do they fill it themselves directly from the tap?

Some uses pumps to fill up from the tap, it's ain't automatic

A.21 Details cost analysis

Details for the prices that are mentioned in the cost analysis can be found below. Every cost is expressed in US Dollar (USD), followed by the original costs in the local currency between brackets.

Investment costs

The investment costs are split into flow meters, pressure meters, supplementary materials and costs for labour during the installation of the system.

Flow meters

The flow meter should be installed to pipes with a diameter of 75mm to 100mm, as there were the physical properties of the optimal sensor allocation analysis.

Harare: The flow meter installed at Ashdown Park had cost around 2000 USD, including the materials for communicating. The meter was attached to a pipe of 200mm. Using an exact cost-estimation of the communication material (follows below), this would estimate the costs of this flow meter at 430 USD. Since it measures pulse-wise, each flow rate is a multitude of 4 m3/h, it should be a mechanical meter, most likely to measure only in one direction. Nairobi: Earlier, a DN100 (so suitable for diameters of 100mm) ultrasonic flow meter was purchased in Nairobi for 4443 USD (482,906 Kshs). A DN80 ultrasonic flow meter was purchased for 3541 USD (384,943 Kshs). Amsterdam: Amsterdam has experience with using two different brands of flange flow meters for pipes up to 200mm. The first brand is Flostar, which uses a single jet meter which can measure in two directions. It costs 846 USD (699 EU) for a 100mm pipe and 729 USD (602 EU) for a 80mm pipe. Woltex is more cheap, but it needs a laminar flow and is sensitive to changes in flow direction. The costs for using a Woltex flow meter are 358 USD (296 EU) for a 100mm pipe and 329 USD (272 EU) for a 80 mm pipe.

The flow meters in the monitoring system should be able to perform well when flow directions change, as this might occur within the DMA. Therefore, the ultrasonic flow meters from Nairobi and the single jet meter from Amsterdam seem the most suitable options. This leaves the flange meters from Flostar (from Amsterdam) as the best financial option.

Pressure meters

Harare: Harare had previously purchased a pressure transducer (pressure meter) for 770 USD. Nairobi: Nairobi does not use pressure measurements in its system (personal com, Mugo). Amsterdam: Not yet available.

The costs for the pressure transducer that has been purchased in Harare shall be used for this thesis.

Supplementary materials

The supplementary materials include a protection chamber, a data transmitter, connection cables, a USB and communication software.

- *Protection chamber*: When installing the flow meter in Ashdown Park earlier, a concrete chamber had been built around the meter to protect it from theft. The costs of this chamber were 2000 USD. In Nairobi, the civil works for constructing concrete protection chambers had costed 17,137 USD (1,862,782 Kshs). It was not clear however, whether these costs were for multiple or a single concrete chamber. Therefore, the costs of a concrete chamber in Harare were used for this costs analysis.
- *Data transmitter*: The data transmitter that is used in Harare to store the data online costs 1350 USD.
- *Connection cables*: The meter is connected to the data transmitter with connection cables. This cables cost 54 USD in total.
- *USB and Communication software*: The USB cable and communication software that was used in Harare to read the measurements on location costs 166 USD.

Visuals of the supplementary materials can be found in appendix A.22.

Transport costs:

The transport costs that were charged earlier in Harare for importing a pressure logger and all the supplementary materials through air from South Africa were 543 USD. This was approximately 20% of the total costs. Therefore, the transportation costs of the monitoring system shall be estimated at 20 % of the material costs.

Labour costs

The estimated workforce to install a meter ranged from a team of 6 people working 5 days to 10 people working 10 days. Since salaries, especially in Zimbabwe, are quite low, it was chosen to hire a team of 10 people for 10 days. The workforce will consist of a supervisor who earns 0.25 USD/hour (ZWL 80/ hour), a foreman who earns 0.22 USD/hour (ZWL 70/ hour) and eight other attendants which earn 0.16 USD/hour (ZWL 50/ hour). This results in a total costs of labour of 140 USD per meter.

Variable costs

The variable costs consist of maintenance and monitoring and the costs for leak repair.

