MCWD
Morgan County Water District
Established Since 1992

June 9, 2022

Linda Bridwell Executive Director
Kentucky Public Service Commission
P.O. Box 615 Frankfort, KY 40602-0615

Re: Case No. 2020-00386

Dear Ms. Bridwell

Pursuant to the Commission's Order of June 9, 2021 in the above-referenced matter, enclosed for filing is Morgan County Water District's annual water loss report. This letter and the enclosed Report are true and accurate copies in paper medium of the electronic version of the documents uploaded to the Public Service Commission's Electronic Filing System this day.

In accordance with 807 KAR 5:001, Section 8, I certify that Morgan County Water District's June 9, 2022 electronic filing of this Report is a true and accurate copy of the same document being filed in paper medium; that the electronic filing was transmitted to the Commission on June 9, 2022 by electronic means in this proceeding.

Sincerely,

Shannon W. Elam
General Manager

Enclosures

1009 Hwy 172, West Liberty, KY 41472 (606) 743-1204 Phone (606) 743-9585 Fax

MCWD

Morgan County Water District

Annual Water Loss Report

2022

PSC Case # 2020-00386 and 2021-00206

Annual Review

Morgan County Water District has made progress in identify and repairing water line leaks in its distribution system in the past year. We have been able to purchase new equipment and have completed the first water line replacement project design to replace water lines that have historically have been a leak problem.

We have been able to add zone meter pits at each of our water tanks and with the purchase of a portable flow meter; we have started to identify leaks that in the past we would not have been able to locate.

We have been able to get additional training for our Leak Detection Specialist through Kentucky Rural Water Association and McKim & Creed

Actual Water Loss Progress

Please see below a three-year comparison of purchase water below. We have been able to purchase less water and at the same time sale more water to our customers.

		(In Millions)
2019		
Water Purchased	264,208	
Water Sold	113,389	
2020		
Water Purchased	243,851	(Change) 20,357 less purchased
Water Sold	115,696	(Change) 2,307 increased sold
2021		
Water Purchased	235,275	(Change) 8,576 less purchased
Water Sold	120, 233	(Change) 4,537 increased sold

Estimated Water Loss Progress

We are making progress toward our goal of 15% water loss and we are on track this year to reduce water purchased by another 10 million gallons in 2022.

Actual Expenditures from Surcharge Account

10/14/2021	Nesbitt Engineering	Corrective Action Plan	$14,702.50
10/29/2021	McKim and Creed	Phocus3 Loggers	$26,995.00
2/22/2022	Beartraxx	Zone Meter Pits	$30,480.00
5/20/2022	Beartraxx	Zone Meter Pit	$5,080.00 (not approved yet)

Total Spent $77,257.50

Estimated Expenditures from Surcharge Account

We will continue to add zone meter pits where we need them throughout our distribution system. We have also been researching the use of pressure monitoring for water leaks. We have had a company come in and do a demo on this technology. We are also looking at implementing Lean Six Sigma for problem solving and leak detection. Hydrant locking system has also been researched to deter theft of service

We have could some research papers regarding pressure monitoring and Lean Six Sigma in Leak Detection.

Additional Zone Meter Pits	$30,000
Pressure Monitoring	$50,000
Lean Six Sigma	$10,000
Hydrant Locks	$30,000

Macrothink Institute™

Environmental Management and Sustainable Development
ISSN 2164-7682
2017, Vol. 6, No. 2

A Six Sigma Approach to Water Savings

Ryland Cairns (Corresponding author)

Water Services Department, Interserve

Whale Way, Mount Pleasant Complex, Falkland Islands, FIQ 1ZZ

E-mail: ryland.cairns@gmail.com

Michael MacPherson

Water Services Department, Interserve

Whale Way, Mount Pleasant Complex, Falkland Islands, FIQ 1ZZ

E-mail: Michael.macpherson@interserve.mod.uk

Received: March 20, 2017 Accepted: May 9, 2017

doi:10.5296/emsd.v6i2.10953 URL: https://doi.org/10.5296/emsd.v6i2.10953

Abstract

The purpose of this paper is to explore the potential of a six sigma approach to reducing water losses through a combination of water efficiency and leak detection on a private distribution system. The paper takes the form of a case study that investigates the implementation of water reduction strategy across an estate with 26 miles of potable water pipe and over 200 facilities. This incorporates methods developed in the water industry such as water loss reduction and water demand management. The paper demonstrates that large water savings could be made through adoption of a six sigma approach. The approach has the potential to be applied to a wide range of situations including sites with limited technology. This case study provides a useful source for Facilities Managers involved in the management of utilities to determine suitable water saving approaches and strategies for large estates with private water distribution networks.

Keywords: Water loss, Six sigma, Water efficiency, Water management, Private water supplies, Facilities management

1. Introduction

There is an increasing global awareness that water scarcity is one of the emerging issues of this century (Morais and Teixeira de Almeida, 2007). The demand for this precious commodity is continuingly escalating through increased population and changes in precipitation patterns brought on by climate change (Bates *et al.* 2008). Effective utilisation of water through the reduction of wastage either within distribution (leaks) or at the point of use (water inefficiency) remains a key problem for both developing and developed countries throughout the world (Morais and Teixeira de Almeida, 2007).

Leakage from water distribution systems has been a widely research topic due to the large percentage of losses encountered even on the smallest systems. Annually the global water sector processes 32 billion cubic meters of water that is lost through leaks (Kingdom *et al.*, 2006) accounting for 9%-43% of the nations total production (Lai, 1991). In addition to the physical loss of water there are other deleterious impacts relating to energy consumption, chemical consumption potential health and environmental risks (Colombo *et al.*, 2009) with an associated total financial of cost of US$15.6 billion (Kingdom *et al.*, 2006). One such issue is the energy consumption associated with treating and pumping leaked water (Colombo and Karney, 2002). This is thought to account for 5-10 billion kWh of power generated a year, in the states alone (AWWA, 2003). Water hygiene issues are also associated with burst or leaks through the intrusion of contaminated groundwater. Karim *et al.*, (2003) demonstrated that a high percentage of soils and potentially harmful pathogenic organism such as coliforms (58% and 70% of water and soil samples respectively) and faecal coliforms (detected in 43% and 50% of the water and soil samples respectively). During low pressure events on a burst line it is possible for these contaminants to be drawn in to the system.

Water demand management (WDM) deals with developing and implementing strategies aimed at influencing water demand by improving efficiency to reduce average water consumption (Brooks, 2006). WMD strategies focus on providing tools, mechanisms and knowledge to enable users to reduce water consumption. Such tools include high efficiency water fixtures and appliances that have demonstrated to support least cost planning strategies for water conservation and are thus a good starting point prior to high cost capital works (Stewart *et al.*, 2010). The provision of knowledge aims to influence attitudes and behaviours which can have a significant effect on water consumption (Willis *et al.*, 2011).

An increased focus on sustainability performance from a range stakeholders has meant that companies must also adapt to support this new core activity (Price *et al.*, 2011; Garza-Reyes, 2015). The built environment is responsible for half of the UK's water consumption (BIFM, 2007) thus sustainability targets are likely to include water usage. It has been reported that UK private sector could realise savings in excess of £3.5 billion a year through the adoption of water efficient practices (Environment Agency, 2012). These practices are also being brought in to the Public Sector through the introduction of the Greening Government Commitments. This has seen the Ministry of Defence (MOD) introduce water reduction targets as part of the new regional prime contracts (FM World, 2013).

To reduce the likelihood of failure the utilisation of frameworks to support improvement initiatives is deemed critical (Oakland and Tanner, 2007). One such framework is Six Sigma which has gained recognition as a quality improvement methodology within the manufacturing and service industries (Basu, 2004). It has been applied successfully to the utilities industry (Inozu *et al.*, 2006; Pheng and Hui, 2004) and can be used to drive sustainability initiatives (Zhang and Awasthi, 2014; Banawi and Bilec, 2014; Cherrafi *et al.*, 2016a; Cherrafi *et al.*, 2016b). Although other frameworks, such as lean, have been proposed to improve sustainability performance it is believed that their integration has not assisted organisations achieve peak performance, unlike that of Six Sigma (Banawi and Bilec, 2014).

This case study investigates a water saving strategy implemented within an MOD estate on the Falkland Islands using a six sigma approach.

2. Case Study

2.1 RAF Mount Pleasant Complex

Mount Pleasant Complex (MPC) is operational British Military theatre situated in the Falkland Islands with a primary mission statement to deter and defend against foreign aggressors. To enable the MOD to focus on their core activities infrastructure service provision, such as power generation and water production, is outsourced.

Water services cover all aspects of water and wastewater treatment. This involves the management of a treatment plant and distribution system in line with The Private Water Supply Regulations. The task inherently involves the metering of water consumption around the estate.

Following an increase in consumption over the years as well as the introduction of environmental targets there has been an increased emphasis on the management of leaks and water efficiency.

2.2 Water Consumption

Over the years, the activities and population of MPC have remained fairly constant. During this time, there has been a steady 27.13% increase in water consumption from 25,339m^3 in 2002 to 32,216m^3 in 2009 (figure 1). This has ranged from a mean consumption of 713m^3/day in May 2003 to 1270m^3/day in June 2009. Despite the repair of a major leak in July 2009, consumption continued to rise.

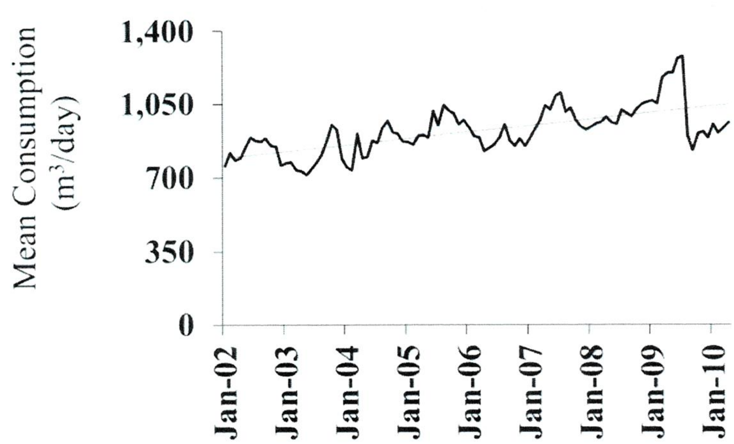


Figure 1. Site Consumption

Average domestic water consumption in the UK has been estimated at 150l/Person/day (Waterwise, 2012). On MPC total water consumption has ranged from 356-635l per person per day from 2002 to 2009, roughly two to four times in excess of that in the UK, however it should be noted that this includes non-domestic water which is used in centralised kitchens, boiler house operations and vehicle wash-down equipment. Due to metering arrangements it is not possible to differentiate between the two.

Under previous Environmental targets, the UK "Greening Government Commitments" the MoD has been set targets of reducing water consumption by 7% on the estate by 2016 relative to 2010/2011 levels.

3. Case Study

"Six Sigma is an organized, parallel-meso structure to reduce variation in the organizational processes by using improvement specialists, a structured method, and performance metrics with the aim of achieving strategic objectives" (Schroeder et al., 2008). Using a data driven improvement methodology Six Sigma aims to identify and significantly reduce operational defects and variability to improve process performance (Zu *et al.*, 2008; Laureani *et al.*, 2009). The six sigma methodology that was used to guide implementation was the five phase DMAIC cycle (define, measure, analyse, improve and control) (Bergman and Klefsjo, 2003) whereby the initial problem is defined prior to the utilisation of various tools to measure, analyse and seek the root causes which can be removed and controlled to ensure performance (Cherrafi *et al.*, 2016a).

In this case study the process problem which was being addressed was water loss and water inefficiencies. Six Sigma tools were used alongside standard water industry approaches to aid with the measurements and analysis.

3.1 Define

The charter of the six sigma project was developed by the Water Services Manager. The business case was based on the fact that water demand was on the verge of exceeding supply across MPC, despite efforts to optimise water processes to maximise the inputs (Cairns *et al.*, 2012). The situation had led to excessive chemical consumption, increased pumping costs and a twofold increase in operator overtime. Due to the age of the infrastructure, water demand was likely to be excessive as a result of large leaks and inefficient water fittings.

The process improvement tool SIPOC; **s**uppliers, **i**nputs, **p**rocess, **o**utputs, and **c**ustomers (Table 1) was used to define the system to allow for better targeting of resources. The SIPOC revealed that the process of water delivery could be split into three distinct levels within the system (main feed, distribution and facilities). This identified that the main goal of reducing overall water consumption could be achieved through a combination of leak reduction techniques on the main distribution system and a water demand management approach within the facilities.

Table 1. SIPOC of water delivery

Suppliers	Inputs	Processes	Outputs	Customer
Water Services	Potable Water	Pumping of water *Main feed → distribution → facilities*	Supplied Water	MPC inhabitants

3.2 Measure

An array of measurement techniques were formulated to cover each of these separate areas and provide a holistic picture of water consumption, losses and inefficiencies, taking into account the limited technology onsite (Figure 2).

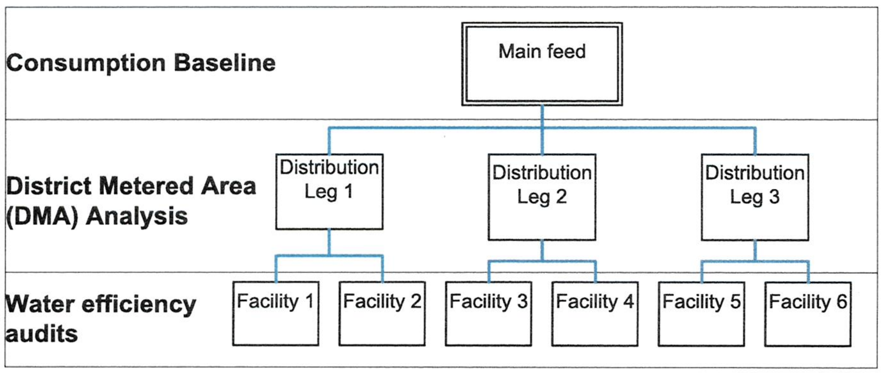


Figure 2. Measurement techniques at various levels of the distribution system

The first stage of the measurement process involved determining the base line consumption figures for the site on a day-to-day basis. This included overall consumption for the site (Figure 1) and minimum night flow (MNF). MNF measurements provide a better indication of leakage on the system compared to daily readings since actual demand is at its lowest during night time (0100hrs-0300hrs) making the leakage component the highest percentage of flow (Morrison, 2004).

The second stage of the measuring process involved the setting up and monitoring of District Metered Areas (DMAs). A DMA is specifically defined area of a distribution system in which the quantities of water entering and leaving the district can be metered. Flow monitoring (particularly MNF) can be used to calculate the level of leakage within a district providing an indication of levels of leakage within different areas and where resources should be focused (Morrison, 2004)

Flow measurements from the final meters on the DMAs provided an indication of consumption within individual facilities and could be fed in to the final stage Water Efficiency Audit. The audit used a standardised form to determine if actual consumption was inline with expected consumption based on staffing levels and fittings within the facility (Affinity Water, 2010). Analysis of night time readings supported this by providing an indication of leaks within individual facilities.

3.3 Analyse

Results from the DMA analysis illustrated that the high site consumption was the result of one particularly problematic distribution leg comprising 4 DMAs.

DMA analysis of this leg (Table 2) showed that 77.28% of the total water was lost to leakage, accounting for 23% of total consumption for the site. The worst areas were DMA 2 and DMA 4 which lost roughly 188m^3 and 24 m^3 a day respectively. DMA 3 and 1 experienced comparatively lower calculated losses at 1 m^3 and 2 m^3 respectively.

Table 2. DMA analysis based off mean daily consumption figures (January 2010)

	Expected Consumption (L)	Actual (L)	Losses (L)	Losses (%)
DMA 1	278,137	276,616	1,521	0.55
DMA 2	263,742	75,256	188,486	71.47
DMA 3	14,484	13,419	1,065	7.35
DMA 4	58,484	34,602	23,882	40.83
		Total	214,954	77.28%

In addition to the high percentage of water loss through leaks on this leg the water efficiency audits revealed that a number of metered facilities experienced excessive consumption (Table 3) totalling 10 a m^{3}day between 2-3 . In all cases, there was significant flow (up to 150 litres hour) during the periods of 0100-0300hrs. Water efficiency audits revealed that it was a faulty urinal flow control devices at the facilities with meters 1, 37 and 40, whilst an external water leak was the cause of high water consumption at meter 6.

Table 3. Water Efficiency Analysis (August 2010)

	Expected daily Consumption (L)	Mean daily consumption (L)	Water consumption per person per day (L)	Potential Losses
Meter 1	1,600	4,635	144	3,035
Meter 6	400	3,130	98	2,730
Meter 37	1,600	3,625	113	2,025
Meter 40	500	2,833	88	2,333

3.4 Improve

Corrective action was prioritised as per (Figure 3). Leaks were located using a combination of visual inspections of the lines, and valve pits in dry conditions and step analysis. Faulty urinal flow monitors were corrected, monitored and replaced if broken. In terms of DMA analysis (Table 4) this lead to a 251m^3 reduction in consumption along the leg and a drop in the water losses from 77.28% to 17.01%. The greatest improvements were realised following work to DMA 2 and DMA 4.

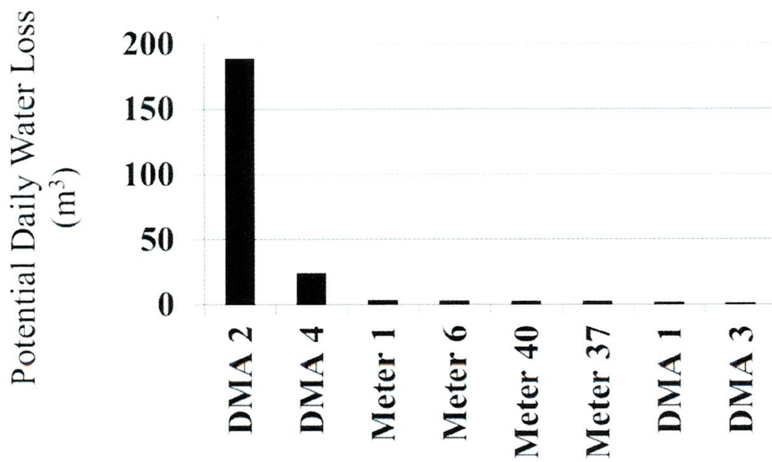


Figure 3. Work Prioritisation based off potential water losses

Table 4. DMA analysis December 2011

	Expected Consumption (L)	Actual (L)	Losses (L)	Losses (%)
DMA 1	28354	25442	2912	10.27
DMA 2	13491	11962	1529	11.33
DMA 3	6419	6190	229	3.57
DMA 4	3897	3743	154	3.95
		Total	4,824	17.01%

Following repairs to the main potable line, building water efficiency was targeted. Night time meter readings revealed either leaks meter side of the building or inefficient water fittings the most common fault was with urinal control devices. These were adjusted where possible and replaced leading to large water efficiency savings within buildings (table 5).

Table 5. Water Efficiency Analysis December 2011

	Mean daily consumption End (L)	Water consumption per person per day (L)	Daily Water Saving (L)
Meter 1	352	11	4,283
Meter 6	194	24	2,936
Meter 37	625	20	3,000
Meter 40	192	19	2,641

From the start of the trials in April 2010 there was a 39% decrease in mean monthly water consumption from 962m^3/day to 584m^3/day in July 2014 (Fig 4). Over the course of a year, this would have provided a saving of 137 million litres. A 39.3% reduction in annual consumption from 2009 compared to 2014 (Fig 5) led to the lowest consumption figures on record and the internal water efficiency targets for 2020.

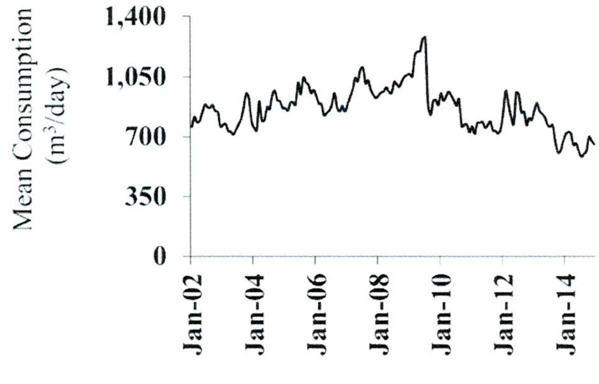


Figure 4. Mean Daily Consumption

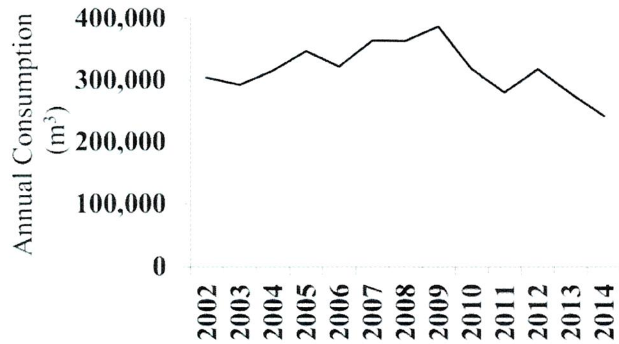


Figure 5. Annual Consumption

3.5 Control

Minimum night flow was used as the main control measurement (Fig 6). Investigations were initiated in the instances where water consumption rose above 22.5m^3/h in the form of further DMA analysis and water efficiency audits around MPC. Following the initial investigations which realised issues with the urinal control systems water efficiency training was provided to cleaning staff. This proved a successful control measure which helped identify several faulty urinals and leaks.

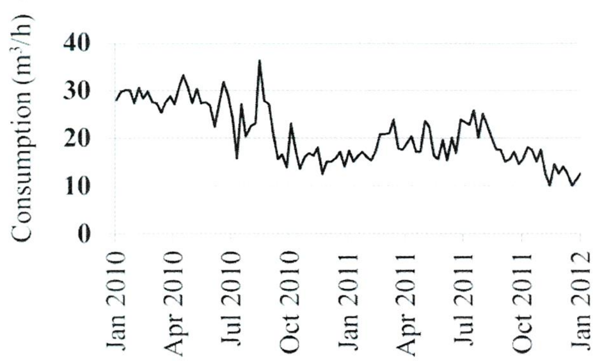


Figure 6. Minimum Night Flow

4. Discussion

The successful application of water management tools utilising a six sigma approach further strengthens the notion that Six Sigma is not just applicable to manufacturing operations (Anthony, 2007; Laureani *et al.*, 2009). The systematic approach of this methodology allows successful investigations to be conducted in-house whilst providing the necessary flexibility to be tailored to distribution systems of varying sizes and sophistication. This is in accordance with previous research which has demonstrated that sustainability performance, especially relating to environmental sustainability can be enhanced through the use of a Six Sigma methodology (Calia *et al.*, 2009; Habidin and Yusof, 2012; Zhang and Awasthi, 2014). Due to the rigorous and disciplined nature of Six Sigma methodologies (Garza-Reyes, 2015) it is not certain if the ease of implementation could have been enhanced through the use of simpler frameworks such as Lean (Ng *et al.*, 2015). In this case, application of the Six Sigma framework achieved a 39% reduction in annual water consumption, however it is not clear if this could have been enhanced through the utilisation of other frameworks was designed to overcome the shortfalls of Six Sigma. Green Lean Six Sigma (GL2S) and was shown on average to produce reductions in resource consumption by 20% to 40% (Cherrafi *et al.*, 2016b) which is similar to the 39% reduction witnessed in this case study however due to the inherent differences in the sectors and organisations in which this was applied a further casestudy on MPC would have to be conducted in order to determine if GL2S provide significant benefits of Six Sigma.

The design and installation of more sophisticated systems will allow for the quick and accurate capture of flow data and has the potential to greatly aid in reducing water consumption. The installation of telemetry systems at a DMA and household level will provide more accurate information on consumption, as well as real time consumption data in a water distribution system, making it possible to identify leaks or inefficient water consumption (Loureiro, 2010). Increased metering would improve the resolution of the leak detection process, reducing the amount of time spend locating leaks or inefficiencies in the system. In addition to metering, various sensing technologies such as IR thermographic imaging could be used (Hawari *et al*, 2017) as well as considering the operating pressure of the distribution network. It has been recognized that leakage rates increase with internal water pressure (Wu *et al*, 2010) and that considering pressure driven leakage in the design stage of the network can reduce overall leaks (Gupta *et al*, 2016). Considering the analytical nature of the six sigma framework, the collection of operational pressure data from the distribution network could be used to identify the areas in which leaks are most likely to occur and thus be used to further reduce site water consumption.

As technological strategies can be costly and would require a detailed cost benefit analysis before any capital works is undertaken: greater benefits are typically achieved by the fitting of efficient fixtures and appliances (Stewart *et al.*, 2010). It has also been documented that high tech equipment and water efficient devices alone do not necessarily provide water savings (Geller *et al*, 1983), as end users must have an appreciation of the functions of the devices and apply them in the correct manner (Elizondo and Lofthouse, 2010). Additionally if these fittings are incorrectly fitted then there is the potential that large volumes of water could be wasted especially in public buildings (Roccaro *et al*., 2011). Both of these issues were apparent within the case study and related to urinal flow control devices which were wasting 2-4 litres a minute. This was a combination of faulty devices, incorrectly fitted devices, and cleaning staff increasing the flush frequency to reduce the amount of cleaning required. Rectification of this issue was tackled through in house training on how the fittings worked, how they should be positioned, and how to spot if they are working correctly. This has led to water savings of up to 90% in some facilities.

The behavioural aspects to water reduction have been shown to play an important part in achieving water savings. Lau *(2012)* demonstrated how the promotion of knowledge and awareness could yield significant water savings in high density housing in Hong Kong. The effectiveness of a similar strategy when applied to a single MOD estate would be questionable due to the high transient population with no visibility of personal water consumption (Randolph and Troy, 2008). This is particularly relevant on this site where the consumption patterns of domestic and non-domestic water consumption is still not fully understood. If targeted at an organisational level so that water conservation was embedded in to the military culture and institutionalised, then there is the potential to realise large water savings across all estates (Randolph and Troy, 2008).

5. Conclusion

- Six Sigma methodology can been successfully utilised to reduced water loss in a Private Water Supply Setting to meet environmental quality targets.
- The underlying principals were successfully applied to both supply side (leakage) and demand side (water efficiency audits) to achieve a 39% reduction in annual consumption.
- The adoption of a systematic approach to reducing water consumption can aid in providing an effective, low cost solution even on sites with limited technology. As water conservation begins to play a more influential role within organisations, industry will naturally have to adapt to support these new requirements.

Acknowledgement

The support of Interserve, Defence Infrastructure Organisation and the Ministry of Defence.

References

Affinity Water. (2010). Water Efficiency Audit Guide. [Online] Available: https://www.affinitywater.co.uk/docs/business-water-audit-guide.pdf (January 2, 2017)

Antony, A. (2007). Six Sigma in service organisations. Benefits challenges and difficulties, common myths, empirical observations and success factors. *International Journal of Quality and Reliability Management, 24*, 294-311. https://doi.org/10.1108/02656710710730889

AWWA Water Loss Control Committee. (2003). Applying worldwide BMPs in water loss control. *Journal of AWWA, 95*, 65-79.

Banawi, A., & Bilec, M. M. (2014). A Framework to Improve Construction Processes: Integrating Lean, Green and Six Sigma. *International Journal of Construction Management,* *14*, 45-55. https://doi.org/10.1080/15623599.2013.875266

Basu, R. (2004). *Implementing Quality: A practical Guide to Tools and Techniques: Enabling the Power of operational excellence*, Thomson Learning, London.

Bates, B. C., Kundzewicz, Z. W., Wu, S., & Palutikof, J. P. (2008). *Climate Change and Water –IPCC Technical Paper VI.* Intergovernmental Panel on Climate Change Secretariat, Geneva.

British Institute of Facilities Management. (2007). Survey shows key role for FMs business continuity and sustainability. [Online] Available: www.bifm.org.uk/bifm/news/3899 (January 2, 2017)

Brooks, D. B. (2006). An operational definition of water demand management. *International Journal of Water Resource Development, 22*, 521-528. https://doi.org/10.1080/07900620600779699

Cairns, R., Sharp, E., Maher, K., & Jefferson, B. (2012). Scope for potential energy savings in the flocculation process. *Water and Environment Journal, 26*, 319-323. https://doi.org/10.1111/j.1747-6593.2011.00291.x

Calia, R. C., Muller-Guerrini, F., & de Castro, M. (2009). The impact of Six Sigma in the performance of a Pollution Prevention program. *Journal of Clean Production, 17*, 1303-1310. https://doi.org/10.1016/j.jclepro.2009.05.001

Cherrafi, A., Elfezazi, S., Chiarini, A., Mokhlis, A., & Benhida, K. (2016a). The integration of lean manufacturing, Six Sig and Sustainability: A literature review and future research directions for developing specific model. *Journal of Cleaner Production, 139*, 828-846. https://doi.org/10.1016/j.jclepro.2016.08.101

Cherrafi, A., Elfezazi, S., Govindan, K., Garza-Reyes, J. A., Benhida, K., & Mokhlis, A. (2016b). A framework for the integration of Green and Lean Six Sigma for superior sustainability performance. *International Journal of Production Research, 55*, 1-35. https://doi.org/10.1080/00207543.2016.1266406

Colombo, A. F., & Karney, B. W. (2002). Energy and costs of leaks: towards a comprehensive picture. *Journal of Water Resources Planning and Management, 128*, 441-450. https://doi.org/10.1061/(ASCE)0733-9496(2002)128:6(441)

Elmualim, A., Shockley, D., Valle, R., Ludlow, G., & Shah, S. (2010). Barriers and commitment of facilities management profession to the sustainability agenda. *Building and Environment, 45*, 58-64. https://doi.org/10.1016/j.buildenv.2009.05.002

Environment Agency. (2012). Businesses urged to use water efficiency as a way to boost profits. [Online] Available, http://www.environment-agency.gov.uk/news/142885 (January 7 2017)

Elizondo, G. M., & Lofthouse, V. (2010). Towards Sustainable use of water at home understanding how much where and why? *Journal of Sustainable Development, 3*, 3-10.

https://doi.org/10.5539/jsd.v3n1p3

FM World. (2013). MOD seeks water saving measures in future deals. [Online] Available: http://www.fm-world.co.uk/news/fm-industry-news/mod-seeks-water-saving-measures-in-future-deals/ (January 16 2017).

Gaza-Reyes, J. A. (2015). Green Lean and the need for Six Sigma. *International Journal of Lean Six Sigma, 6*, 226-248. https://doi.org/10.1108/IJLSS-04-2014-0010

Geller, E. S., Erickson, J. B., & Buttram, B. A. (1983). Attempts to promote residential water conservation with educational behavioural and engineering strategies. *Population and Environment, 6*, 96-112. https://doi.org/10.1007/BF01362290

Gupta, R., Nair, A. G. R., & Ormsbee, L. (2016). Leakage as Pressure-Driven Demand in Design of Water Distribution Networks. *Journal of Water Resources Planning and Management, 142*, 1-14. https://doi.org/10.1061/(asce)wr.1943-5452.0000629

Habidin, N. F., & Yusof, S. M. (2012). Relationship between lean six sigma, environmental management systems and organizational performance in Malaysian automotive industry. *International Journal of Automotive Technology, 13*, 1119-1125. https://doi.org/10.1007/s12239-012-0114-4

Hawari, A., Khader, M., Hirzallah, W., Zayed, T., & Moselhi, O. (2017). Integrated sensing technologies for detection and location of leaks in water distribution networks. *Water Science and Technology: Water Supply*, 17. https://doi.org/10.2166/ws.2017.044

Inozu, B., Niccolai, M. J., Whitcomb, C. A., Mac Claren, B., Radovic, I., & Bourg, D. (2006). New horizons for ship building process improvement. *Journal of Ship Production, 22*, 87-98.

Karim, M. R., Abbaszadegan, M., & LeChevallier, M. (2003). Potential for pathogen intrusion during pressure transients. *Journal of American Water Works Association, 95*, 134-146.

Kingdom, B., Liemberger, R., & Martin, P. (2006). The challenge of reducing non-revenue water (NRW) in developing countries: how the private sector can help; a look at performance-based service contracting. *Water Supply Sanitation sector board discussion paper series, 8*, 8-12

Lai, C. C. (1991). Unaccounted for water and the economics of leak detection. *Water Supply, 9*, 1-8.

Lau, S. S. Y., Yang, F., & Ma, A. Y. W. (2012). Potable water savings in high density housing. *Journal of Facilities Management, 10*, 226-240. https://doi.org/10.1108/14725961211246027

Laureani, A., Antony, J., & Douglas, A. (2009). Lean Six Sigma in a call centre: a case study. *International Journal of Productivity and Performance Management, 59*, 757-768. https://doi.org/10.1108/17410401011089454

Loureiro, D. (2010). *Consumption analysis methodologies for the efficient management of water distribution systems*, PhD Thesis, Universidade Técnica de Liboa, Lisbon, Portugal.

Morais, D. C., & de Almedia, T. (2007). Group decision-making for leakage management strategy of water network. *Resources, Conservation and Recycling, 52*, 441-459. https://doi.org/10.1016/j.resconrec.2007.06.008

Morrison, J. (2004). Managing Leakage by District Metered Areas: A Practical Approach. *Water, 21*(2), 45-46.

Ng, R., Low, J. S. C., & Song, B. (2015). Integrating and implementing Lean and Green practices based on proposition of Carbon-Value efficiency metric. *Journal of Cleaner Production, 95*, 242-255. https://doi.org/10.1016/j.jclepro.2015.02.043

Oakland, J. S., & Tanner, S. J. (2007). A new framework for managing change. *The TQM Magazine, 19*, 572-589. https://doi.org/10.1108/09544780710828421

Phen, L. S., & Hui, M. S. (2004). Implementing and applying six sigma in construction *Journal of Construction and Engineering Management., 130*, 482-9. https://doi.org/10.1061/(ASCE)0733-9364(2004)130:4(482)

Price, S., Pitt, M., & Tucker, M. (2011). Implications of a sustainability policy for facilities management organisation. *Facilities, 29*, 391-410. https://doi.org/10.1108/02632771111146314

Randolph, B., & Troy, P. (2008). Attitudes to conservation and water consumption. *Environmental Science and Policy, 11*, 441-445. https://doi.org/10.1016/j.envsci.2008.03.003

Roccaro, P., Falciglia, P. P., & Vagliasindi, F. G. (2011). Effectiveness of water saving devices and educational programs in urban buildings. *Water Science and Technology, 63*, 57-65. https://doi.org/10.2166/wst.2011.190

Stewart, R. A., Willis, R. M., Giurco, D., Panuwatwanich, K., & Capati, G. (2010). Web based knowledge management system: linking smart metering to the future of urban water planning. *Australian Planner, 47*, 66-74. https://doi.org/10.1080/07293681003767769

Waterwise. (2012). Water- the facts: Why do we need to think about water [Online] Available: http://www.waterwise.org.uk/data/resources/25/Water_factsheet_2012.pdf (May 08 2017).

Zhang, M., & Awasthi, A. (2014). Using Six Sigma to achieve sustainable manufacturing. Innovative Design and Manufacturing. *Proceedings of the 2014 International Conference on innovative design and manufacturing*, Montreal, August 13-15.

Willis, R. M., Stewart, R. A., Panuwatwanich, K., Williams, P. R., & Hollingsworth. (2011). Quantifying the influence of environmental and water conservation attitudes on household end use water consumption. *Journal of Environmental Management, 92*, 1996-2009. https://doi.org/10.1016/j.jenvman.2011.03.023

Wu, Z. Y., Sage, P., & Turtle, D. (2010). Pressure-dependent Leak Detection Model and its Application to a District Water System. *Journal of Water Resources Planning and Management, 136*, 116-128. https://doi.org/10.1061/(ASCE)0733-9496(2010)136:1(116)

Macrothink Institute™
Environmental Management and Sustainable Development
ISSN 2164-7682
2017, Vol. 6, No. 2

Glossary

DMA: District Metered Area

DMAIC: Define, Measure, Analyse, Improve, Control

MNF: Minimum Night Flow

MOD: Ministry of Defence

MPC: Mount Pleasant Complex

SIPOC: Suppliers, Inputs, Processes, Outputs, Customers

WDM: Water Demand Management

ice Publishing

Badr and Venables
ISBN 978-0-7277-6525-3
https://doi.org/10.1680/oicwe.65253.091

MANAGEMENT OF WATER LOSSES IN WATER DISTRIBUTION SYSTEMS USING LEAN SIX SIGMA FRAMEWORK

J.S. AlDairi[1] and A. Badr[2]*

[1,2]Military Technological College, Muscat, Oman
**Corresponding author*

ABSTRACT There is a wide range of leak management tools and techniques in water distribution systems. This paper investigates the application of Lean Six Sigma methodology in managing losses of water distribution systems. A DMAIC5S-TPM conceptual framework has been developed to facilitate this approach. The proposed DMAIC5S-TPM is based on the integration of DMAIC, 5S and TPM. The 5S and TPM are often used together to develop integrated management systems. However, in the proposed framework, the DMAIC methodology is integrated in between 5S and TPM. This framework presents the use of Lean Six Sigma dimensions as a sustainable utility management method that could be used by water utilities' managers and engineers. In addition to managing water losses, this method will enhance the effectiveness of using water resources. It could also serve as a quality assurance tool for satisfying customer needs and expectations.

1. Introduction

Many countries are suffering from a shortage of water. Kılıç (2020) stated that only (0.3%) of the water resources around the world are classified as usable water. Azevedo and Saurin (2018) highlighted that in water distribution systems, a level of 20% to 30% of the worldwide produced treated water is lost. As a result, it is estimated that more than one billion of people are living without adequate potable water. Therefore, it is necessary to utilise a responsible sustainable approach towards managing water resources and maintenance of distribution systems.

It is widely accepted that water loss from distribution systems has become a universal challenge that requires tackling through a sustainable approach. This is a worrying problem even for well developed countries that have good infrastructure and excellent operating practices. However, the problem is more serious for developing countries, with poor distribution system infrastructure, because it causes a serious risk and significant economic consequences (Farley and Trow, 2005).

This paper investigates the application of Lean Six Sigma methodology in managing losses of water distribution systems. It presents a systematic management approach based on solid and well known international continuous improvement industrial practise. In fact, the contribution of this paper is to introduce a novel conceptual framework that will assist water utility managers and maintenance practitioners in achieving their business excellence.

This research is applying mixed method. It aims to proceed further in implementing the proposed framework in real applications. And this will require, indeed, collecting qualitative and quantitative data that will be used to evaluate and test the framework.

2. Losses in water distribution system (WDS)

A water distribution system (WDS) delivers potable water to consumers in a metropolis by interconnecting reservoirs, pipes, and related accessories (Liu and Yu, 2014). It is a complex socio-technical system that is interconnected logically, geographically, and physically with related infrastructure critical items to deliver a potable water to various consumers (Azevedo and Saurin, 2018). This type of interconnection will help in hastening identifying root causes of WDS losses through defining its external perspectives, since the whole system is integrating the social and technology aspects under one platform.

In a recent study, Fezai et al. (2021) defined the WDS as a complex process that is composed of multi-hydraulic elements (e.g., reservoirs and consumption nodes) which are interconnected together by links (e.g., pumps, valves, and pipes). They have proposed a statistical hypothesis test for leak detection which aims to evaluate the existence of leaks in WDS. Water losses in any type of distribution systems depend on many attributes, including specifications of the WDS pipe network and other internal and external factors such as the utility provider operational management practise, the level of proficiency and technology deployed to control the system (Farley and Trow, 2005). Chirica *et al.* (2018) reported that water losses (the percentage of the annual non-revenue water) reach a level of 11% in developed countries and rise up to 65% in developing countries.

It is almost impossible to eliminate losses in WDSs due to the complexity of these socio-technical systems. Azevedo and Saurin (2018) acknowledged the difficulty of controlling losses in WDSs as they captured more than 130 methods from literatures that focus on controlling different types of WDS losses. Their systematic review of previous studies on losses in WDSs focused on the factors influencing the complexity attributes (e.g., engineering attributes and extensive distribution system), non-linear interactions (e.g., leaks not detected and variations in water consumption), diversity of elements (e.g., different categories of consumers and different supply sources), unexpected variability (e.g., WDS infrastructure conditions), and resilience (e.g., continuous supply).

Lambert (2002) presented the first IWA (International Water Association) water loss task force performance indicators as illustrated in Table 1; also known as IWA Standard Water Balance. This approach is used widely by technical organisations, utilities, consultants, and regulators, in many countries. Globally, the Non-Revenue Water (NRW) is representing the main challenge that utilities providers are facing and need addressing.

Water losses could also be classified into apparent and real losses. The apparent losses can be determined from unauthorised consumption and metering inaccuracies. On the other hand, real losses result from leakage on transmission and/or distribution mains, leakage on overflows at utility's storage tanks, and leakage on service connections up to the measurement point. Azevedo and Saurin (2018) described the IWA as a worldwide solid applied methodology used to assess water losses in various types of WDSs. Although, the performance indicators are comprehensive and covering the many aspects, it has its limitation due to the complexity of such socio-technical systems.

Table 1 IWA standard water balance, adopted from Lambert (2002)

System Input Volume	Authorised Consumption	Billed Authorised Consumption	Billed Metered Consumption (including water exported)	Revenue Water
			Billed Unmetered Consumption	
		Unbilled Authorised Consumption	Unbilled Metered Consumption	Non-Revenue Water (NRW)
			Unbilled Unmetered Consumption	
	Water Losses	Apparent Losses	Unauthorised Consumption	
			Metering Inaccuracies	
		Real Losses	Leakage on Transmission and/or Distribution Mains	
			Leakage and Overflows at Utility's Storage Tanks	
			Leakage on Service Connections up to the Measurement Point	

3. Lean six sigma (LSS)

Lean Six Sigma (LSS) as an integrated solution has become one of the most used business process improvement methodology in the last two decades (Antony *et al.*, 2020). Lean and Six Sigma have different historical developments. The concept of Lean was explored globally after the success of a Lean Production System created by Toyota manufacturing company. It is basically focused on eliminating waste by determining the value-added activities or steps from non-value-added activities or steps. However, Six Sigma is a data driven approach that is focusing on reducing process variations which affect the product performance. Literally saying, Six Sigma means having a process that produce 3.4 defects per million opportunities (DPMO) (Franchetti, 2015).

Environmental Protection Agency (EPA) is committed to eliminate the world's known eight deadly wastes in water utilities by Lean management technique. These wastes are defects or rework, overproduction, waiting, non-utilisation of talents, transportation, excessive inventory, motion, and overprocessing. Based on the researcher experience, Table 2 illustrates clear examples of those wastes from daily practises in WDS field.

Deploying Lean management in isolation cannot improve the process variation and, similarly, deploying Six Sigma in isolation cannot eliminate all types of process waste. In most of the Lean management applications, the solution is quite visible, and it just needs a set of tools to implement the known solution. However, when the problems need understanding of critical process parameters and the output varies due to number of factors, then Six Sigma is the one to use (Corbett, 2011; Antony *et al.*, 2020).

In many of the process improvement cases, it was realised that applying Six Sigma or Lean methodology alone might not be enough as both are complementary to each other. For example, it is preferable sometimes to reduce the process variations in addition to reducing the average lead time. This approach of combination is taking the advantage of six sigma in reducing the process variations and the advantage of Lean management in the sense of waste elimination (Antony *et al.*, 2020). In fact, this approach was initially triggered by George (2002); as he stated that Lean on itself cannot resolve complex problems that require statistical data analysis techniques. Antony *et al.* (2017) emphasised that integrating Lean and Six Sigma will help in achieving high efficiency and effectiveness which will instantly boost the organisation into achieving the desired business excellence. This is technically valid as LSS provides the required tools and techniques that could prepare leaders for managing changes during the cycle of process improvement.

Table 2 Types of wastes with examples from WDS field

Waste	Examples
Defects	Improper repair of joint leak in main water distribution line.
Overproduction	Exceeding the optimum limit of water production.
Overprocessing	Injecting non-added-value chemicals to the water treatment process.
Waiting	Unavailability of critical spare parts.
Non-utilised or under-utilised resources/talents	Plant operator is working as a clerk.
Transportation	Water distribution network contains many unnecessary lines.
Inventory	Storing larger amount of chemicals than required.
Motion	Unnecessary movement of technicians inside the utility plant due to improper equipment layout.

4. The applications of lean and six sigma in water sector

Cairns and MacPherson (2017) have investigated the effectiveness of deploying Six Sigma in private WDSs with the aim of reducing water losses. They implemented water reduction plans in over 200 water facilities. The findings revealed the potential of huge water savings in both supply and demand sides. It was also shown that this approach can be successfully applied on sites with limited technologies. Although they did not focus on Lean, they used DMAIC (Define, Measure, Analyse, Improve and Control) methodology as a key aspect for organisations to achieve peak operational performance.