Costs of maintenance and monitoring

In Harare, a fee of 121 USD/year was paid to use the online platform that stores the measurements. Other costs of monitoring would be to hire a person which will be in charge of operating the monitoring system. An appropriate salary of such a person would probably be comparable to the salary of a foreman in leak repair, so 0.22 USD/hour (ZWL 70/hour) Magedi MacDonald (2020). Assuming this would be a full-time function (40 hours/week, 46 weeks/year), this person would earn 405 USD/year. This is extremely low compared to European standards, since the salaries in Zimbabwe are very low.

In Nairobi, a one-off fee of 12.42 USD (1,350 Kshs) was paid for the activation of a per post paid sim card. However, it was not clear if the measurements where stores on an online platform or directly send to NCWSC. Therefore, the maintenance costs in Harare will be used as a guidelines for the maintenance costs of the platform.

Repairing costs

The most recent repair of leaks in Harare involved placing two couplings (brand: Viking Johnson) of 33 USD each. Furthermore, also a new pipe might be useful. Pipes in Zimbabwe are usually sold for around 35 USD for a 6m PVC pipe **?**. Of course, this price depends on the pipe diameter and its material. It was estimated that it required four people to repair the leak, one supervisor, one foreman and two attendants. Using the same salary as was used for the installing of the sensors and assuming that the leak can be repaired in one day, this amounts to a labour cost of 6.32 USD per leak. At last, fuel is needed to travel towards the leak. A proper estimation of the fuel costs would be 8 USD, since fuel is sold for 1.20 USD/L, travel distance can be estimated at 40 km and the car usually drives 6 km on 1 L of diesel. The total repair costs will therefore be approximately 115 USD per leak.

A.22 Supplementary materials

Visuals of the supplementary materials and their functions are given below.

Data transmitter

The transmitter that was used in Harare to transmit data is the Cello 4S data transmitter. It can be seen in the smart measuring system in Figure A.72 and the individual machine is shown in Figure A.73. The machine was bought in South Africa for 16.700 R, which was 367.400 ZWL at the time of this thesis.

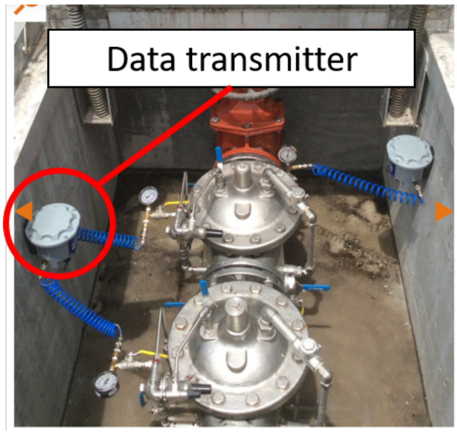

Figure A.72: A data transmitter in a smart measuring system.

Figure A.73: The Cello 4S data transmitter.

A pressure transducer that was used in Harare can be seen in Figure A.74. It had costs 11.500 R, which converted to 253.000 ZWL during this thesis.

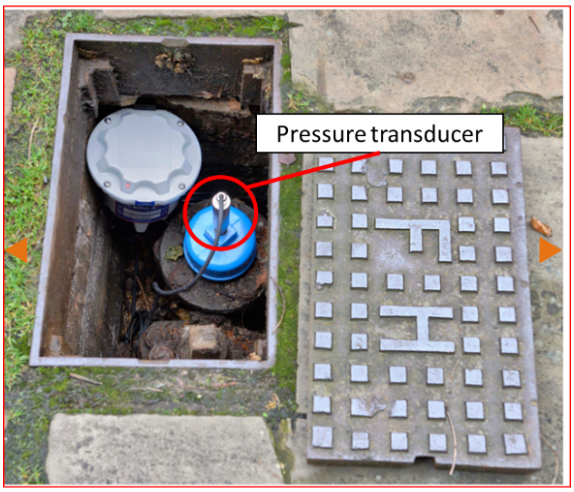

Figure A.74: A pressure transducer.

The connection between the Cello 4S and the pressure transducer (Figure A.75) had cost 805 R, so 17.710 ZWL.