Kung *et.al* (2008) applied the principles of Lean management in a case study related to water and sewage system in Edmonton, Canada. They validated the effect of the proposed improvements using Lean techniques in productivity. There was a potential of high effectiveness, but they stated that such improvements need to be quantified and measured, suggesting that Six Sigma is the right method to be integrated in their future work.

EPA (2012) created Effective Utility Management (EUM) as a guide and framework for identifying priorities, strengths, and weaknesses of water utilities. It has some attributes and keys that assess the effectiveness of managing water utilities. It was claimed that LSS facilitated the roadmap into achieving the EUM attributes, and hence it enabled the water-sector to increase the profitability and improve the product quality and employee morale.

Almutairi (2020) applied LSS principles in a selected case studies and data collected from literature review in water maintenance operations. He tested the degree of improvements in maintenance and quality by calculating MTTR (Mean Time to Repair), MTBF (Mean Time Between Failure), availability, reliability, and the 5S audit scores. The findings illustrated a clear improvement in maintenance operation (i.e., the MTTR was reduced and/or the MTBF was increased). In addition, similar improvements were noticed in other parameters (i.e., availability, reliability, and the 5S audit scores).

Lean and Six Sigma contain many tools and techniques that can be integrated based project selection and problem-solving methodology (Corbett, 2011; Antony *et al.*, 2020). However, this study will focus only in presenting the proposed integration of DMAIC (as a Six Sigma problem-solving methodology) along with Five S and Total Productive Maintenance (TPM) as Lean management tool and technique.

4.1 Five S (5S)

The 5S tool is the basic tool that represents a proven and solid method towards sustainable warehousing that has been used widely in very successful and leading-edge organisations. This tool can show many hidden elements that act as obstacles, and it can motivate the employees for upcoming culture change. 5S is a Japanese workplace organisation tool which contains five words that start with the letter "S": seiri ("sort" or clear selected area from unrelated/unneeded items), seiton ("set in order" or rearranging tools and portable equipment layout based on usage frequency), seiso ("shine" or cleaning and checking), seiketsu ("standardise" or applying standard procedures/instructions within workplace), and shitsuke ("sustain" or ensuring the continuity of previous steps). It is used widely with TPM to develop an integrated management system (Ahuja and Khamba, 2008). However, for the proposed framework, the DMAIC methodology will be integrated in between.

4.2 DMAIC

DMAIC is a Six Sigma problem-solving method that can be elaborated as:

- Define: The business value and outcomes besides customer needs using critical to quality (CTQ), process mapping or voice of customer techniques.
- Measure: To evaluate data that help in setting criteria and priorities.
- Analyse: By determining root causes and deep understanding of the process and problem (using tools like Fishbone, Value Stream Mapping, and hypothesis test).
- Improve: By refining goal statements and developing achievable solutions.
- Control: To monitor the changes by developing a tracking process, applying Statistical Process Control, and maintaining a document management system (AlDairi, 2017; Franchetti, 2015).

4.3 Total productive maintenance (TPM)

TPM is a technique that aims to maximise the efficiency of equipment through its lifetime by preventing unexpected quality defects, speed losses, and breakdown throughout the process (Ahuja and Khamba, 2008). They classified 16 major losses in the manufacturing process for example, losses in overall equipment efficiency (OEE). equipment loading time, worker efficiency, and the use of production resources. These losses can be eliminated by adopting the (8) main pillars of TPM identified by the authors:

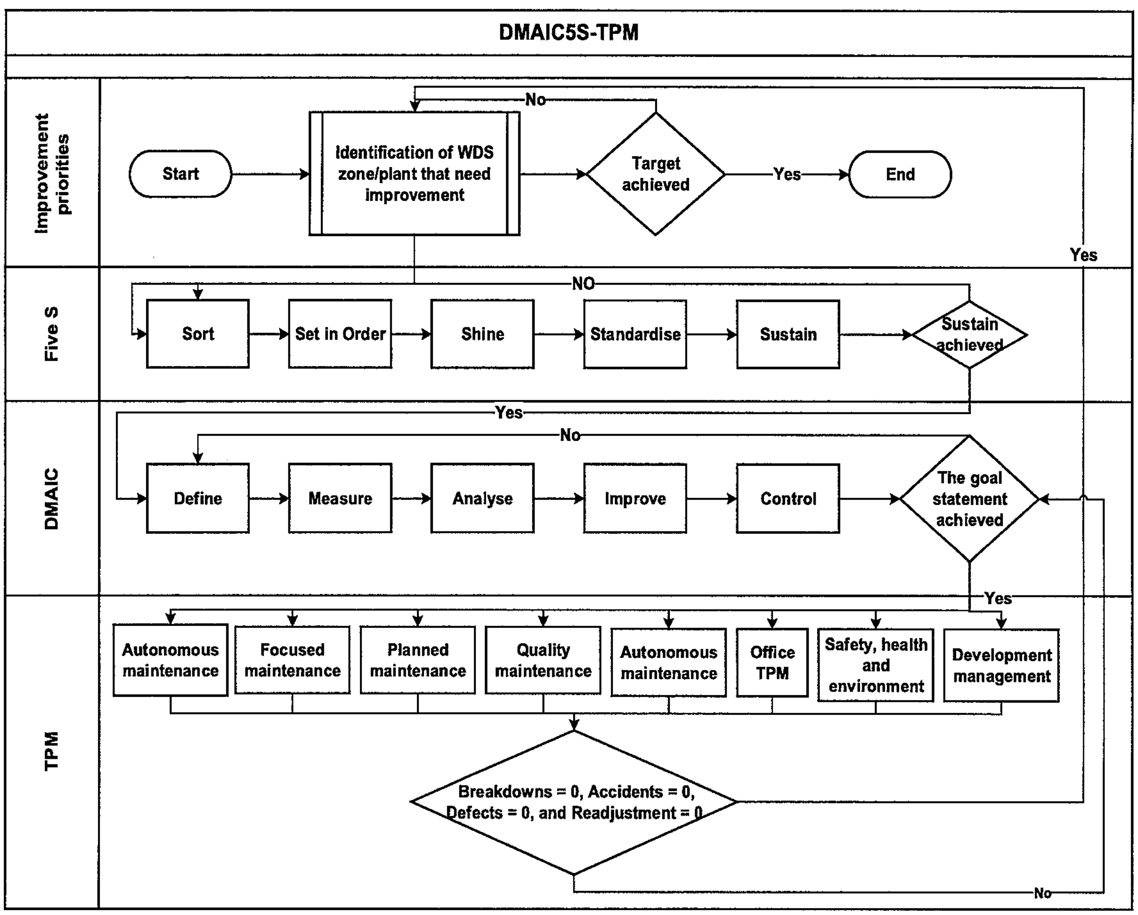
- Autonomous maintenance: (e.g., cleaning, lubricating, and tightening).
- Focused maintenance: (e.g., increasing efficiency, monitoring OEE).
- Planned maintenance: (e.g., improving MTBF and MTTR).
- Quality maintenance: (e.g., working towards zero defects).
- Education and training: (e.g., boosting employees' skills).
- Office TPM: (e.g., improving cross-functional collaboration).
- Safety, health, and environment: (e.g., ensuring safety at workplace).
- Development management: (e.g., encouraging employee's initiatives).

Swamidass (2002) highlighted the importance of integrating TQM (Total Quality Management), TPM, JIT (Just in Time), and employee involvement in the manufacturing environment. He emphasised that TPM is a critical success factor in improving equipment performance and organisational capabilities. Using data collected from 179 manufacturer, McKone et al (2001) tested the effect of TPM on performance and suggested that it improves the manufacturing process by supporting JIT and TQM. In order to perform a realistic Lean maintenance, Mostafa et al. (2015) emphasised that TPM practitioners have to ensure proper implementation of autonomous maintenance, planned maintenance, Root Cause Analysis (RCA), safety improvement, OEE, and work order system.

5. The DMAIC5S-TPM conceptual framework

Despite the flexibility of adopting various management tools and techniques, water utility managers have raised the concern of not having a standard framework that can improve the operation and maintenance aspects and fit to all sizes of utilities (EPA, 2012). Building on the water losses scheme described by the IWA standard water balance, this paper is presenting a conceptual framework that can be used by WDS managers and practitioners to maximise the system efficiency and organisation profitability. To apply standardisation in tackling the investigation of different types of losses, the proposed framework aims to maintain a sustainable and reliable approach. This approach will be achieved by integrating the 5S with DMAIC and TPM in a platform named as DMAIC5S-TPM which can be illustrated as shown in Figure 1.

Figure 1 The proposed DMAIC5S-TPM conceptual framework



The first step is to identify the WDS improvement priorities, whether it is a utility, small plant, or a selected zone. The complicated structure of WDS requires pre-identification of areas/zones that need critical improvement. This depends on rate of losses, geographical spread and diversity of system components which must be investigated by a dedicated LSS team. After selecting the priority zone that require improvement, an immediate plan has to be launched deploying the 5S technique embedded with visual management.

6. Implementation of the conceptual framework

In practice, the implementation of LSS methodology is carried out as a small project, which might require from 4 to 6 months to complete (Antony *et al.*, 2020). A closer look at most of WDSs utilities and site locations can create a quick impression of how much improvement can be made. The unawareness of cleaning, lubricating equipment, calibrating gauges…etc, are not just due to lack of supervision. In fact, most water utilities processes can be dealt with as manufacturing processes (i.e., producing/delivering water with certain specification and standards/quality that meets customer needs). Therefore, it requires adopting a solid scientific and sustainable method of management. In practice, managing the workplace is the easiest part to start with for such improvements, and therefore, utilising the 5S tool enables the managers to:

- Sort out all critical and value-added WDS parts/items from those that are not adding any value and should be removed from the site (e.g., storing un-used pipes, pumps, and accessories around the plant may cause a hazard within a utility).
- Set in order, so that maintenance equipment and hand tools are placed based on frequent usage.
- Shine, where every part within the WDS must be cleaned from debris, oil, or any kind of waste.
- Standardise, for instance by creating routine operation and maintenance checklists, color-coding the waterflow lines and gauges.
- Sustain, by ensuring the continuous improvements to previous steps.

The fast results of deploying the 5S tool will ease the implementation of the next step in the proposed framework, which is the DMAIC problem-solving method. Before starting the DMAIC cycle, the LSS project must be selected carefully with a clear goal statement to avoid unexpected future implications (Antony *et al.*, 2020). In this study, the selected project goal statement might be "reducing particular WDS zone apparent water losses by 99%". Then, comes the Measure phase which includes collecting relevant historical and concurrent data of water supply and consumptions that will be needed for the next phase. Many organisations build complex data collection and information management systems without really understanding how the data collected can benefit the organisation (LSS Black Belt Manual, 2013). In fact, having such valuable data without using them for example in testing hypothesis or investigating root causes of critical issues means creating waste of information and manpower efforts. In DMAIC cycle, the findings of the "Measure phase" will be carried forward to the Analysis phase.

In the Analysis phase, the LSS project team will test whether the initial scope has been validated. Additionally, the project team has to ensure identifying the value added and non-value-added steps within the specified process. For example, redundant pumps' setup or other unnecessary installations. The project team will also look at whether the cause-and-effect technique was used to determine potential problem causes, and to prioritise those causes. This will drive the LSS team to start looking into the Improve phase; where the process starts by identifying the valid root causes derived from the previous step. This is followed by selecting the best solutions that fit with the Plan, Do, Check, Act (PDCA) continuous improvement cycle, and completed by preparing an action plan and updating the project documents.

In the Control phase, the LSS project team begins the process by ensuring that the customer needs are met, the Value Stream Mapping (VSM), and the eight major wastes have been re-identified. Moreover, the team will evaluate actions taken for managing the documentation process. Documentation is very important to ensure what has been learned from the LSS projects is systematically shared within the organisation for implementing solutions and supporting on-going training (LSS Black Belt Manual, 2013). Finally, the team will check the existence of a process monitoring tool (e.g. control chart) to secure the objective of continuous improvement. Process monitoring is very important in the Control phase. According to Woodall and Montgomery (2014), it should monitor input variables so that improvement from the previous stage can be maintained over time.

Despite the improvements in the selected LSS project zone, WDS's assets are still in need of a sustainable maintenance management tool, where lack of maintenance awareness can be addressed. Therefore, integrating the TPM technique will boost the overall performance which will enhance the infrastructure stability and operational resilience of the WDS. The idea behind integrating TPM after the DMAIC is due to the fact that TPM will be continuously running as a standalone Lean method after succeeding in completing the LSS project in that specific zone. In other words, the operation and maintenance team responsible for that zone are committed towards achieving optimum targets of zero breakdowns, zero accidents, zero defects, and zero readjustment (Ahuja and Khamba, 2008). Ngadiman *et. al* (2012) emphasised that TPM acts as a proactive method that aims to determine operational issues and plan to prevent their occurrences. In fact, the TPM is integrating both maintenance and problem-solving activities within the responsibilities of any plant operator rather than depending only on maintenance staff. For example, along with basic maintenance skills, operators have to be trained in dealing with six major types of pump losses, including setup and adjustment losses, speed losses, breakdown losses, stoppage losses, quality defect losses, and equipment losses. These will lead to continuous improvement of the overall system OEE, availability, reliability, and maintainability. The whole DMAIC5S-TPM implementation cycle is repeated for other zones based on operation and maintenance priorities.

7. Conclusion

This paper proposed a conceptual framework that can be used by WDS managers and practitioners to maximise the system efficiency and organisational profitability. The proposed DMAIC5S-TPM is based on the integration of DMAIC, 5S and TPM. The 5S and TPM are often used together to develop integrated management systems. However, in the proposed framework, the DMAIC methodology is integrated between 5S and TPM.

The proposed DMAIC5S-TPM conceptual framework presents the use of Lean Six Sigma dimensions as a sustainable utility management method and aims to maintain a sustainable and reliable approach.

The framework must be evaluated and tested in real application. The implementation cycle starts by applying 5S as a Lean management tool followed by the DMAIC Six Sigma problem-solving technique and accomplished by adopting the TPM towards achieving optimum targets of zero breakdowns, zero accidents, zero defects, and zero readjustment.

It has been shown that the implementation of this framework could help practitioners in developing continuous improvement of the overall WDS networks through proper management of workplace. It could also help in achieving standard excellence of OEE, availability, reliability, and maintainability that will satisfy customer needs.

8. Acknowledgement

This work is supported by the Military Technological College, Sultanate of Oman.

9. References

Ahuja IP and Khamba JS (2008) Total productive maintenance: literature review and directions. *International Journal of Quality & Reliability Management,* 25 (7): 709-756.

AlDairi JS et al. (2017) Knowledge-based Lean Six Sigma Maintenance System for Sustainable Buildings. *International Journal of Lean Six Sigma*, 8(1): 109-130.

Almutairi NA (2020) Using Lean Principles in Water Maintenance Operations with Application to Water Treatment Plant. *Journal of Environmental Treatment Techniques*, *8*(4): 1345-1351.

Antony J et al. (2017) Lean Six Sigma: Yesterday, today and tomorrow. *International Journal of Quality & Reliability Management*, 34(7): 1073–1093.

Antony J et al. (2020) *The Ten Commandments of Lean Six Sigma: A Guide for Practitioners*. Emerald Publishing Limited, Bingley, U.K.

Azevedo B and Saurin T (2018) Losses in Water Distribution Systems: A Complexity Theory Perspective. *Water Resources Management*, 32(9): 2919–2936.

Cairns R and MacPherson M (2017) A six sigma approach to water savings. *Environmental Management and Sustainable Development*, *6*(2): 98-111.

Chirica S et al. (2018) Water Loss Management In Order To Protect Available Water Resources. *Land Reclamation, Earth Observation and Surveying, Environmental Engineering,* VII: 122-127.

Corbett LM (2011) Lean Six Sigma: The contribution to business excellence. *International Journal of Lean Six Sigma*, 2(2): 118–131.

EPA (Environmental Protection Agency) (2012) *Resource Guide to Effective Utility Management and Lean Improving Performance and Addressing Key Management Priorities at Water-Sector Utilities.* Environmental Protection Agency (EPA), Washington, U.S.

Farley M and Trow S (2005). *Losses in Water Distribution Networks: A Practitioner's Guide to Assessment.* IWA Publishing, London, UK.

Fezai R et al. (2021) A Novel Leak Detection Approach in Water Distribution Networks. *18th International Multi-Conference on Systems, Signals & Devices (SSD).* IEEE, Monastir, Tunisia (Virtual), pp. 607-612.

Franchetti, M (2015) *Lean Six Sigma for engineers and managers.* CRC Press, USA.

George M (2002) *Lean Six Sigma: Combining Six Sigma quality with lean production speed.* McGraw-Hill, New York, USA.

Kılıç Z (2020) The importance of water and conscious use of water, *International Journal of Hydrology*, 4(5):239–241.

Kung D et al. (2008) Application of lean thinking to improve the productivity of water and sewer service installations. *Canadian Journal of Civil Engineering*, 35: 418-430.

Lambert AO (2002) International report: water losses management and techniques. *Water Science and Technology: Water Supply*, *2*(4): 1-20.

Liu J, Yu G (2014) Analysis of demand and leakage distributing uniformly along pipes. *16th Conference on Water Distribution System Analysis*. Elsevier Ltd, Bari, Italy, vol. 89, pp. 603–612

LSSBBM (LSS Black Belt Manual) (2013) *Lean Six Sigma Black Belt Manual XL v11*. An e-Careers Limited Publication, Langley, UK.

McKone KE et al. (2001) The impact of total productive maintenance practices on manufacturing performance. *Journal of operations management,* 19 (1): 39-58.

Mostafa S et al. (2015) Lean thinking for a maintenance process. *Production & Manufacturing Research,* 3 (1): 236-272.

Ngadiman Y et al. (2012) A study of total productive maintenance implementation in manufacturing industry for generating greater profits. *Proceedings International Conference of Technology Management Business and Entrepreneurship 2012 (ICTMBE2012)*, Melaka, Malaysia, pp. 157-173.

Swamidass, P. M. (2002) *Innovations in competitive manufacturing.* Kluwer Academic Publisher, Norwell, USA

Woodall WH and Montgomery DC (2014) Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology*, *46*(1): 78-94.

water

MDPI

Article

Optimal Number of Pressure Sensors for Real-Time Monitoring of Distribution Networks by Using the Hypervolume Indicator

Bruno Ferreira [1,*], Nelson Carriço [1] and Dídia Covas [2]

[1] INCITE, Barreiro School of Technology, Polytechnic Institute of Setúbal, 2910-761 Setúbal, Portugal; nelson.carrico@estbarreiro.ips.pt
[2] CERIS, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal; didia.covas@tecnico.ulisboa.pt
* Correspondence: bruno.s.ferreira@estbarreiro.ips.pt

Abstract: This article proposes a novel methodology to determine the optimal number of pressure sensors for the real-time monitoring of water distribution networks based on a quality hypervolume indicator. The proposed methodology solves the optimization problem for different numbers of pressure sensors, assesses the gain of installing each set of sensors by means of the hypervolume indicator and determines the optimal number of sensors by the variation of the hypervolume indicator. The methodology was applied to a real case study. Several robustness analyses were carried out. The results demonstrate that the methodology is hardly influenced by the method parameters and that a reasonable estimation of the optimal number of sensors can be easily achieved.

Keywords: real-time monitoring; pressure sensors; optimization; multi-objective; hypervolume

Citation: Ferreira, B.; Carriço, N.; Covas, D. Optimal Number of Pressure Sensors for Real-Time Monitoring of Distribution Networks by Using the Hypervolume Indicator. *Water* **2021**, *13*, 2235. https://doi.org/10.3390/w13162235

Academic Editor: Athanasios Loukas

Received: 29 June 2021
Accepted: 15 August 2021
Published: 17 August 2021

1. Introduction

Water distribution networks (WDN) constitute vital infrastructure to ensure drinking water in sufficient quantity and quality for users. Real-time monitoring of these networks is essential to ensure the quality of the service provided. This is achieved by employing on-site sensors to measure certain water quality (e.g., pH, temperature, contaminants) and hydraulic (i.e., pressure and flowrate) parameters [1]. These sensors produce different types of data (e.g., flowrate, pressure, contaminants concentration) that can be used in different activities. Only a limited number of sensors can be installed in a given WDN due to budget constraints. Moreover, these WDN may rapidly increase in complexity (as a function of the area served), resulting in a challenging task of real-time monitoring of large networks when few sensors are installed [2].

Numerous methods have been developed over the recent decades to find the optimal location for sensors to detect contamination events in real time and mitigate the associated effects. In the Battle of the Water Sensor Networks [3], fifteen different approaches of optimal sensor location for contaminant detection were compared according to four different objectives. Aral et al. [4] blended four distinct criteria into one single objective which was solved using genetic algorithms. According to Weickgenannt et al. [5], the risk of contamination is explicitly evaluated as the product of likelihood of not detecting the contaminant intrusion and the corresponding consequence (water consumed). In a different approach, Zhao et al. [6] attempted to minimize the consumption of contaminated water prior to contamination detection. More recently, Naserizade et al. [7] used multi-objective optimization and a multicriteria decision-making technique whilst Ponti et al. [8] proposed a new multi-objective evolutionary algorithm which showed improvements over NSGA-II. However, sensor location techniques aiming at monitoring water quality parameters may not provide the best solutions for hydraulic parameters since they have different goals in the utility activities.

In particular, pressure data are widely used by water utilities in the real-time monitoring of WDN for creating alerts and use in hydraulic modelling calibration and pipe burst detection. Alarm systems can be designed to respond to the occurrence of abnormal pressure values (low or high pressure) at specific locations of the network [9]. The calibration process of hydraulic models requires pressure data from specific locations of the network to calibrate high uncertainty variables such as nodal demands and pipe roughness coefficient [2,10]. Pipe burst detection and location techniques use pressure data in a variety of ways, for instance, to carry out inverse analysis of the pipe burst location (using hydraulic simulation and optimization) [11,12] or by using a data-driven classifier approach [13].

Distinct methods have been proposed in the literature to optimally locate pressure sensors in WDN, most of them focusing on the total gain of sensor placement considering different aims. Cao et al. [14] positioned pressure sensors in such a way as to represent the pressure patterns of the homogeneous area of the WDN. Zhao et al. [15] aimed at maximizing the detection coverage rate of pipe burst events whilst Raei et al. [16] aimed at reducing the detection time of these events. Both Casillas et al. [17] and Steffelbauer et al. [18] focused on maximizing the percentage of leak scenarios correctly identified (according to the introduced criteria). Sarrate et al. [19] focused on maximizing the leak detectability performance. Furthermore, the maximization of nodal pressure sensitivities was considered both by de Schaetzen et al. [20] and Francés-Chust et al. [21]. The maximization of the accuracy of the hydraulic model was also considered by Kapelan et al. [22] and Behzadian et al. [23].

Ideally, the optimal location and number of pressure sensors should be determined simultaneously. Nonetheless, most of the previously presented methods are not able to directly provide the optimal number of pressure sensors to be installed. As a result, the number of pressure sensors to be installed in a WDN is usually defined using different criteria. Sarrate et al. [19] and Sanz et al. [24] defined the number of pressure sensors for pipe burst detection by budget limitations whilst both Sophocleous et al. [11] and Wu et al. [25] used a metric of the number of households per pipe unit length. For the same objective of pipe burst detection, Soldevila et al. [26] used a metric based on the network size, and Quintiliani et al. [27] used engineering good sense. Regarding pressure sensors for model calibration, de Schaetzen et al. [20] used the number of households per unit length to decide the number of sensors whilst Wéber et al. [2] emphasized the minimum percentage of calibration errors to be achieved.

This article presents a novel methodology to determine the optimal number of pressure sensors in a WDN. The major advantage is the possibility to be coupled with the existing optimization methods of sensor location (such as those previously presented), thus providing the optimal number according to the considered objective functions. Furthermore, the methodology does not require previous knowledge on sensor installation costs nor the consideration of costs in the optimization problem itself. The methodology uses the hypervolume indicator [28] to measure the total gain of installing different numbers of pressure sensors (according to the objective functions being considered) with a single value. In order to reduce the number of solved optimization problems, a discrete set of numbers of sensors was established in a distributed and representative way for which the optimization problem would be solved. A trade-off function was then fitted to the hypervolume data, which can be interpreted as a compromise between the total gain (using the hypervolume) and the number of sensors. Finally, the optimal number of pressure sensors was determined by assessing the evolution of the hypervolume indicator as a function of the number of sensors, specifically by identifying the point of maximum curvature of the hypervolume data. From that point onwards, the inclusion of another sensor was rather not worth the increment on the total gain.

This methodology is demonstrated through application to a real WDN. The pressure sensors were located using an optimization method based on the maximization of nodal sensitivities to both pipe burst events and variations of the pipe roughness coefficient. The obtained results are discussed and the most relevant conclusions are drawn.

The main novel contributions are as follows: (i) the proposal and application of the hypervolume indicator to quantify the total gain associated with a given number of pressure sensors and (ii) the application of different techniques to detect the point of maximum curvature aiming to determine the optimal number of pressure sensors.

2. Methodology

2.1. General Approach

A flowchart of the proposed methodology is present in Figure 1 and is composed of six main steps: (1) definition of the maximum number of sensors; (2) establishment of a discrete set of numbers of sensors; (3) optimization of the pressure sensor locations; (4) assessment of the total gain of sensors; (5) trade-off function calculation; (6) determination of the optimal number of pressure sensors. The following sections further detail each step of the proposed methodology.

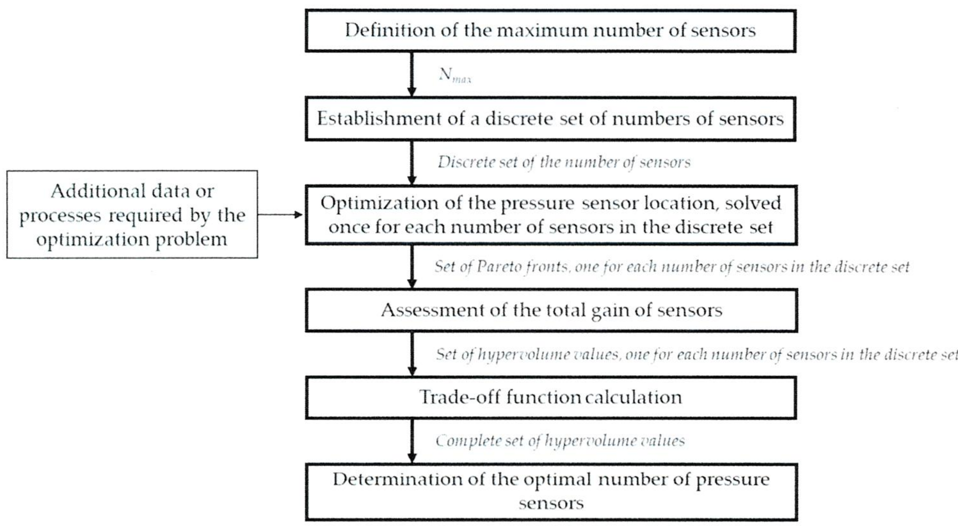


Figure 1. Flowchart of the proposed methodology.

2.2. Definition of the Maximum Number of Sensors

The optimal number of pressure sensors is determined by analyzing the variation of the total gain (using the results of optimization problems) with the number of sensors. However, solving optimization problems for an incremental number of sensors (e.g., between 1 and 150) can be computationally expensive due to the exponential increase in the search space. For example, considering 20, 21 and 22 sensors in a total set of 150 possible locations results in 3.63×10^{24}, 2.24×10^{25} and 1.31×10^{26} possible combinations of sensors, respectively. Therefore, the maximum number of installed sensors, N_{max}, is firstly defined, aiming at minimizing the computational burden by limiting the number of optimization problems to be solved.

Different approaches can be used for the definition of the maximum number of pressure sensors in a WDN, for instance, by using economic criteria based on the size of the WDN [18] or by determining when the gain of installation of an additional sensor has no significant improvement [15], taking into account the total number of possible locations [16] or the available budget [23,27].

2.3. Establishment of the Discrete Set of Numbers of Sensors

A discrete set of numbers of sensors is established between 1 and the maximum number of sensors N_{max}, for which the optimization problem will be solved. The established set does not have to be a continuous sequence of natural numbers, for instance, it could be 1, 5, 10, 15, … , N_{max}. The aim is to reduce the number of optimization problems to be solved whilst leaving enough results to characterize the total gain as a function of the number of sensors between 1 and N_{max}. As such, different sets of the number of sensors can be considered, for instance, with an evenly distributed number of sensors between 1 and N_{max} or with a higher density of observations in the lower number of sensors. The effect of different discrete sets of the numbers of sensors is assessed in this article.

2.4. Optimization of the Pressure Sensor Locations

The total gain with the installation of pressure sensors depends on the method used for locating those sensors. Different pressure sensor location methods can be coupled to the proposed methodology, based on different optimization problem formulations [2,15–23,29].

A method to optimally locate pressure sensors based on the maximization of nodal sensitivities is considered to demonstrate the performance of the proposed methodology. This method is based on the formulation of an unconstrained multi-objective optimization problem, in which the decision variables are the pressure sensor locations (using nodes as possible locations). It requires the prior computation of two pressure sensitivity matrices $S1$ and $S2$ (one for the pipe roughness coefficient and the other for the pipe burst size, respectively). The size of the obtained matrices is the number of pipes $\times$ the number of nodes and the number of nodes $\times$ the number of nodes, respectively. As such, $S1_{i,j}$ refers to the variation of the pressure in node j with respect to the variation of the pipe roughness coefficient of pipe i. Similarly, $S2_{i,j}$ refers to the variation of the pressure in node j with respect to the variation in the pipe burst size in node i. Further details regarding pressure sensitivity analysis can be found in the works by de Schaetzen et al. [20] and Lansey et al. [30].

Accordingly, two objective functions are formulated for a given sampling design X of sensor locations. The first function f_1 is defined according to de Schaetzen et al. [20] by using a compromise programming formulation. It aims at maximizing the sensitivity to the pipe roughness coefficient that the sensors cover (f_A) whilst ensuring the evenly spread geographical distribution of sensor locations by maximizing the entropy according to Shannon's definition (f_B). Functions f_A and f_B are defined as follows:

$$Maximize\ f_A = \sum_{i=1}^{Np} a_i, \quad where\ a_i = \max_{j \in X}\left(S1_{i,j}\right) \tag{1}$$

$$Maximize\ f_B = -\sum_{i=1}^{Np} p_i \times \log_2(p_i), \quad where\ p_i = \frac{a_i}{\sum_{i=1}^{Np} a_i} \tag{2}$$

where Np is the number of pipes.

Finally, the two functions f_A and f_B are combined together into f_1 by taking the weighted sum of the two functions.

The second function f_2 aims at directly maximizing the sensitivity of sensors to pipe burst events and can be computed as follows:

$$Maximize\ f_2 = \sum_{i=1}^{Nn} a_i, \quad where\ a_i = \max_{j \in X}\left(S2_{i,j}\right) \tag{3}$$

where Nn is the number of nodes.

Further details regarding the formulation of the objective functions can be found in the work by de Schaetzen et al. [20].

The optimization problem should be solved for each number of sensors in the discrete set. The results of each optimization problem are a single set of pressure sensor locations (for single objective optimization) or a set of optimal combinations of sensors (for multiple objective functions). In the latter case, a Pareto front of solutions is obtained; each solution of a Pareto front refers to a specific combination of sensors for which no other combination exists that presents better results for both objectives. Each optimal combination is characterized by the values associated with each objective function. These values are used to assess the gain of installing the different number of sensors in the WDN. Note that the same objective functions and optimization method should be used for all numbers of sensors in the discrete set.

2.5. Assessment of the Total Gain of Sensors

The result of the previous step is a Pareto front for each number of sensors in the discrete set. Each Pareto front embraces all the possible combinations of sensors for that specific number of sensors.

The assessment of the total gain for a given number of sensors is not a trivial task as the gain can be described by multiple objective functions and, therefore, can have multiple different values; consider the example presented in Figure 2a, where a single optimal configuration of five sensors is depicted with a red dot. The gain of that specific configuration of sensors is characterized by three different values in accordance with some hypothetical objective functions 1, 2 and 3 to maximize (these are deliberately different from f_1 and f_2 since the methodology can be adopted to different optimization problem formulations with different dimensions). Furthermore, multiple optimal configurations might exist for a given number of sensors due to a trade-off between the objectives. This is represented in Figure 2b with a Pareto front of four optimal combinations of five sensors depicted with red dots. Note that the increase of objective function 3 will lead to the decrease of objective functions 1 and 2 as a trade-off between these contradictory objectives.

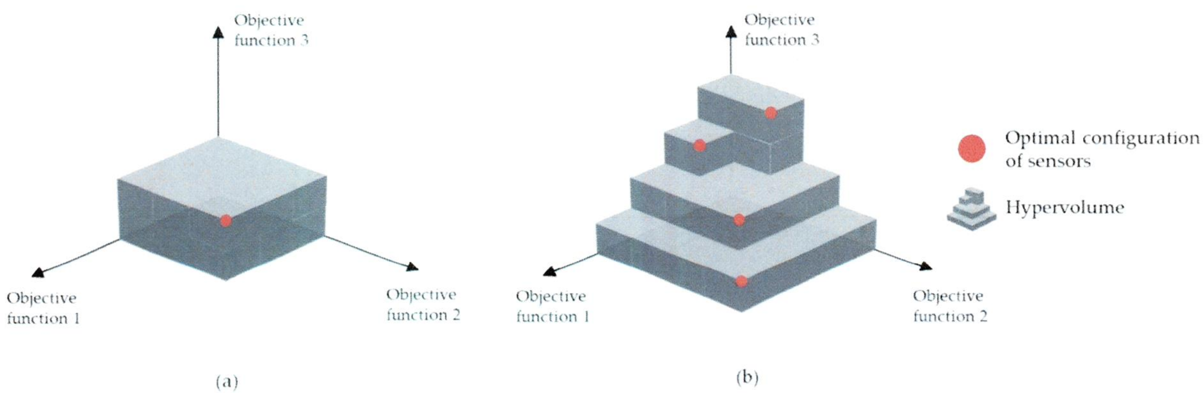


Figure 2. Characterization of the total gain for a given number of sensors using the hypervolume indicator.

In order to assess the total gain of installing a different number of pressure sensors, the quality of each obtained Pareto front is assessed using a quality measure [28,31–36]. This calculation leads to a single value for each Pareto front, characterizing the total gain associated with each specific number of sensors.

The use of the hypervolume indicator [28] is proposed herein to measure the quality of each Pareto front. It is one of the most widely used quality measures, the main advantages whereof being its easy interpretation and properties, such as guaranteeing strict monotonicity regarding Pareto dominance [37]. Nonetheless, this indicator also has disadvantages, such as sensitivity to the presence or absence of extreme points in a Pareto front [38]. In sum, this indicator considers the "volume" (i.e., in as many dimensions as the number of the objective functions) of the region of the objective space dominated by the set of optimal solutions in relation to the global worst point. Note that the hypervolume has no defined units, and its magnitude is related to the objective functions used. The hypervolume is graphically depicted in Figure 2 as a grey volume and considering the global worst point as (0,0,0). The formal mathematical definition of the hypervolume can be found in the work by Guerreiro et al. [37].

Different methods of the hypervolume calculation have been developed [37–42], with different efficiency levels against high dimensionality problems [37]. The efficiency of the hypervolume calculation in this particular methodology is not a major concern because no high dimensionality is expected and a reduced number of hypervolume calculations is required (specifically, one calculation for each Pareto front).

2.6. Trade-Off Function Fitting

The result of the previous step is a set of hypervolume values, with one value associated with each number of sensors in the discrete set previously defined in Section 2.3. As the number of installed pressure sensors increases, the hypervolume also increases since more sensors lead to a higher total gain. This can be interpreted as a trade-off between the total gain of installing pressure sensors (using the hypervolume) and the number of sensors.

Therefore, observations of the hypervolume are used to derive a trade-off function that describes the behavior of the hypervolume (interpreted as the total gain) as a function of the number of sensors. By doing so, it is possible to estimate the total gain associated with the number of sensors for which the optimization problem is not solved.

Different algorithms can be used to fit a function to the hypervolume values by solving the nonlinear least-squares problem. A review on numerical methods for nonlinear least-squares optimization problems can be found in the work by Yuan [43]. The use of the Levenberg–Marquardt method is proposed to optimally fit a function by finding its optimal set of parameters. It is a well-established method with fast convergence, although it may be sensitive to the initial parameters guess.

Distinct trade-off functions (with different mathematical formulations and numbers of parameters) can be fitted to the same hypervolume data. In this study, the well-known root-mean-square error (RMSE) is used to assist in deciding which function (after the fitting) best describes the hypervolume data. The smaller the value of the RMSE is, the better the fitted function represents the behavior of the hypervolume.

The selected fitted function is used to derive estimations of the hypervolume for the complete set of numbers of sensors between 1 and N_{max}.

2.7. Determination of the Optimal Number of Pressure Sensors

The hypervolume estimations are analyzed in order to determine the optimal number of sensors. This optimal number corresponds to the point where the inclusion of another sensor does not significantly improve the increment on the total gain. Such a point is defined as the point of maximum curvature in the hypervolume curve. The mathematical formulation for the point of maximum curvature can be found in the work by Satopaa et al. [44]. This point is defined herein as the "knee/elbow" of the trade-off function and can be identified using different automatic techniques [44–47]. The two techniques used herein are based on the works of Satopaa et al. [44] and Salvador and Chan [46], respectively. The performance of both techniques is assessed later in the article.

The first technique, the Kneedle method, starts by rescaling the estimated hypervolume data to [0,1], leading to the normalized hypervolume values, HV_{norm}. The number of sensors between 1 and N_{max} should also be rescaled to [0,1], leading to the normalized number of sensors values, N_{norm}. These values are represented in Figure 3 as black circles and squares, respectively, considering $N_{max} = 15$ in this example. Finally, the differences Δ between HV_{norm} and N_{norm} are calculated, as exemplified with grey arrows and triangular markers in Figure 3. The optimal number of sensors, N_{opt}, is selected as the number of sensors that presents the maximum difference.

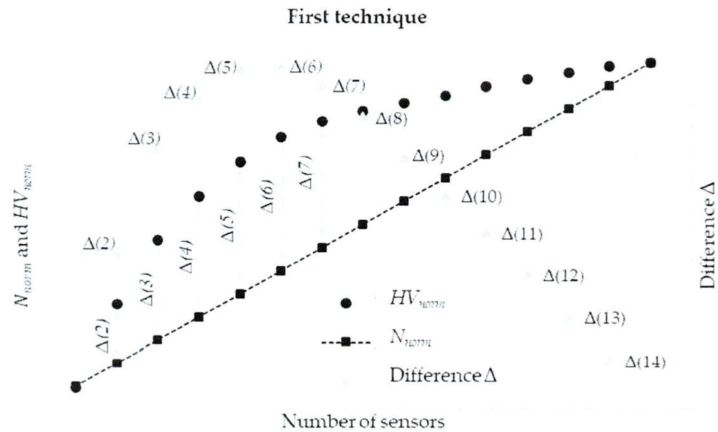


Figure 3. Determination of the optimal number of sensors based on the evolution of the hypervolume by using the Kneedle method.

The second technique, the L-method, determines the "knee/elbow" as the point that best divides the hypervolume data into two straight lines. Pairs of lines are fitted to the hypervolume data using a linear regression procedure, and their junction point is the looked-for "knee/elbow". Each line must contain at least two points and start at either end of the hypervolume data. Figure 4 shows all the possible pairs of straight lines for an example of the hypervolume data considering N_{max} = 7. The seven hypervolume values are presented both as blue or orange circles (either they are on the left or right side of the junction point). The corresponding fitted straight lines are represented with blue or red dashed lines, respectively. The number of sensors tested for the optimal number is in a blue diamond. The optimal number of sensors is the junction point that minimizes the weighted RMSE for the two linear parts of the hypervolume data.

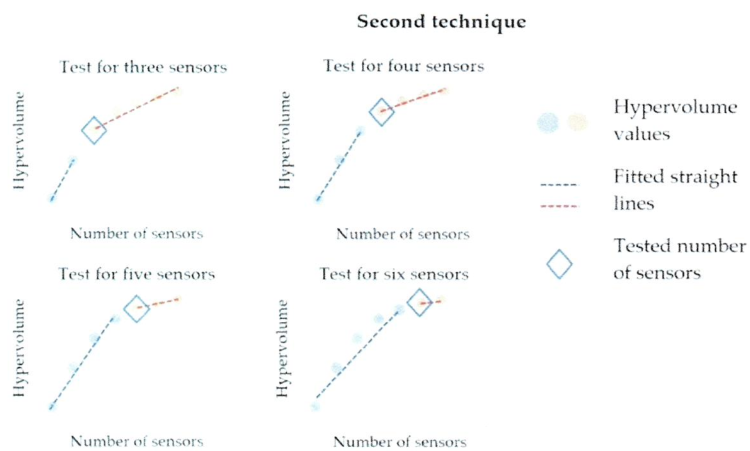


Figure 4. Determination of the optimal number of sensors based on the evolution of the hypervolume using the second technique.

The final optimal locations can be found in the Pareto front related to the optimal number of sensors N_{opt}. Nonetheless, the Pareto front might not have yet been obtained (i.e., when N_{opt} was not considered in the discrete set). In these cases, the optimization problem should be solved for the number of sensors equal to N_{opt}.

3. Case Study

This methodology is demonstrated in a real WDN with a total network extension of ca. 70 km, 2200 service connections and an average inlet flowrate of about 60 m^3/h. The supplied area consists mainly of single-family houses with irrigated gardens, most of them with a swimming pool. The WDN also supplies a few golf courses. The hydraulic simulation model was developed in EPANET [48] and included four storage tanks, 4474 pipes and 4429 nodes, of which 2200 were consumers with an individual hourly demand pattern. Figure 5 presents the layout of the hydraulic network.

Figure 5. Layout of the case study.

The maximum number of installed sensors, N_{max}, was determined based on the intrinsic network characteristics, namely, on network length. A maximum of one sensor per kilometer of pipe network was adopted according to several Portuguese expert opinions, leading to N_{max} = 70 sensors (i.e., one sensor per km).

Five discrete sets of numbers of sensors were defined. The objective was to assess if major differences in the optimal number of sensors were found when considering distinct sets of numbers of sensors. These five sets presented distinct characteristics and were defined as follows:

$$Set\ 1 = [1, 10, 20, 30, 40, 50, 60, 70] \tag{4}$$

$$Set\ 2 = [1, 10, 30, 50, 70] \tag{5}$$

$$Set\ 3 = [1, 10, 20, 30, 70] \tag{6}$$

$$Set\ 4 = [1, 5, 10, 15, 20, 25, 30, 70] \tag{7}$$

$$Set\ 5 = [1, 2, 3, \ldots, 24, 25, 70] \tag{8}$$

Note that both *Set* 1 and *Set* 2 attempted to describe evenly distributed observations of the hypervolume between 1 and N_{max} whilst *Set* 3, *Set* 4 and *Set* 5 incorporated a higher density in the lower number of sensors. Furthermore, the number of observations highly varied between the sets: *Set* 2 and *Set* 3 contained only five numbers of sensors (leading to five optimization problems to be solved) whilst *Set* 5 contained 26 numbers of sensors (leading to a total of 26 optimization problems).

The method described in Section 2.4 was used to optimally locate the pressure sensors. Note that other methods can be used with distinct objective functions (both in number and formulation). Two pressure sensitivity matrices were computed prior to the optimization problem itself. A variation of Hazen-Williams's pipe roughness coefficient of 10 was considered for the computation of the first matrix. The second matrix was obtained by generating a burst of fixed size for every node of the hydraulic model with a single emitter coefficient of 0.25 (leading to an average burst of 5 L/s). The hydraulic simulations were carried in EPANET [48].

The optimal sensor location for a given number of sensors was formulated as an unconstrained multi-objective optimization problem. The decision variables were the nodes where pressure sensors could be potentially installed, and all the nodes were considered as possible locations. The two objective functions presented in Section 2.4 were considered, aiming at the maximization of nodal pressure sensitivities to both pipe roughness coefficient variations (f_1) and pipe burst events (f_2).

The problem was solved once for each number of sensors in each of the five discrete sets. When the number of sensors appeared in more than one set (e.g., 70), the problem was only solved once. The NSGA-II algorithm was applied in the Python environment using the Pymoo package [49] to solve several optimization problems. Discrete variables and integer encoding were used for problem formulation; each discrete location (i.e., node) was translated by an integer value. A population member is a set of locations for the pressure sensors and each variable of a population member represents a possible pressure sensor location.

The following NSGA-II parameters were used: random integer sampling and selection operators; integer polynomial mutation with the probability p_m = 0.05 and index parameter n_m = 20; simulated binary crossover with probability p_c = 0.95 and index parameter n_c = 20.