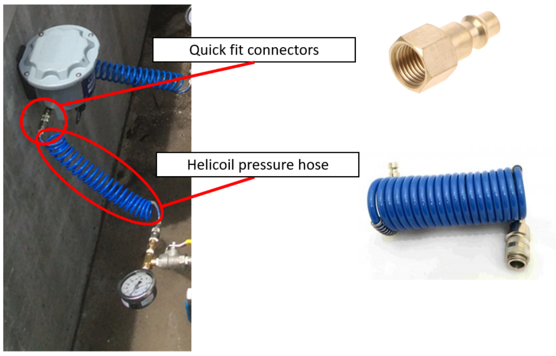

Figure A.75: Connection cable smart meter.

The USB cable and communication software (Figure A.76) had cost 2480 R, so 54.560 ZWL. In a smart meter system it can be considered as an “extra feature” since the data is already transmitted to an online platform.

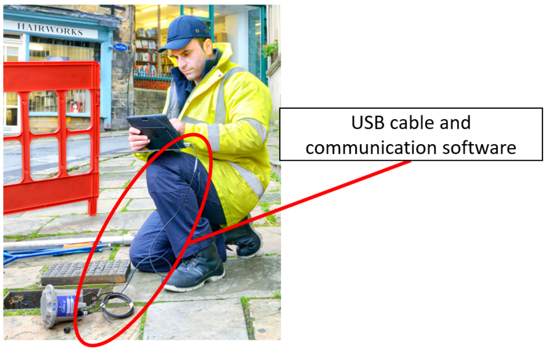

Figure A.76: A USB cable and communication software.

A.23 Cost-savings graphs for monitoring system with pressure sensors

The graphs below show the costs and savings for using a monitoring system with a certain number of pressure sensors and a pressure dependent DBM. The costs and savings for detecting leaks with a size of 9 m3/h can be seen in Figure A.78. This figure shows that a leak of 9 m3/h is unlikely to be find within 10 years, resulting in no savings. If the detection system would be improved to find leaks of 6 m3/h, a payback period of five years could be reached when applying 3 pressure sensors (Figure A.78).

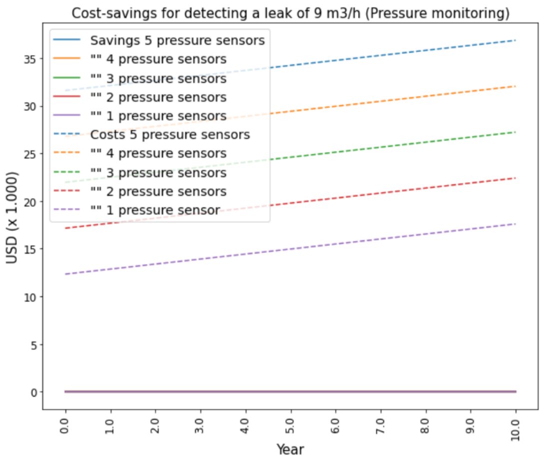

Figure A.77: The costs and savings that result from detecting leaks of 9 m3/h with pressure sensors and a pressure dependent DBM.

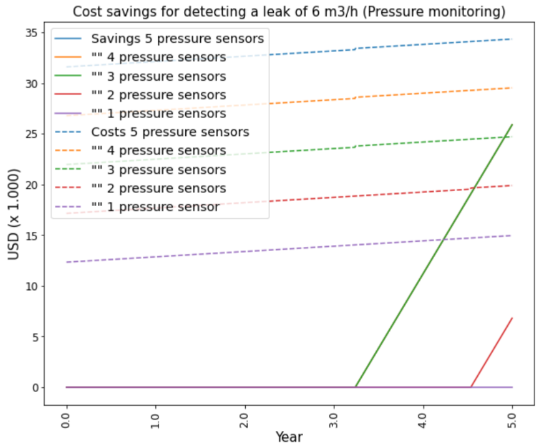

Figure A.78: The costs and savings that result from detecting leaks of 6 m3/h with pressure sensors and a pressure dependent DBM.