A population size of 100 was considered and all NSGA-II runs were carried out for 500 generations. As such, each NSGA-II run (i.e., solving an optimization problem for a given number of sensors in a given set) resulted in 500 × 100 = 50,000 objective function evaluations. The optimization problems were solved using an Intel Core i5-8250U processor of 1.80 GHz and 8 GB of memory, with a total running time of around 1 h per optimization problem. Table 1 presents the number of objective function evaluations associated with each of the five discrete sets; compare the computation effort (i.e., the number of objective function evaluations) in *Set* 2 and *Set* 3 (both with five optimization problems to be solved) with that of *Set* 5 (with 26 optimization problems to be solved).

Table 1. Number of objective function evaluations for each discrete set.

Sets of Numbers of Sensors	Number of Observations	Objective Function Evaluations
$Set\ 1 = [1, 10, 20, 30, 40, 50, 60, 70]$	8	400,000
$Set\ 2 = [1, 10, 30, 50, 70]$	5	250,000
$Set\ 3 = [1, 10, 20, 30, 70]$	5	250,000
$Set\ 4 = [1, 5, 10, 15, 20, 25, 30, 70]$	8	400,000
$Set\ 5 = [1, 2, 3, \ldots, 24, 25, 70]$	26	1,300,000

The result of each optimization problem was a Pareto front of optimal pressure sensor locations, according to the objective functions f_1 and f_2. For the sake of simplicity, the obtained Pareto fronts for 10, 20, 30 and 40 sensors are depicted in Figure 6. Each circle represents an optimal configuration of either 10, 20, 30 or 40 sensors.

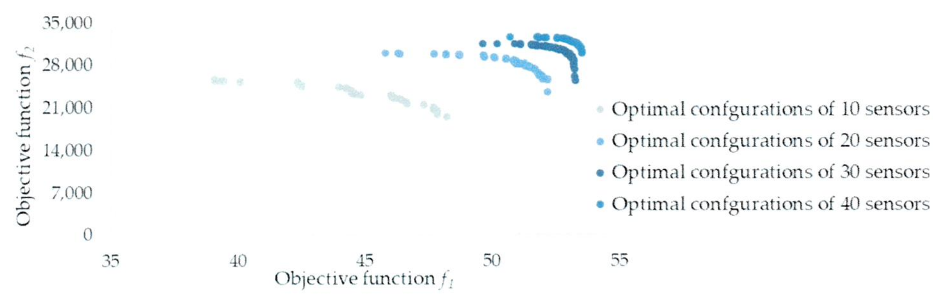


Figure 6. Example of obtained Pareto fronts for 10, 20, 30 and 40 sensors.

The hypervolume indicator was calculated for each Pareto front of each set of numbers of sensors, leading to a total of five sets of Hypervolume values (i.e., one for each set of numbers of sensors). The global worst point was considered equal to (0,0) since both functions aimed at the maximization of the objective function. The calculation method for the hypervolume was based on variant 3 of the algorithm proposed by Fonseca et al. [42]. The five sets of hypervolume values are depicted in Figure 7 as black triangles.

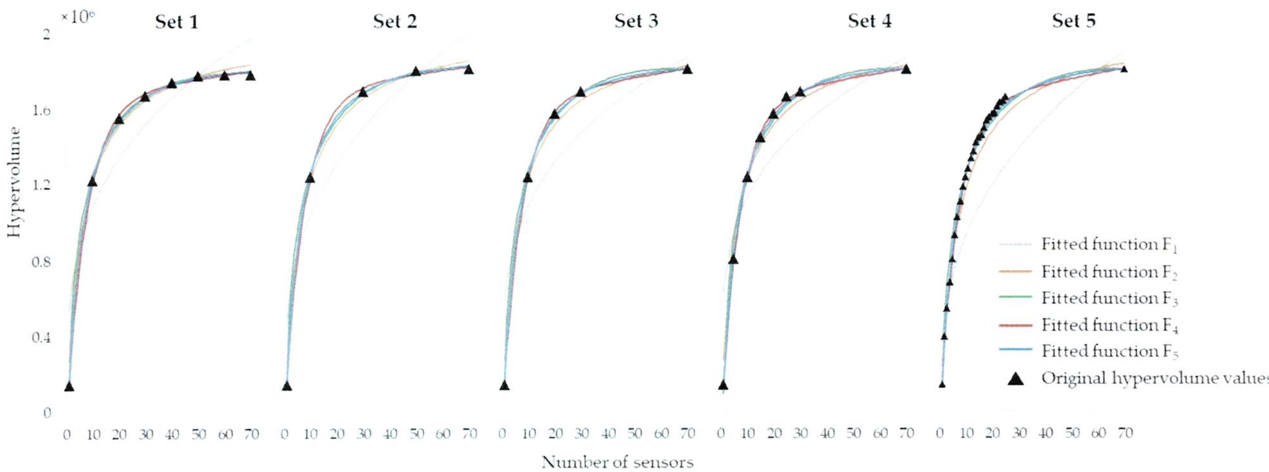


Figure 7. Estimated hypervolume values using different fitted functions F_i and different discrete sets of numbers of sensors.

Five different functions, F_i, were considered to describe the hypervolume data as a function of the number of sensors, N:

$$F_1(N) = aN^b \tag{9}$$

$$F_2(N) = (aN)/(b + N) \tag{10}$$

$$F_3(N) = a + blog_{10}(N) - c\,N \tag{11}$$

$$F_4(N) = a\,e^{b\,N} + c\,e^{d\,N} \tag{12}$$

$$F_5(N) = a(N + b)^c + d \tag{13}$$

where a, b, c and d are the function parameters. These parameters, once optimized, can represent the curve of the hypervolume as a function of the number of sensors. Note that different functions can be considered to describe the increasing behavior one can see in Figure 7 in black triangles.

The Levenberg–Marquardt method was used to fit each function F_i to each set of hypervolume values by finding the optimal set of parameters. This was done by using MATLAB's Curve Fitting Toolbox. A reasonable initial parameter guess was found by coarsely gridding the parameter space. This was carried out to cope method sensitivity to the initial parameters. Figure 7 presents in colored lines the estimated hypervolume values obtained by using each model F_i. The original hypervolume values are represented as triangular black markers.

Table 2 presents the five functions' specific parameters found using the Levenberg–Marquardt method for *Set* 1 of hypervolume values.

Table 2. Fitted functions' parameters for *Set* 1 of hypervolume values.

Fitting Function	a	b	c	d
$F_1(N) = aN^b$	532,809	0.308	-	-
$F_2(N) = (aN)/(b + N)$	2,007,063	6.465	–	–
$F_3(N) = a + blog_{10}(N) - c\,N$	144,141	1,190,045	7835	–
$F_4(N) = a\,e^{b\,N} + c\,e^{d\,N}$	1,665,878	0.001	−1,733,051	−0.129
$F_5(N) = a(N + b)^c + d$	−136,128,665	10.779	−1.772	1,861,283

The RMSE was calculated for each function F_i of each set of numbers of sensors in order to assist in deciding which function F_i best describes the hypervolume data. The obtained results are presented in Table 3, with the best fitted function for each set of numbers of sensors presented in bold and shaded areas. Finally, the optimal number of sensors N_{opt} (considered as the point of maximum curvature) was determined using an automatic detection technique. The two distinct techniques presented in Section 2.7 were used for each function F_i of each set of numbers of sensors, with the obtained results presented in Table 3. Figure 8 presents the application of both Kneedle method and L-method techniques to the estimated hypervolume for the F_5 function of *Set* 1.

Table 3. RMSE and the optimal number of sensors N_{opt} for the different fitted functions and discrete sets of sensors. The results for the best-fitted function of each discrete set are presented in bold and shaded area.

Number of Sensors' Set		$F_1(N)$	$F_2(N)$	$F_3(N)$	$F_4(N)$	$F_5(N)$
Set 1 = [1, 10, 20, 30, 40, 50, 60, 70]	RMSE	2.2×10^5	6.0×10^4	1.9×10^4	2.0×10^4	$\mathbf{1.5 \times 10^4}$
	N_{opt} (Kneedle method)	20	17	16	18	**17**
	N_{opt} (L-method)	20	16	16	17	**16**
Set 2 = [1, 10, 30, 50, 70]	RMSE	2.7×10^5	7.6×10^4	$\mathbf{2.2 \times 10^4}$	3.0×10^4	2.6×10^4
	N_{opt} (Kneedle method)	20	18	**16**	18	17
	N_{opt} (L-method)	20	17	**16**	16	16
Set 3 = [1, 10, 20, 30, 70]	RMSE	3.3×10^5	8.9×10^4	2.9×10^4	$\mathbf{1.0 \times 10^4}$	3.3×10^4
	N_{opt} (Kneedle method)	19	17	16	**17**	17
	N_{opt} (L-method)	20	16	16	**16**	16
Set 4 = [1, 5, 10, 15, 20, 25, 30, 70]	RMSE	2.8×10^5	7.9×10^4	6.1×10^4	$\mathbf{1.4 \times 10^4}$	2.2×10^4
	N_{opt} (Kneedle method)	19	17	16	**17**	17
	N_{opt} (L-method)	19	16	16	**16**	16
Set 5 = [1, 2, 3, ..., 24, 25, 70]	RMSE	4.1×10^5	1.3×10^5	6.4×10^4	$\mathbf{1.5 \times 10^4}$	2.4×10^4
	N_{opt} (Kneedle method)	20	18	16	**17**	17
	N_{opt} (L-method)	21	17	16	**16**	16

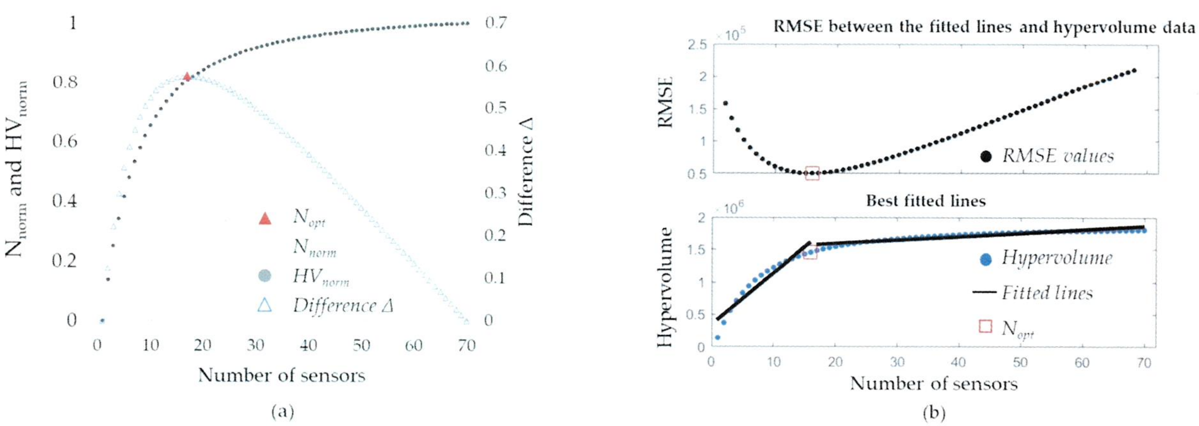


Figure 8. Optimal number of sensors for the F_5 function of *Set* 1 using (**a**) the Kneedle method and (**b**) the L-method.

4. Discussion

The increase in the number of installed sensors led to a higher total gain. Nonetheless, the marginal gain decreased when considering additional sensors. This is visible in Figure 7 in black triangles. This figure shows that regardless of the considered number of sensors in the set, the function F_1 (grey line) clearly fails to mathematically describe the trend of the hypervolume as a function of the number of sensors, whereas the remaining functions can approximately describe the hypervolume variation.

Table 3 presents better insights regarding the fitting capabilities of the different functions F_i for each discrete set of numbers of sensors. The best fit corresponds to the smallest values of the RMSE. Complex models F_4 and F_5 (i.e., both with a larger number of parameters) presented the best fitting results (in bold) for most sets of numbers of sensors, except for *Set* 3, for which F_3 led to the lowest RMSE values.

The first robustness analysis was carried out to determine the effect of the different functions F on the optimal number of sensors. For each technique (Kneedle method and L-method) of each set of numbers of sensors, it was possible to verify that the optimal number of sensors did not significantly vary with the fitted functions F_i to estimate the hypervolume value (note that results from F_1 were excluded from analysis). Technique 1 presented higher variability in the optimal number of sensors, ranging from 16 to 18. Technique 2 presented smaller variability, between 16 and 17. Therefore, and as long as selected functions F_i (and respective parameters) can properly represent the hypervolume as a function of the number of sensors, a good estimation of N_{opt} can be obtained (considering either the Kneedle method or the L-method).

The second robustness analysis was carried out to analyze the effect of different discrete sets of numbers of sensors on the optimal number of sensors. For each technique and function F_i, it was possible to verify that the optimal number was relatively stable for the different sets, and no differences greater than one were found. Thus, and as long as the selected discrete set of numbers of sensors can properly represent the hypervolume as a function of the number of sensors, a good estimation of N_{opt} can be obtained. This allows for a significant reduction in the number of optimization problems to be solved as similar results are obtained when both five (in *Set* 2 and *Set* 3) or 26 (in *Set* 5) optimization problems are solved.

The third robustness analysis was carried out to assess the effect of the two different "knee/elbow" detection techniques (Kneedle method and L-method) on the optimal number of sensors. No differences greater than two in the optimal number of sensors were found by comparing the pairs of results for both techniques in each discrete set of numbers of sensors and function F_i as the optimal number varied between 16 and 18 sensors. Furthermore, when the best function F_i was considered for each set (results in bold in Table 3), the optimal number of sensors was equal to 16 or 17 (depending on the "knee/elbow" detection technique) regardless of the discrete set of numbers of sensors. Based on these results, the Kneedle method was recommended given the straightforward implementation and easier concept understanding.

Overall, the proposed methodology was robust regarding different discrete sets of numbers of sensors, fitting functions and "knee/elbow" automatic detection techniques. The optimal number of sensors for the best fitting function was equal to 16 or 17 (depending on the "knee/elbow" detection method) regardless of the chosen discrete set of numbers of sensors. This allowed for a great reduction in the computational burden associated with the overall analysis as similar results were obtained when considering both *Set* 2 and *Set* 3 (with five optimization problems to be solved) and *Set* 5 (26 optimization problems to be solved). Furthermore, a good approximation between 16 and 18 sensors was found considering any combination of function F (with the exception of F_1), any discrete set of numbers of sensors and any "knee/elbow" detection technique.

Based on these results, the optimal number of sensors was selected as 16. The optimization problem was already solved for 16 sensors (as part of *Set* 5), and the obtained Pareto front is presented in Figure 9a. Each grey point in Figure 9a represents a combination of 16 sensors. Figure 9b depicts a possible optimal location for the 16 sensors in green triangles, associated with the highlighted solution from the Pareto front.

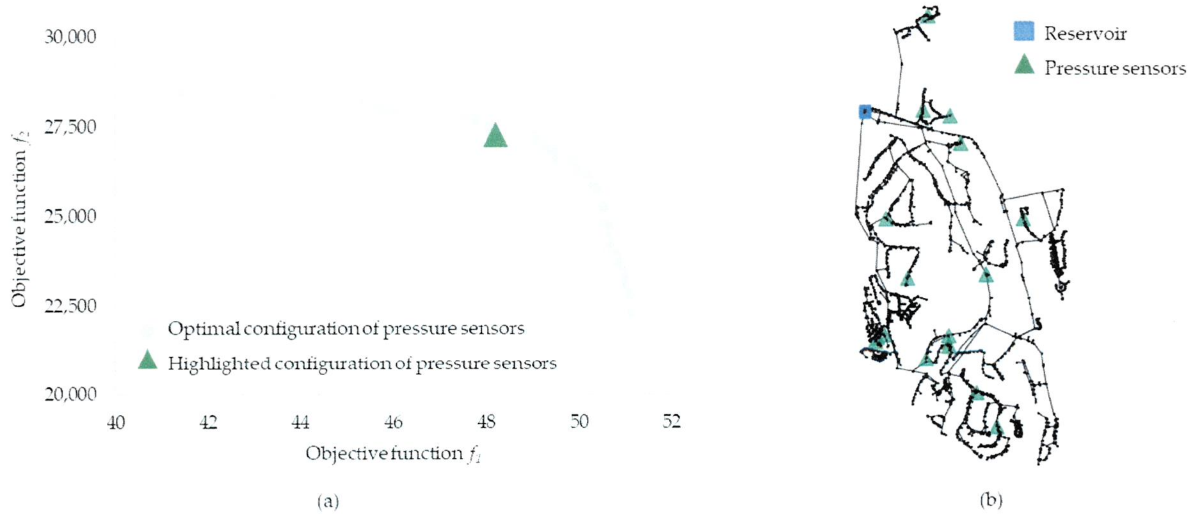


Figure 9. (**a**) Pareto front for 16 sensors; (**b**) optimal pressure sensor location for 16 sensors.

5. Conclusions

This article proposes a six-step methodology to determine the optimal number of pressure sensors. The major advantage of the proposed methodology is that it does not require neither prior knowledge about sensor installation costs nor the consideration of costs in the optimization problem itself to decide on the optimal number of sensors. Furthermore, the methodology can be used with different optimization methods of the pressure sensor location.

In the proposed methodology, the optimal number of pressure sensors is determined by analyzing the relationship between the results of optimization problems and the number of sensors. The total gain of installing a given number of sensors is characterized by using the hypervolume indicator. A trade-off function is derived to describe the gain of installing sensors as a function of the number of sensors. Finally, a technique is used to determine the optimal number of sensors as the point where the inclusion of another sensor is not worth the increment on the total gain.

A real case study was used to demonstrate, assess and discuss the performance of the proposed methodology. Three distinct robustness analyses were carried out and, according to the obtained results, the following key conclusions related to the proposed methodology can be drawn:

- The hypervolume indicator can be used to characterize the total gain of installing pressure sensors whose locations are obtained by solving multi-objective optimization problems.
- A trade-off function can be derived, allowing the characterization of the total gain of installing sensors (i.e., hypervolume) as a function of the number of sensors. This trade-off function fitting process allows for a great reduction in the computational effort associated with the overall analysis, as similar results are obtained when solving both five and 26 multi-objective optimization problems.
- The obtained results are not affected by considering different trade-off functions or different sets of numbers of sensors as long as the selected function and the selected discrete set can properly represent the hypervolume as a function of the number of sensors.
- Both Techniques 1 and 2 lead to similar results. The use of Technique 1 (Kneedle method) is further recommended given the straightforward implementation and easier concept understanding.

In future work, the robustness of the proposed methodology could be further assessed by comparing the obtained results with those obtained by using two distinct methods of locating those sensors. Furthermore, different quality measures to compare consecutive Pareto fronts and additional techniques to automatically detect the "knee/elbow" could be compared.

Author Contributions: Conceptualization, B.F.; methodology, B.F.; investigation, B.F.; writing—original draft preparation, B.F.; writing—review and editing, N.C. and D.C.; visualization, B.F.; supervision, N.C. and D.C.; project administration, N.C.; funding acquisition, N.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fundação para a Ciência e a Tecnologia through a studentship (reference number SFRH/BD/149392/2019) and the WISDom project (reference number DSAIPA/DS/0089/2018). The APC was funded completely by the WISDom project.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Kara, S.; Karadirek, I.E.; Muhammetoglu, A.; Muhammetoglu, H. Real time monitoring and control in water distribution systems for improving operational efficiency. *Desalin. Water Treat.* **2016**, *57*, 11506–11519. [CrossRef]
2. Wéber, R.; Hős, C. Efficient Technique for Pipe Roughness Calibration and Sensor Placement for Water Distribution Systems. *J. Water Resour. Plan. Manag.* **2020**, *146*, 04019070. [CrossRef]
3. Ostfeld, A.; Uber, J.G.; Salomons, E.; Berry, J.W.; Hart, W.E.; Phillips, C.A.; Watson, J.-P.; Dorini, G.; Jonkergouw, P.; Kapelan, Z.; et al. The Battle of the Water Sensor Networks (BWSN): A Design Challenge for Engineers and Algorithms. *J. Water Resour. Plan. Manag.* **2008**, *134*, 556–568. [CrossRef]
4. Aral, M.M.; Guan, J.; Maslia, M.L. Optimal design of sensor placement in water distribution networks. *J. Water Resour. Plan. Manag.* **2010**, *136*, 5–18. [CrossRef]
5. Weickgenannt, M.; Kapelan, Z.; Blokker, M.; Savic, D.A. Risk-Based Sensor Placement for Contaminant Detection in Water Distribution Systems. *J. Water Resour. Plan. Manag.* **2010**, *136*, 629–636. [CrossRef]
6. Zhao, Y.; Schwartz, R.; Salomons, E.; Ostfeld, A.; Poor, H.V. New formulation and optimization methods for water sensor placement. *Environ. Model. Softw.* **2016**, *76*, 128–136. [CrossRef]
7. Naserizade, S.S.; Nikoo, M.R.; Montaseri, H. A risk-based multi-objective model for optimal placement of sensors in water distribution system. *J. Hydrol.* **2018**, *557*, 147–159. [CrossRef]
8. Ponti, A.; Candelieri, A.; Archetti, F. A New Evolutionary Approach to Optimal Sensor Placement in Water Distribution Networks. *Water* **2021**, *13*, 1625. [CrossRef]
9. Pérez-Padillo, J.; Morillo, J.G.; Ramirez-Faz, J.; Roldán, M.T.; Montesinos, P. Design and implementation of a pressure monitoring system based on iot for water supply networks. *Sensors* **2020**, *20*, 4247. [CrossRef]
10. Zanfei, A.; Menapace, A.; Santopietro, S.; Righetti, M. Calibration procedure for water distribution systems: Comparison among hydraulic models. *Water* **2020**, *12*, 1421. [CrossRef]
11. Sophocleous, S.; Savić, D.; Kapelan, Z. Leak Localization in a Real Water Distribution Network Based on Search-Space Reduction. *J. Water Resour. Plan. Manag.* **2018**, *145*, 04019024. [CrossRef]
12. Hajibandeh, E.; Nazif, S. Pressure Zoning Approach for Leak Detection in Water Distribution Systems Based on a Multi Objective Ant Colony Optimization. *Water Resour. Manag.* **2018**, *32*, 2287–2300. [CrossRef]
13. Sun, C.; Parellada, B.; Puig, V.; Cembrano, G. Leak Localization in Water Distribution Networks Using Pressure and Data-Driven Classifier Approach. *Water* **2019**, *12*, 54. [CrossRef]
14. Cao, H.; Hopfgarten, S.; Ostfeld, A.; Salomons, E.; Li, P. Simultaneous sensor placement and pressure reducing valve localization for pressure control of water distribution systems. *Water* **2019**, *11*, 1352. [CrossRef]
15. Zhao, M.; Zhang, C.; Liu, H.; Fu, G.; Wang, Y. Optimal sensor placement for pipe burst detection in water distribution systems using cost–benefit analysis. *J. Hydroinform.* **2020**, *22*, 606–618. [CrossRef]
16. Raei, E.; Shafiee, M.E.; Nikoo, M.R.; Berglund, E. Placing an ensemble of pressure sensors for leak detection in water distribution networks under measurement uncertainty. *J. Hydroinform.* **2019**, *21*, 223–239. [CrossRef]
17. Casillas, M.; Puig, V.; Garza-Castañón, L.; Rosich, A. Optimal Sensor Placement for Leak Location in Water Distribution Networks Using Genetic Algorithms. *Sensors* **2013**, *13*, 14984–15005. [CrossRef] [PubMed]

18. Steffelbauer, D.B.; Fuchs-Hanusch, D. Efficient Sensor Placement for Leak Localization Considering Uncertainties. *Water Resour. Manag.* **2016**, *30*, 5517–5533. [CrossRef]
19. Sarrate, R.; Blesa, J.; Nejjari, F.; Quevedo, J. Sensor placement for leak detection and location in water distribution networks. *Water Supply* **2014**, *14*, 795–803. [CrossRef]
20. de Schaetzen, W.B.; Walters, G.; Savic, D. Optimal sampling design for model calibration using shortest path, genetic and entropy algorithms. *Urban Water* **2000**, *2*, 141–152. [CrossRef]
21. Francés-Chust, J.; Brentan, B.M.; Carpitella, S.; Izquierdo, J.; Montalvo, I. Optimal Placement of Pressure Sensors Using Fuzzy DEMATEL-Based Sensor Influence. *Water* **2020**, *12*, 493. [CrossRef]
22. Kapelan, Z.S.; Savic, D.A.; Walters, G.A. Optimal Sampling Design Methodologies for Water Distribution Model Calibration. *J. Hydraul. Eng.* **2005**, *131*, 190–200. [CrossRef]
23. Behzadian, K.; Kapelan, Z.; Savic, D.; Ardeshir, A. Stochastic sampling design using a multi-objective genetic algorithm and adaptive neural networks. *Environ. Model. Softw.* **2009**, *24*, 530–541. [CrossRef]
24. Sanz, G.; Pérez, R.; Kapelan, Z.; Savic, D. Leak Detection and Localization through Demand Components Calibration. *J. Water Resour. Plan. Manag.* **2016**, *142*, 04015057. [CrossRef]
25. Wu, Z.Y.; Sage, P.; Turtle, D. Pressure-Dependent Leak Detection Model and Its Application to a District Water System. *J. Water Resour. Plan. Manag.* **2010**, *136*, 116–128. [CrossRef]
26. Soldevila, A.; Blesa, J.; Tornil-Sin, S.; Fernandez-Canti, R.M.; Puig, V. Sensor placement for classifier-based leak localization in water distribution networks using hybrid feature selection. *Comput. Chem. Eng.* **2018**, *108*, 152–162. [CrossRef]
27. Quintiliani, C.; Vertommen, I.; van Laarhoven, K.; van der Vliet, J.; van Thienen, P. Optimal Pressure Sensor Locations for Leak Detection in a Dutch Water Distribution Network. *Environ. Sci. Proc.* **2020**, *2*, 2040. [CrossRef]
28. Auger, A.; Bader, J.; Brockhoff, D.; Zitzler, E. Theory of the hypervolume indicator: Optimal μ-Distributions and the Choice of the Reference Point. In Proceedings of the Tenth ACM SIGEVO workshop on Foundations of Genetic Algorithms—FOGA '09, Orlando, FL, USA, 9–11 January 2009; ACM Press: New York, NY, USA, 2009; pp. 87–102.
29. Blesa, J.; Nejjari, F.; Sarrate, R. Robust sensor placement for leak location: Analysis and design. *J. Hydroinform.* **2016**, *18*, 136–148. [CrossRef]
30. Lansey, K.E.; El-Shorbagy, W.; Ahmed, I.; Araujo, J.; Haan, C.T. Calibration Assessment and Data Collection for Water Distribution Networks. *J. Hydraul. Eng.* **2001**, *127*, 270–279. [CrossRef]
31. Laszczyk, M.; Myszkowski, P.B. Survey of quality measures for multi-objective optimization: Construction of complementary set of multi-objective quality measures. *Swarm Evol. Comput.* **2019**, *48*, 109–133. [CrossRef]
32. Wang, S.; Ali, S.; Yue, T.; Li, Y.; Liaaen, M. A practical guide to select quality indicators for assessing pareto-based search algorithms in search-based software engineering. In Proceedings of the 38th International Conference on Software Engineering, Austin, TX, USA, 14–22 May 2016; ACM: New York, NY, USA, 2016; pp. 631–642.
33. Tan, K.C.; Lee, T.H.; Khor, E.F. Evolutionary algorithms for multi-objective optimization: Performance assessments and comparisons. In Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No.01TH8546), Seoul, Korea, 27–30 May 2001; IEEE: Piscataway, NJ, USA, 2002; Volume 2, pp. 979–986.
34. Bandyopadhyay, S.; Pal, S.K.; Aruna, B. Multiobjective GAs, Quantitative Indices, and Pattern Classification. *IEEE Trans. Syst. Man Cybern. Part B* **2004**, *34*, 2088–2099. [CrossRef]
35. Collette, Y.; Siarry, P. Three new metrics to measure the convergence of metaheuristics towards the Pareto frontier and the aesthetic of a set of solutions in biobjective optimization. *Comput. Oper. Res.* **2005**, *32*, 773–792. [CrossRef]
36. Zitzler, E.; Deb, K.; Thiele, L. Comparison of Multiobjective Evolutionary Algorithms: Empirical Results. *Evol. Comput.* **2000**, *8*, 173–195. [CrossRef] [PubMed]
37. Guerreiro, A.P.; Fonseca, C.M.; Paquete, L. The Hypervolume Indicator: Problems and Algorithms. *arXiv* **2020**, arXiv:2005.00515.
38. While, L.; Hingston, P.; Barone, L.; Huband, S. A faster algorithm for calculating hypervolume. *IEEE Trans. Evol. Comput.* **2006**, *10*, 29–38. [CrossRef]
39. Daulton, S.; Balandat, M.; Bakshy, E. Differentiable Expected Hypervolume Improvement for Parallel Multi-Objective Bayesian Optimization. *arXiv* **2020**, arXiv:2006.05078.
40. Jaszkiewicz, A. Improved quick hypervolume algorithm. *Comput. Oper. Res.* **2018**, *90*, 72–83. [CrossRef]
41. Zhao, G.; Arroyave, R.; Qian, X. Fast Exact Computation of Expected Hypervolume Improvement. *arXiv* **2018**, arXiv:1812.07692.
42. Fonseca, C.M.; Paquete, L.; López-Ibáñez, M. An improved dimension-sweep algorithm for the hypervolume indicator. In Proceedings of the 2006 IEEE International Conference on Evolutionary Computation, Vancouver, BC, Canada, 16–21 July 2006; pp. 1157–1163. [CrossRef]
43. Yuan, Y.-X. Recent advances in numerical methods for nonlinear equations and nonlinear least squares. *Numer. Algebr. Control Optim.* **2011**, *1*, 15–34. [CrossRef]
44. Satopaa, V.; Albrecht, J.; Irwin, D.; Raghavan, B. Finding a "Kneedle" in a Haystack: Detecting Knee Points in System Behavior. In Proceedings of the 2011 31st International Conference on Distributed Computing Systems Workshops, Minneapolis, MN, USA, 20–24 June 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 166–171.
45. Zhao, Q.; Hautamaki, V.; Fränti, P. Knee Point Detection in BIC for Detecting the Number of Clusters. In *Advanced Concepts for Intelligent Vision Systems (ACIVS'08), Proceedings of the 10th International Conference, ACIVS 2008, Juan-les-Pins, France, 20–24 October 2008*; Springer: Berlin/Heidelberg, Germany, 2008; Volume 5259, pp. 664–673; ISBN 3540884572.

46. Salvador, S.; Chan, P. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms. In Proceedings of the 16th IEEE International Conference on Tools with Artificial Intelligence, Boca Raton, FL, USA, 15–17 November 2004; IEEE: Piscataway, NJ, USA, 2004; pp. 576–584.
47. Antunes, M.; Gomes, D.; Aguiar, R.L. Knee/Elbow Estimation Based on First Derivative Threshold. In Proceedings of the 2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService), Bamberg, Germany, 26–29 March 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 237–240.
48. Rossman, L. *EPANET 2 User's Manual*; United States Environmental Protection Agency: Washington, DC, USA, 2000.
49. Blank, J.; Deb, K. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* **2020**, *8*, 89497–89509. [CrossRef]

Control Engineering Practice 19 (2011) 1157–1167

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

Methodology for leakage isolation using pressure sensitivity analysis in water distribution networks

Ramon Pérez [a,*], Vicenç Puig [a,b], Josep Pascual [a], Joseba Quevedo [a], Edson Landeros [c], Antonio Peralta [d]

[a] Advanced Control Systems Group (SAC), Universitat Politècnica de Catalunya (UPC), Rambla Sant Nebridi, 10, 08222 Terrassa, Spain
[b] IRI Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Spain
[c] CETAQUA Water Technological Center, Spain
[d] AGBAR Barcelona Water Company, Spain

ARTICLE INFO

Article history:
Received 6 September 2010
Accepted 2 June 2011
Available online 12 July 2011

Keywords:
Pressure sensitivity
Leakage localisation
Sensor placement

ABSTRACT

Leaks are present to some extent in all water-distribution systems. This paper proposes a leakage localisation method based on the pressure measurements and pressure sensitivity analysis of nodes in a network. The sensitivity analysis using analytical tools is not a trivial job in a real network because of the huge non-explicit non-linear systems of equations that describe its dynamics. Simulations of the network in the presence and the absence of leakage may provide an approximation of this sensitivity. This matrix is binarised using a threshold independent of the node. The binary matrix is assumed as a signature matrix for leakages. However, there is a trade-off between the resolution of the leakage isolation procedure and the number of available pressure sensors. In order to maximise the isolability with a reasonable number of sensors, an optimal sensor placement methodology, based on genetic algorithms, is also proposed. These methodologies have been applied to the Barcelona Network using PICCOLO simulator. The sensor placement and the leakage detection and localisation methodologies are applied to several district management areas (DMA) in simulation and in reality.

1. Introduction

Water loss in distribution system networks is an issue of great concern for water utilities, strongly linked with operational costs and water resource savings. Continuous improvements in water loss management are applied and new technologies are developed to achieve higher levels of efficiency. Usually a leakage detection method in a District Metered Area (DMA) starts by analysing input flow data, such as minimum night flows and consumer metering data (Lambert, 1994; MacDonald, 2005). Once the water distribution district is identified to have a leakage, various techniques are used to locate the leakage for pipe replacement or repair. Methods for locating leaks range from ground-penetrating radar to acoustic listening devices or physical inspection (Colombo, Lee, & Karney, 2009; Farley & Trow, 2003). Some of these techniques require isolating and shutting down part of the system. The whole process could take weeks or months with a significant volume of water wasted. Techniques based on locating leaks from pressure monitoring devices allow a more effective and less costly search in situ.

This paper presents a model-based methodology to detect and localise leaks. It has been developed within a project carried out by Aguas Barcelona, Water Technological Centre CETaqua, and the Technical University of Catalonia (UPC). The objective of this project is to develop and apply an efficient system to detect and locate leaks in a water distribution network. It integrates methods and technologies available and in use by water companies, including DMA and flow/pressure sensor data, in conjunction with mathematical hydraulic models. The method is based on the analysis of pressure variations produced by leakage in the water distribution network (Pudar & Ligget, 1992). This technique differs from others in the literature, such as the reflection method (LRM) or the inverse transient analysis (ITA), since it is not based on the transient analysis of pressure waves (Ferrante & Brunone, 2003a, 2003b; Misiunas, Lambert, Simpson, & Olsson, 2005; Verde, Visairo, & Gentil, 2007). Alternatively, the leakage detection procedure is performed by comparing real pressure and flow data with their estimation using the simulation of the mathematical network model. Simulation of the network in presence and absence of leakage provides an approximation of pressure sensitivity of nodes in a network when a leak is present in a node. The approximation is used to generate a sensitivity matrix that is binarised using a threshold independent of the node. In order to successfully apply this methodology, the characterisation of district metered areas and consumers, considered a critical issue for a correct model calibration, should be also addressed but is not described in this paper (see, e.g. Perez, de las Heras, Aguilar,

* Corresponding author. Tel.: +34937398620; fax: +349373928.
E-mail address: ramon.perez@upc.edu (R. Pérez).

0967-0661/$ - see front matter
doi:10.1016/j.conengprac.2011.06.004

Pascual, & Peralta, 2009a, for further details). Another critical point is the data validation of DMA sensors that can be addressed as it is described for flowmeters in Quevedo et al. (2010). The paper also proposes a methodology for placing pressure sensors within a DMA that optimises leakage detection using a minimum number of sensors based on the approach proposed in Pérez et al. (2009b). Finally, the leakage detection methodology proposed will be tested with sensors installed in a DMA used as case study.

Section 2 reviews water distribution network modelling and presents the case study used to illustrate the proposed methodologies. Model-based fault detection and isolation techniques described in Section 3 are used for the leakage detection and location. Section 4 presents how the leak signature matrix is obtained from the pressure sensitivity matrix. Since the sensor placement is a critical issue for maximising discriminability, an algorithm is presented in Section 5. The signature matrix is generated for the set of sensors selected. This matrix has to be compared with the signature obtained comparing the model and the real measurements. From this comparison, the leakage is located in a set of possible nodes. This methodology is presented in Section 6 and is illustrated by simulation and real results. Finally, Section 7 summarises the conclusions.

Fig. 1. Case study network: *Plaça del Diamant*.

2. Water distribution systems: plaça del diamant case study

A water distribution system consists of three major components: pumps, distribution storage, and distribution piping network. Most systems require pumps to supply lift to overcome differences in elevation, and energy losses caused by friction. Pipes may contain flow-control devices, such as regulating or pressure-reducing valves (Brdys & Ulanicki, 1994). The purpose of a distribution system is to supply the system's users with the amount of water demanded, under adequate pressure for various loading conditions. A loading condition is a spatial pattern of demands that defines the users' flow requirements.

2.1. Mathematical modelling

The governing laws for flow in pipe systems under steady conditions are conservation of mass and energy. The law of conservation of mass states that the rate of storage in a system is equal to the difference between the inflow to and outflow from the system. In pressurised water distribution networks, no storage can occur within the pipe network, although tank storage may change over time. Therefore, in a pipe, or a junction node, the inflow and the outflow must balance. For a junction node

$$\sum q_{in} - \sum q_{out} = q_{ext} \tag{1}$$

where q_{in} and q_{out} are the pipe flow rates into and out of the node and q_{ext} is the external demand or supply. Conservation of energy states that the difference in energy between two points is equal to the energy added to the flow in components between these points minus the frictional losses. An energy balance can be written for paths between the two end points of a single pipe, between two fixed graded nodes (a node for which the total energy is known, such as a tank) through a series of pipes, valves, and pumps, or around a loop that begins and ends at the same point. In a general form for any path

$$\sum_{i \in J_p} h_{P,j} - \sum_{i \in I_p} h_{L,i} = \Delta E \tag{2}$$

where $h_{L,i}$ is the headloss across component i along the path, $h_{P,j}$ is the head added by pump j, and ΔE is the difference in energy between the end points of the path. The primary network component is a pipe. The relationship between pipe flow (q) and energy loss caused by friction (h_L) in individual pipes can be represented by a number of equations, including the Darcy–Weisbach and Hazen–Williams equations. The general relationship is of the following form:

$$h_L = Kq^r \tag{3}$$

where K is a pipe coefficient that depends on the pipe's diameter, length, and material and r is an exponent in the range of 2.

2.2. Plaça del Diamant DMA case study

The case study used to illustrate the leak localisation methodology presented in this paper is based on Plaça del Diamant DMA at the Barcelona Water Network (see Fig. 1). This DMA is used for illustrating the methodology. Its model contains 1600 nodes and 41.153 m of pipes. This DMA is simulated using PICCOLO software. Demands are assumed to occur in the nodes. In this paper, it will also be assumed that leaks occur at the nodes. Such assumption introduces a minor imprecision compared with those due to the methodology and the uncertainty of the model itself. Distance from the real leakage to the closest junction is much shorter than the diameter of the search zone obtained in the best case. It will be clear with results because the areas obtained include some pipes and nodes. Under such assumption, leaks can be seen as additional demands but with unknown location and quantity.

Simulated leaks introduced in the network are of 1 l/s, more or less 3% of the total demand of the sector (in the nighttime). The demand distribution all over the network is the most variable parameter of the model. Some uncertainty in the demand has also been included in order to test the robustness of the method.

3. Leakage detection and isolation methodology foundations

The methodology of leakage localisation proposed in this paper is mainly based on standard theory of model-based diagnosis described for example in (Gertler, 1998) that has already been applied to water networks to detect faults in flow metres (Ragot & Maquin, 2006) or in open channel with dynamic models (Bedjaoui & Weyer, 2011; Nejjari, Pérez, Escobet, & Traves, 2006).

Model-based diagnosis can be divided in two subtasks: fault detection and fault isolation. The principle of model-based fault detection is to check the consistency of observed behaviour while fault isolation tries to isolate the component that is in fault. The consistency check is based on computing residuals, $\mathbf{r}(k)$, obtained from measured input signals $\boldsymbol{u}(k)$ and outputs $\boldsymbol{y}(k)$ using the sensors installed in the monitored system and the analytical relationship which are obtained by system modelling:

$$\mathbf{r}(k) = \boldsymbol{\Psi}(\boldsymbol{y}(k), \boldsymbol{u}(k)) \tag{4}$$

where $\boldsymbol{\Psi}$ is the residual generator function that depends on the type of detection strategy used (parity equation (Gertler, 1998) or observer (Chen & Patton, 1999)). At each time instance, k, the residual is compared with a threshold value (zero in ideal case or almost zero in real case). The threshold value is typically determined using statistical or set-based methods that take into account the effect of noise and model uncertainty (Blanke, Kinnaert, Lunze, & Staroswiecki, 2006). When a residual is bigger than the threshold, it is determined that there is a fault in the system; otherwise, it is considered that the system is working properly. In practice, because of input and output noise, nuisance inputs and modelling errors affecting the considered model, robust residual generators must be used. The robustness of a fault detection system means that it must be only sensitive to faults, even in the presence of model-reality differences (Chen & Patton, 1999).

Robustness can be achieved at residual generation (active) or evaluation phase (passive). Most of the passive robust residual evaluation methods are based on an adaptive threshold changing in time according to the plant input signal and taking into account model uncertainty either in the time or frequency domain (Puig, Quevedo, Escobet, Nejjari, & de las Heras, 2008). In this paper, a passive method in time domain has been proposed for robust fault detection, where the detection threshold has been obtained using the method described in Section 4. Robust residual evaluation allows obtaining a set of observed fault signatures $\boldsymbol{\phi}(k) = [\phi_1(k), \phi_2(k), \cdots, \phi_{n_\phi}(k)]$, where each indicator of fault is obtained as follows:

$$\phi_i(k) = \begin{cases} 0 & \text{if } |r_i(k)| \le \tau_i(k \\ 1 & \text{if } |r_i(k)| > \tau_i(k) \end{cases} \tag{5}$$

where τ_i is the threshold associated to the residual $r_i(k)$ generated from sensor i.

Fault isolation involves identifying the faults affecting the system. It is carried out on the basis of observed fault signatures, $\boldsymbol{\phi}$, generated by the detection module and its relation with all the considered faults, $\boldsymbol{f}(k) = \{f(k)_1, f_2(k), \ldots, f_{n_f}(k)\}$ that are compared with theoretical signature matrix FSM (Gertler, 1998). One element of this matrix FSM_{ij} will be equal to one, if a fault $f_j(k)$ is affected by the residual $r_i(k)$. In this case, the value of the fault indicator $\phi_i(k)$ must be equal to one when the fault appears in the monitored system. Otherwise, the element FSM_{ij} will be zero. A given fault $f_i(k)$ is proposed as a fault candidate when the observed fault signature matches with its theoretical fault signature.

4. Leakage sensitivity analysis

The theoretical signature matrix needed to apply the isolation method presented in previous section can be obtained from a leakage sensitivity analysis. This analysis evaluates the effect of a leakage on the pressure in a node. If this process is repeated for each node and possible leak, the sensitivity matrix (Pudar & Ligget, 1992) is obtained as follows:

$$S = \begin{pmatrix} \frac{\partial p_1}{\partial f_1} & \cdots & \frac{\partial p_1}{\partial f_n} \\ \cdot & \cdots & \cdot \\ \frac{\partial p_n}{\partial f_1} & \cdots & \frac{\partial p_n}{\partial f_n} \end{pmatrix} \tag{6}$$

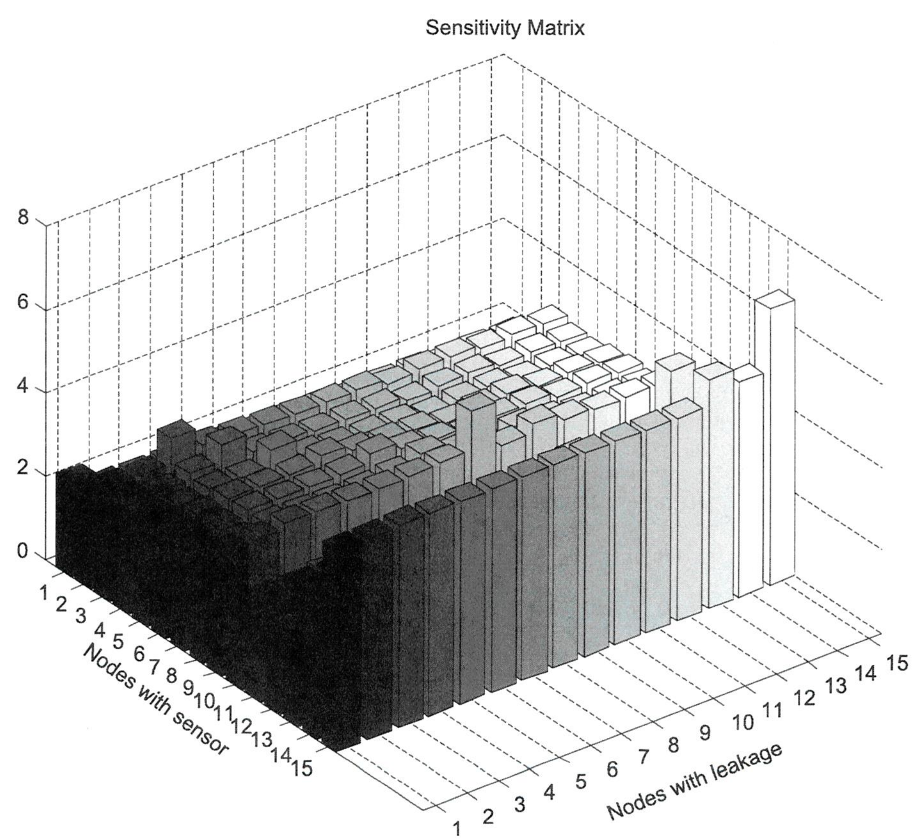


Fig. 2. Sensitivity matrix.

where each element s_{ij} measures the effect of leak f_j in the pressure of node p_i. It is extremely difficult to calculate S analytically in a real network because a water network is a large scale problem described by a multivariable non-linear and non-explicit system of equations as described in Section 2. This work proposes instead generating the sensitivity matrix by simulation as follows: The same leakage is introduced in each node and the increment of pressure is measured in each node. It implies 1600 simulations where 1600 pressures are measured. It has been verified that the analytical and the simulated sensitivity converge for small leakages. The sensitivity matrix depends on the working point that is, on the demand and boundary conditions (Vento & Puig, 2009).

In Fig. 2, the sensitivity matrix for the case study network of Fig. 1 is shown graphically. It has been plotted for 15 nodes distributed homogenously in the DMA as illustration.

Some sensors are much more sensitive to all leakages than others. Thus, a normalisation of sensitivity is needed so that the information provided by any node is comparable. Each row corresponding to a node with a sensor is divided by the maximum value of this row that corresponds to the leakage most important for that node. This procedure leads to the normalised sensitivity matrix:

$$\overline{S} = \begin{pmatrix} \frac{s_{11}}{\sigma_1} & \cdots & \frac{s_{1n}}{\sigma_1} \\ \cdot & \cdots & \cdot \\ \frac{s_{n1}}{\sigma_n} & \cdots & \frac{s_{nn}}{\sigma_n} \end{pmatrix} \tag{7}$$

where $\sigma_i = \max\{s_{i1},\ldots,s_{in}\}$, $i=1,\ldots,n$. This matrix is shown in Fig. 3 for the considered example. It shows how the most relevant leak is the one on the node itself, the maximum normalised sensitivity is on the diagonal. Columns correspond to nodes with leak and rows correspond to nodes with sensors.

Finally, from the normalised sensitivity matrix (7), the FSM matrix introduced in Section 3 can be derived. Each element FSM_{ij} is equal to zero when leakage j does not affect pressure in node i and it is equal to 1 when leakage j affects node i. The aim is to generate the signature matrix from the normalised sensitivity matrix. In Fig. 3, it can be seen that all leakages affect all pressures. Algorithm 1 presents how the Binarised Sensitivity Matrix ($\overline{S}^b$) is generated.

A process inspired in the ε-method proposed by Sezer and Siljak (1986) is proposed with the aim of identifying the strongest relations between leaks and pressure measurements. In this process, it is absolutely essential to choose conveniently the threshold that controls if a leak has or not an effect on a given pressure. The process proceeds as follows: those leaks that have an effect less than the given threshold are considered as a '0' in the leak signature matrix (5). Otherwise, their effect is considered as a '1'. In this way, the sensitivity matrix is binarised based on the selected threshold. Normalisation allows using a unique threshold for all sensors but the choice of the threshold is most relevant in the process. For small thresholds, all binarised matrix elements are 1 and only detection is possible. As the threshold increases more 0s appear. When threshold approaches 1, then only the diagonal of the signature matrix is 1 and localisation is perfect (or almost perfect, simulation precision makes some nodes equally sensitive to some leakages) but all sensors are needed. Fig. 4 shows how the number of 1s decreases as threshold approaches 1. Number of signatures increases but the significance of each sensor decays.

Algorithm 1. Binarised Sensitivity Matrix Generation **is**

input: n_y is the number of sensors, n_f is the number of leaks
$d(k)$ are the DMA demands
$p(k)$ are the boundary pressures
s_{th} is the binarisation threshold
k time instant when sensitivity matrix is calculated

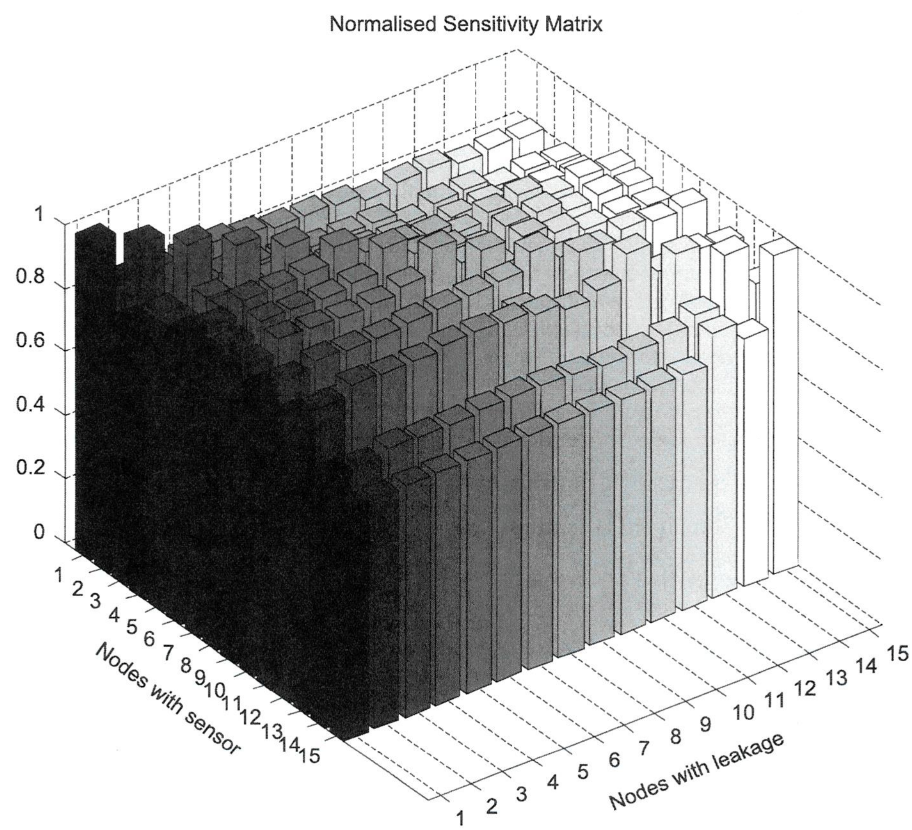


Fig. 3. Normalised sensitivity matrix.

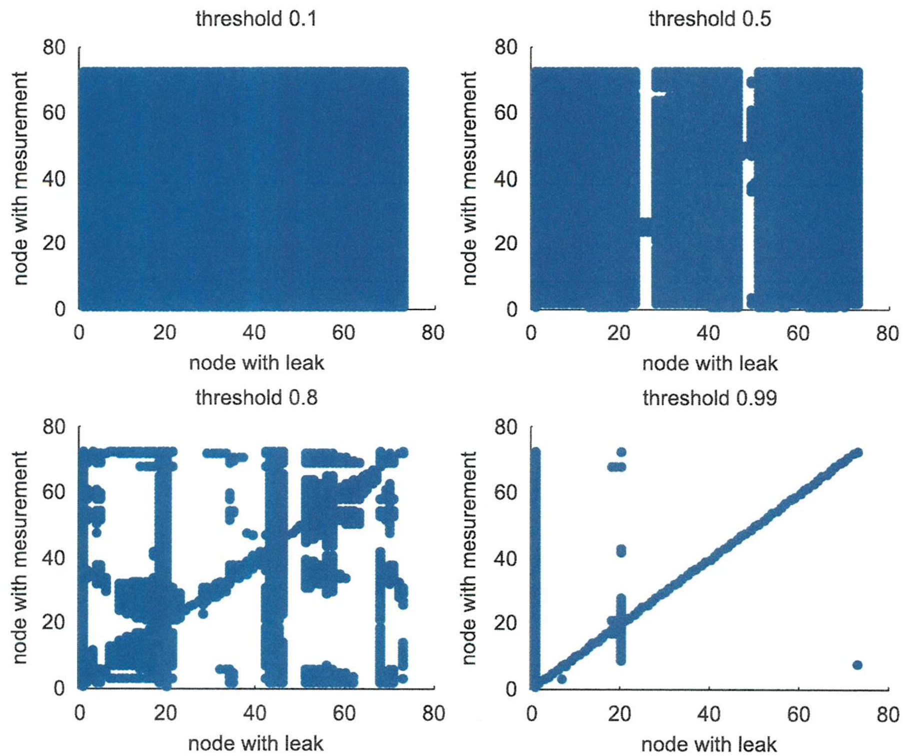


Fig. 4. Number of 1s and 0s depending on threshold.

output: $\vec{S}^b$ and $\sigma_{i...ny}$

for each sensor $i=1,\ldots,n_y$

compute the simulated pressures without leak $\widehat{y}_{i,0}(k)$

for each leak $j=1,\ldots,n_f$

compute the simulated pressures with leak at node j $\widehat{y}_{i,j}(k)$

compute $s_{ij}(k)=\widehat{y}_{i,j}(k)-\widehat{y}_{i,0}(k)$

end

end

for each sensor $i=1,\ldots,n_y$

$\sigma_i(k)=\max_{j=1\ldots nf}(s_{i,j}(k))$

for each leak $j=1,\ldots,n_f$

$\bar{s}_{i,j}(k)=\frac{s_{i,j}(k)}{\sigma_i(k)}$

if $\bar{s}_{i,j}(k)<s_{th}$

$\bar{s}_{i,j}{}^b(k)=0$

else

$\bar{s}_{i,j}{}^b(k)=1$

end

end

end

return $\vec{S}^b(k)$ and $\sigma_{i...ny}(k)$

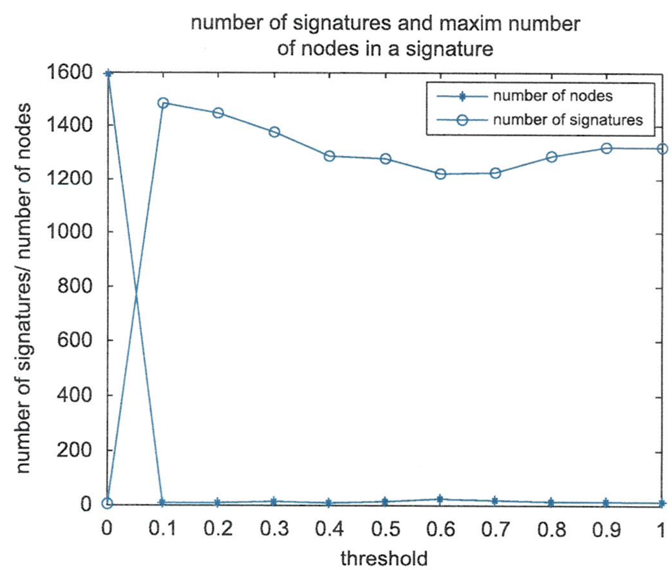


Fig. 5. Evolution of the signature matrix depending on threshold.

Fig. 5 shows the evolution of the number of signatures present in the matrix and the maximum number of leakages with the same signature. It corresponds to the 1613 nodes of the network in Fig. 1. Theoretically with 11 sensors (rows) there may be 2047 (that corresponds to $2^{11}-1$ since signature with all 0 is discarded as detection is imposed) different signatures for leakages (columns). In order to get maximum number of signatures, a necessary condition is to have in each column 2^{n-1} 1s, where n is the number of sensors (rows). This necessary condition is fulfilled for the threshold where both lines in Fig. 5 cross ($\sim$0.1). This is the threshold used.

Algorithm 2 summarises the leakage detection and isolation procedure using the binarised sensitivity matrix.

Algorithm 2. Leakage Detection and Isolation

input: n_y is the number of sensors, n_f is the number of leaks

N is the time horizon

$d(0)\ldots d(N)$ are the demands

$p(0)\ldots p(N)$ are the boundary pressures

$y_{i...ny}(0)\ldots y_{i...ny}(\mathrm{N})$ are the measured pressures

s_{th} is the binarisation threshold determined using Algorithm 1
output: $f_{1...nf}$ contains the number of incidences of each leak detected in the time horizon N.
initialise $f_{1...nf}=0$
for each instant $k=0,\ldots,N$
  for each sensor $i=1,\ldots,n_y$
   compute simulated pressures without leak $\hat{y}_{i,0}(k)$
   evaluate the residual $r_i(k)=y_i(k)-\hat{y}_{i,0}(k)$
   compute $\overline{S}^b(k)$ and $\sigma_{i\ldots ny}(k)$ from Algorithm 1
   the normalised residual $\overline{r}_i(k)=\frac{r_i(k)}{\sigma_i(k)}$
   if $\overline{r}_i(k)<s_{i,th}$
    $\phi_i(k)=0$
   else
    $\phi_i(k)=1$
   end
  end
  for each leak $j=1,\ldots,n_f$
   if HammingDistance($\phi(k)$,$\overline{S}^b(:\ j)(k)$) $\neq 0$
    $f_j=f_j$
   else
    $f_j=f_j+1$
   end
  end
end
return $f_{1...nf}$

5. Sensor placement algorithm

An optimal sensor placement is defined as a sensor configuration that achieves the minimum economical cost (number of sensors) while observing pre-specified performance criteria (groups of nodes that are not isolable with a minimum number of elements). Since this issue has been addressed in many applications in particular in Song, Chen, Sastry, and Tas (2009) a good literature review is provided.

A model of water network can be represented as a graph $G=(V,E)$, where E is the set of edges that represent the pipes and V is the set of vertices (nodes) where pipes meet. Vertices can represent sources, such as reservoirs or tanks, where water is introduced or sinks (demand points) where water is consumed. Each pipe connects two vertices v_i and v_j and usually is denoted as (v_i,v_j).

Using the graph representation, the problem of optimal sensor placement can be formulated as an integer programming problem, where each decision variable x_i associated to a node v_i of the network can be 1 or 0, meaning that the sensor will be or will not be installed in this node (Bagajewicz, 2000). The starting point of the algorithm is the leakage sensitivity matrix obtained by simulation binarised using the process described in Section 4. Every row corresponds to a hypothetical position of a sensor in a node while every column corresponds to a possible leak in a node. Thus, if a given element of this binary matrix contains a "1", it means that installing a sensor in the node corresponding to this row it would be able to detect the fault associated to the column of this element assuming a single leakage. A particular distribution of sensors (solution) is achieved by instantiating the value of decision variables x_i to "1" (meaning installing the sensor) or to "0" (meaning non installing the sensor). For any particular distribution, a set of groups of indiscernible leaks appear, each group with n_i leaks. The objective of the sensor placement algorithm is to find the sensor distribution that minimises the number of elements for the largest set of leaks with the same signature. The objective (cost) function is therefore

$$J=\min_{x_1,\cdots,x_n}\max\{n_1,\cdots,n_{n_f}\} \tag{8}$$

where $x_1,\ldots,x_n$ are the decision variables that determines a particular sensor distribution and n_i is the number of nodes in group i of indiscernible nodes for a given leakage f_i. In order to increase isolability, this cost should be minimised but at the same time keeping the economical cost reasonable, that is installing the less number of sensor that is possible. The problem is solved for a number of sensors; this number is increased till the cost does not decrease subtantialy. A constraint is included such that all leaks should be detected. It is introduced by forcing that signature with all 0s is not accepted.

This optimisation problem can be solved using either deterministic method based for example in Branch and Bound or heuristic methods based for example in Genetic Algorithms. The first type of methods guarantee the optimal solution but the computation time tends to be exponential with the number of nodes/faults (Sarrate, Puig, Escobet, & Rosich, 2007). On the other hand, the second type of methods just guarantees a suboptimal solution that tends to the optimal one when the size of considered population tends to infinity. Besides the formulation of solutions in series of 1s and 0s are most convenient for a GA. Algorithm 3 describes in detail how the optimal sensors distribution is done.

Algorithm 3. Optimal Sensor Distribution

input: n_f are the number of leaks (nodes), n_y are the number of sensors
 d are the DMA demands
 p are the boundary pressures
output: sensors x, and cost $J(n_y)$
 Solve $\min_x(J)$
 subject:
$\sum_{i=1}^{n_f} x_i=n_y$
 where the cost function J is computed using Algorithm 4
return x

Algorithm 4. Cost function J

input: $\overline{S}^b$ is the binarised sensitivity matrix
  x are the optimisation variables
output: J is the cost of the solution $j=0$
for each node $i=1\ldots n_f$
 if $x(i)==1$
   $\overline{S}^b{}_x(:\ j)=\overline{S}^b(:\ ,i)$
   $j=j+1$
 end
end
for each leak $j=1,\ldots,n_f$
   $m(j)=\mathrm{dec}(\overline{S}^b(:\ j))$
 where dec is the conversion of binary to decimal.
end
for $i=1\ldots 2^{ny}$
   n_i=number of i in m
 end
 $J=\max\ (n_i)$
return J

In Fig. 6, the evolution of cost function is presented. The cost has been taken as the number of nodes in the biggest group of possible leakage isolated with a number of sensors and a threshold between 0.1 and 0.4. The set of sensors should, in the leakage localisation, signal a group of nodes that may include a leak. Optimisation tries that the size of this group is as small as possible. A sharp improvement appears with the first sensors but adding more than 7 or 8 sensors introduce little improvement for any threshold. Therefore only 8 sensors are used.

In Fig. 7, the different groups of nodes with the same leakage signature are shown. There are 39 groups and the biggest contains 190 nodes. The localisation of the sensors after the optimisation process is presented in the last figure.

In an ideal situation with a well calibrated network model, a leakage should be searched in one of these regions instead of the whole sector. It is important to note that regions are connected and geographically coherent. Such coherence is a major issue for further search in situ. For further details see Pérez et al. (2009b).

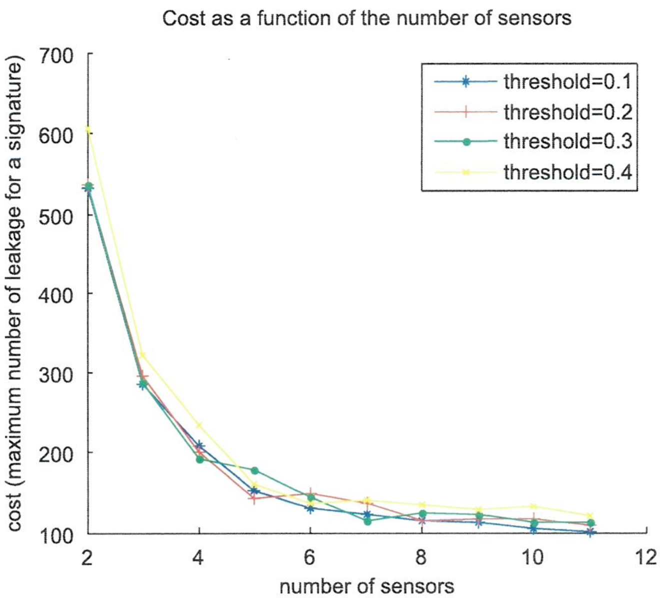


Fig. 6. Evolution of the cost function depending on number of sensors and threshold.

6. Leak isolation results

6.1. Simulation results

The proposed approach of localisation of leakages is first applied in simulation to the *Plaça del Diamant* using the optimal distribution of the sensors obtained in Section 5 consisting in 8 sensors. The process of leak localisation is based on Algorithm 2. If the model were perfect (no uncertainty in demands) and no noise, the leak should be localised with one measurement. However, because of modelling uncertainty and noise, the test has been done during 15 days of simulation (only the lowest consume hour is used each day that corresponds when uncertainty in demands is minimal) and then three options are used to assign the observed leakage signature to a group:

- mean of the sensitivities;
- mean of binarised sensitivities; and
- voting scheme (all days the leak is assigned to a group). The group with more assignations (votes) is the elected.

Results, even without uncertainty/noise, were not good using any of the three decision criteria. It was due to the changing boundary conditions (pressures and flows) that affected very

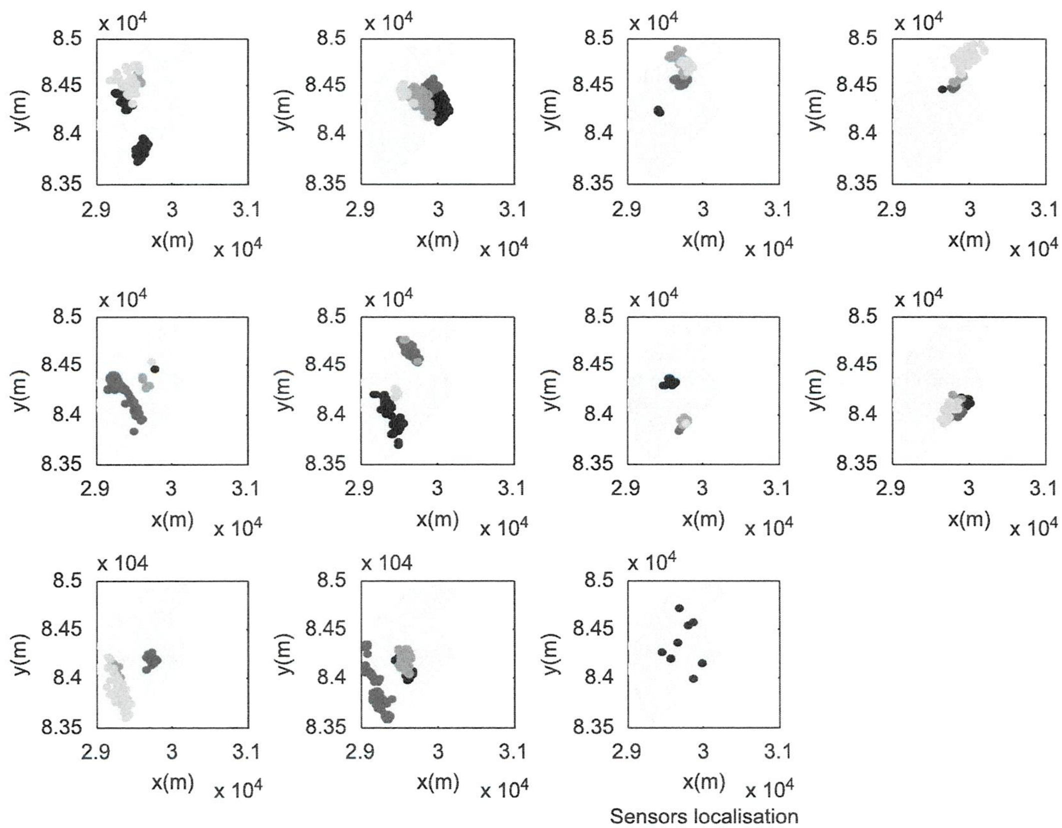


Fig. 7. Groups of nodes with the same leakage signature with 8 sensors and placement of sensors.

much the sensitivity matrix. It is necessary to generate the sensitivity matrix ad hoc for each day with proper boundary conditions that are known. Thus, for each new conditions all the simulations, normalisation and binarisation described in Section 4 are carried out. When a new signature matrix for each day is generated the two first approaches are useless because signature change for each iteration and mean values are meaningless. Thus the third one is tested. It provided perfect results without uncertainty, 100% localisation. It means that each day the group that was signalled suitable to have a leakage contained the node with leakage. These groups were all different each instant and signature matrix is adapted to boundary conditions thus only the voting method had sense. Thus, there are different probabilities of having a leak in a node. This appears in Table 1. It shows the number of nodes that have been signalled 0–15 times (each one for each day). The shadowed line cell corresponds to the one that contains the node that has real leakage. In this case, it corresponds always to the node number 15. It has been done for the 39 groups (one leakage for each) that appeared in sensor distribution (Fig. 7). In Fig. 8, the nodes are presented in grey scale representing the times that have been signalled to be suitable of containing a leakage. The one that contained it appears in the black area.

In order to test the methodology under uncertain parameters in the model, uncertainty in demands was introduced. Uncertainty was estimated using the monthly variation for a demand. It was of 18% of the total demand. Uncertainty was introduced as a coefficient multiplied to the demand of each node generated as a random number between 0.8 and 1.2. The global demand has been kept equal because it is a measured variable and affects greatly the sensitivity.

Table 1
Results using voting criteria adapting signature matrix.

Number of detections	Leak 1	2	3	4	5	6	7	8	9	10	11	12	13
0	529	316	503	316	986	798	782	884	1245	489	1253	1343	363
1	629	505	639	519	438	615	491	518	325	761	131	64	778
2	175	559	311	545	48	147	95	52	64	279	15	9	345
3	126	88	12	88	18	37	76	5	3	9	29	12	59
4	32	57	11	51	17	15	32	22	0	17	1	6	22
5	19	39	21	37	30	0	63	12	0	15	0	7	33
6	54	35	9	11	11	15	30	11	0	9	16	2	13
7	31	8	5	39	7	0	14	5	0	4	117	6	10
8	9	2	15	3	3	0	13	9	0	5	2	9	2
9	1	1	31	0	12	0	2	6	1	5	11	94	4
10	10	2	56	1	18	0	6	17	0	1	5	46	9
11	3	1	15	2	10	2	7	55	0	26	11	7	0
12	1	6	8	1	6	0	6	33	1	0	2	10	0
13	9	18	2	9	0	1	15	10	0	5	4	2	0
14	5	0	1	0	29	6	6	0	0	6	8	20	0
15	7	3	1	18	7	4	2	1	1	9	35	3	2

Number of detections	Leak 14	15	16	17	18	19	20	21	22	23	24	25	26
0	357	1311	1363	354	489	539	477	490	1396	1430	996	959	807
1	869	66	44	853	552	575	778	512	122	92	372	347	533
2	304	21	13	317	156	179	255	182	12	11	82	132	134
3	50	83	5	57	170	59	65	47	55	52	100	90	53
4	14	18	6	14	10	70	4	141	3	5	39	15	50
5	12	119	9	11	7	53	5	4	7	8	2	23	1
6	9	5	5	11	5	95	12	25	3	7	4	8	7
7	9	2	35	7	25	18	33	31	4	1	2	19	1
8	1	0	11	1	29	12	0	17	3	11	5	18	23
9	0	0	23	0	53	9	2	25	11	3	1	11	9
10	2	4	0	3	69	4	1	21	2	0	6	14	14
11	4	8	9	4	11	8	4	46	2	5	15	1	0
12	6	1	0	2	36	3	1	35	0	0	2	1	0
13	2	1	17	4	9	8	0	23	9	11	8	1	6
14	0	0	44	1	15	1	1	24	0	0	2	0	0
15	1	1	56	1	4	7	2	17	11	4	4	1	2

Number of detections	Leak 27	28	29	30	31	32	33	34	35	36	37	38	39
0	596	769	993	594	597	593	544	593	490	490	593	877	568
1	477	585	293	471	597	593	429	593	512	518	595	357	610
2	362	208	90	209	79	109	246	109	182	163	116	43	158
3	109	45	97	55	115	76	127	76	47	5	81	105	48
4	56	2	77	120	121	42	176	42	140	59	40	34	50
5	9	2	34	19	43	16	28	18	5	43	34	11	30
6	0	7	21	32	12	44	24	40	25	76	32	4	52
7	1	5	3	35	33	19	16	26	35	74	13	7	35
8	2	2	8	33	11	51	2	43	18	16	27	0	18
9	6	1	10	14	9	9	2	16	22	101	23	2	8
10	1	1	4	23	13	6	5	6	38	9	9	40	14
11	6	5	2	5	2	4	23	3	25	79	27	139	8
12	1	7	2	16	6	11	6	23	30	3	7	13	30
13	1	0	5	8	0	40	4	22	38	2	27	1	10
14	8	0	0	1	1	12	6	21	7	1	5	6	0
15	5	1	1	5	1	15	2	9	26	1	11	1	1

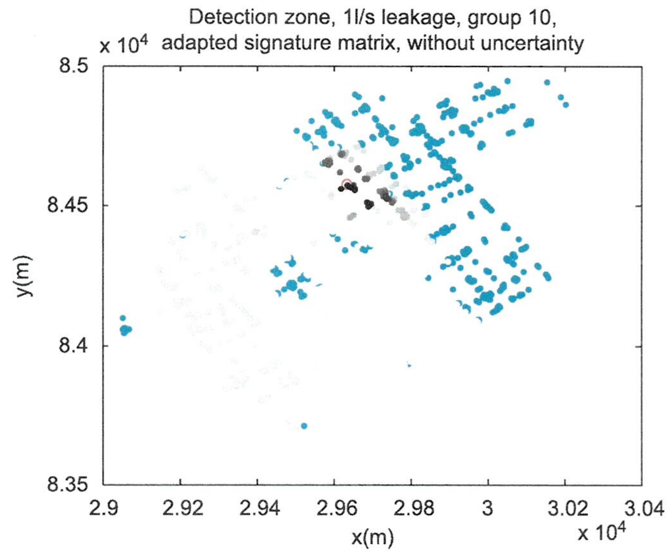


Fig. 8. Localisation of a leak in the correct zone with adapting signature matrix.

Results are presented in Table 2 and Fig. 9. In this case, the leaky node is not always exactly in the most signalled group and the dark grey in the figure does not correspond to 15 but to 9 days. It means that the nodes that more times have been signalled have been signalled thirteen times out of the fifteen. In Fig. 9, the grey scale is lighter than in Fig. 8 because there are less correct detections due to the uncertainty.

Increasing uncertainty interval, the proposed localisation methodology produces poorer results. For a 50% uncertainty, leaks were not well localised but they were localised in a neighbour zone.

The main handicap of the methodology is that in a highly looped network pressure drops due to a leak are not very significant. Therefore it demands high accuracy in transducers. Table 3 show the maximum and minimum pressure drop for leaks 0.5–10 l/s. In high demand hour, the difference is higher but the uncertainty in demand is higher too. Thus, the high cost of sensors may not guarantee good results because of uncertainties in demands.

6.2. *Real results*

Results from simulation test showed that high accuracy sensors are required. Such sensors exist but represent a major investment. Before such investment is authorised, real test with existing sensors were carried on. Few sensors with non-optimal distribution are available. Measurements have not been taken in best conditions (lower demand time). Nevertheless these results were interesting for the company in order to take further decisions and are presented in this section. A scenario based on a leakage forced in *Enamorats* DMA, in Barcelona network too, is used. This DMA have no qualitative difference with *Plaça del Diamant*. All the steps of methodology exposed so far are applied identically. Only the sensor distribution is not applied because the existing ones are used.

Enamorats DMA model contains 260 nodes and two water input points, where a flow metre and a pressure metre are installed. Input flows in the network and pressures at these points are fixed in the simulation model as boundary conditions. In addition to this information, this DMA has 3 installed pressure sensors, which have been used to apply leakage localisation

Table 2
Results using voting criteria with uncertainty 18%.

Number of detections	Leak 1	2	3	4	5	6	7	8	9	10	11	12	13
0	810	504	688	504	1027	1185	539	688	794	754	1223	1343	765
1	412	646	754	646	207	370	418	708	517	762	164	67	778
2	219	270	15	270	233	28	414	63	326	27	11	6	19
3	34	61	5	61	38	20	73	9	0	23	27	12	37
4	39	83	26	83	17	9	123	29	0	10	17	3	14
5	17	35	25	35	18	0	15	10	0	5	93	9	10
6	71	8	19	8	9	0	18	20	0	11	22	9	9
7	12	2	20	2	4	1	12	26	0	38	6	9	6
8	4	1	42	1	10	17	4	42	1	8	10	94	2
9	17	2	20	2	16	3	6	37	0	2	5	46	0
10	5	22	26	4	25	4	8	8	1	0	23	7	0
11	0	6	0	24	31	3	10	0	1	0	39	10	0
12	0	0	0	0	5	0	0	0	0	0	0	2	0
13	0	0	0	0	0	0	0	0	0	0	0	20	0
14	0	0	0	0	0	0	0	0	0	0	0	3	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0

Number of detections	Leak 14	15	16	17	18	19	20	21	22	23	24	25	26
0	645	1322	1363	623	637	528	1121	638	1272	1272	833	953	811
1	895	66	44	911	554	479	274	561	247	235	535	325	551
2	55	10	5	61	161	187	193	152	14	15	82	119	144
3	11	83	10	11	22	251	42	23	3	12	100	100	58
4	12	18	7	8	3	106	2	3	54	9	39	24	14
5	6	13	9	6	1	34	5	1	4	51	2	44	1
6	1	111	6	5	6	11	3	30	17	13	4	11	8
7	4	1	1	6	28	2	0	40	13	17	2	23	34
8	9	1	36	3	29	1	0	32	1	5	5	20	14
9	1	0	7	0	82	3	0	70	15	11	1	18	2
10	1	4	85	6	94	25	0	64	0	0	6	2	3
11	0	4	58	0	12	10	0	26	0	0	18	1	0
12	0	7	9	0	11	3	0	0	0	0	9	0	0
13	0	0	0	0	0	0	0	0	0	0	4	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0

Number of detections	Leak 27	28	29	30	31	32	33	34	35	36	37	38	39
0	596	596	597	594	597	595	526	593	638	638	595	1165	570
1	477	477	601	478	597	619	464	613	501	511	620	10	636
2	362	367	100	270	85	94	197	99	16	21	106	101	116
3	103	153	118	68	145	85	120	87	80	6	98	106	88
4	29	15	118	50	80	52	192	29	70	0	28	34	54
5	37	3	42	34	42	17	68	32	67	101	19	17	50
6	5	7	23	45	39	18	13	21	40	151	65	4	47
7	3	6	9	30	19	64	8	59	54	23	24	2	12
8	7	1	14	24	15	15	8	21	37	100	9	3	41
9	5	1	8	10	7	5	3	7	44	9	19	38	10
10	8	8	4	10	6	20	2	3	93	73	35	144	16
11	7	6	4	10	4	34	16	39	0	4	20	9	0
12	1	0	2	12	3	20	14	15	0	3	2	7	0
13	0	0	0	5	1	2	9	20	0	0	0	0	0
14	0	0	0	0	0	0	0	2	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0

methodology. The water company provided boundary conditions (pressure and flow) and pressure inside the DMA (three sensors) data with 10 minute time step. This information was for 5 days in the last day a leakage was forced. Table 4 shows information about this leakage.

The first step is to verify that the hydraulic model provided is correctly calibrated. A four days simulation without any leakage has been done considering pressure values in three internal pressure sensors. The result is the pressure evolution during each day in internal pressure sensors. Differences between the model and reality are important because of demand uncertainty. The worst consequence of these results is that pressure difference caused by a leak can be hidden by the differences due to misfitting of demand model and real demand. To solve this problem, model is corrected with the mean error during no leakage days. Real corrected pressure using these mean errors in each sensor is shown in Fig. 10, compared with the simulation ones.

Although a correction to real pressure has been applied, no difference in the period of leakage can be observed. Thus localisation methodology is applied to see if it is possible to show more

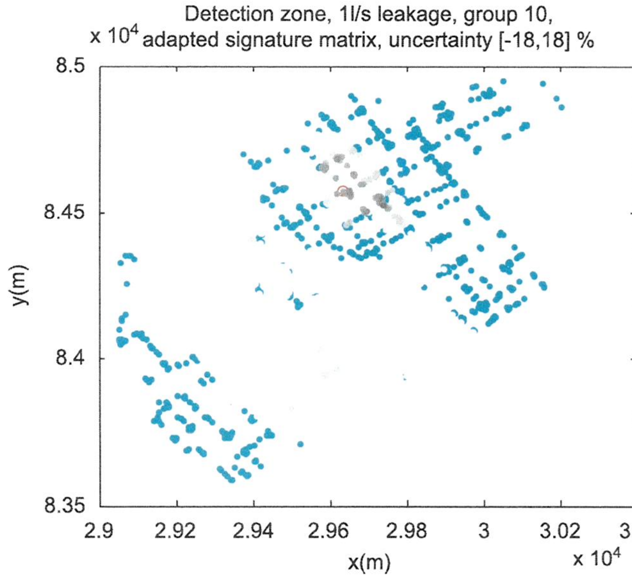


Fig. 9. Localisation of a leak in the correct zone with 18% uncertainty in the demand.

Table 3
Maximum and minimum pressure drop.

Leakage flow (l/s)	Minimum demand hour		Maximum demand hour	
	Minimal ΔP (m)	Maximal ΔP (m)	Minimal ΔP (m)	Maximal ΔP (m)
0.5	0.01	0.02	0.01	0.03
1	0.01	0.04	0.01	0.06
2	0.01	0.09	0.01	0.12
3	0.01	0.14	0.01	0.18
4	0.01	0.19	0.01	0.24
5	0.01	0.24	0.01	0.31
6	0.01	0.29	0.01	0.38
8	0.01	0.37	0.01	0.52
10	0.01	0.44	0.01	0.67

Table 4
Leakage information in *Enamorats* DMA.

Flow (m^3/h)	Flow (l/s)	Leak location	Start time	End time
18	5	Lepant/Aragó	10:20	10:35
14	3.9	Lepant/Aragó	10:37	10:52
9	2.5	Lepant/Aragó	10:53	11:08
6	1.7	Lepant/Aragó	11:10	11:25
16	4.4	Aragó 79	11:53	12:08

information not seen in previous figures. Leakage period duration is about one hour. For leakage period, five second step time data is given by the water company, but only pressures, not flows. If a ten minute step time data is used, in an hour period only 6 samples can be taken. To increase the number of samples a minute time step is proposed. To calibrate the model pressures at the input points are calculated by the mean of the last 30 s data (6 samples) and input flow is taken as a constant during 10 min.

To find discriminable zones obtained with installed sensors, a leak is moved for all 260 possible nodes using the model. For the leakage period two simulations are done: the first one without any leakage and the second one with a leakage moved for 260 nodes. Forced leakage flow is not constant, as it can be seen in Table 3, but only ten minutes data is given for each case. For this reason it is assumed that the leakage flow (5 l/s) is one of them for the whole

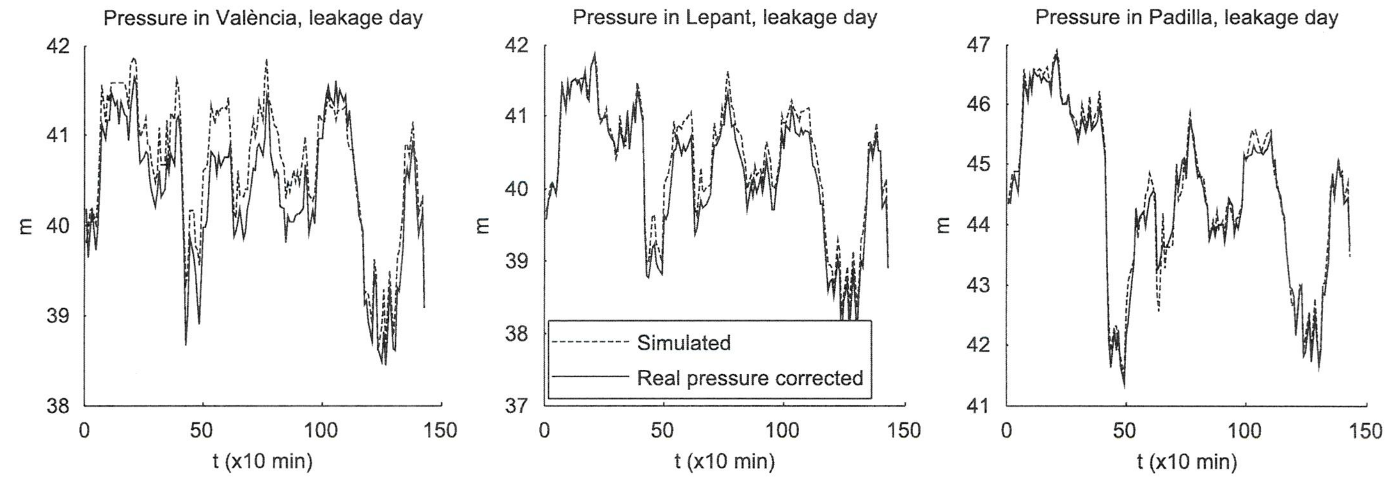
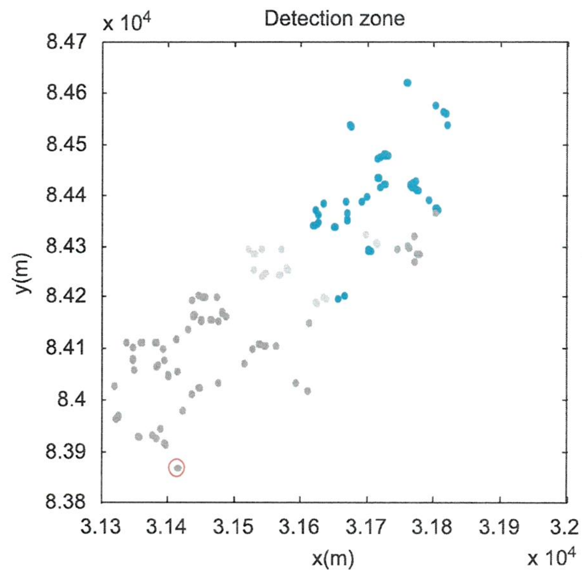


Fig. 10. Corrected real pressures compared with simulated ones.



Threshold = 0.4				
Signature			Number of nodes	Number of detections
València	Lepant	Padilla		
0	0	0	55	0
1	0	0	23	17
1	1	0	88	31
1	1	1	94	4
Total (max. = 64)				52

Fig. 11. Leakage localisation results with a threshold of 0.4 in a leakage period.

period. This assumption can be justified due the fact that in a real situation the leakage flow may be variable and unknown.

In the same way as in simulation tests, leakage methodology has been applied during more than one time step. In next figures some results are shown. The first case corresponds to the leakage period. As well as in simulation, signature matrix has been calculated depending on boundary conditions. Due to the little quantity of data, the test has been done during the whole leakage period, taking a pressure measurement every minute. Results for a 0.4 threshold are shown in Fig. 11. Although some leakages are not detectable (55 nodes zone), the real leakage is outside this zone. Sensors are not located optimally, so these undetectable leakages were expected. The number of discriminable zones is four, including the non-detectable one. Leakage zone corresponds to the third group, which contains 88 nodes.

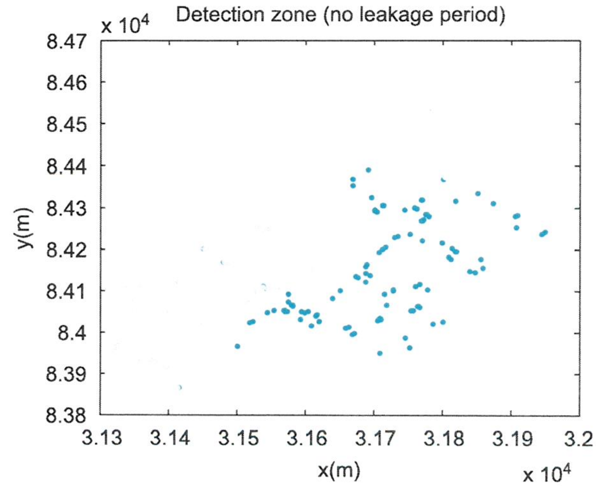
The leakage is given in the circled node: 31 of 64 detections signalled the correct leakage zone. After this test a non leakage period is chosen to apply the methodology. At night discrepancies between reality and the model are smaller than during the day; so it is the best time to do the test. Although an important zone is signalled as a possible leakage zone, the number of detections is only 9 on 42. These results are shown in Fig. 12.

7. Conclusions

A leakage localisation method based on the pressure measurements and sensitivity analysis of nodes in a network has been proposed. The leakage localisation methodology is founded in standard model-based fault diagnosis well established theory.

In order to maximise the isolability with a reasonable number of sensors, an optimal sensor placement methodology based on genetic algorithms is also proposed. The objective function in the minimisation process was the size of the maximum group discriminated. The confidence of the information provided by pressure sensors about leakage could be studied using the Fisher Information Matrix generated using the sensitivity matrix. This new approach is studied as a possible way to define the sensor placement avoiding the optimisation process.

To assess the validity of the proposed approach, it has been applied to a DMA of Barcelona network in real and simulated leak scenarios. Models and information were provided by the water company. For these sectors (DMA), the sensor placement and the leakage detection and localisation methodologies have been applied with successful results even in presence of demand uncertainty in simulation.



Threshold = 0.4				
Signature				
València	Lepant	Padilla	Number of nodes	Number of detections
0	0	0	29	0
0	0	1	94	0
1	0	0	62	4
1	1	0	75	5
			Total (max. = 42)	9

Fig. 12. Leakage localisation results with a threshold of 0.4 in a non leakage period.

In real test where sensors used where already installed results were poorer. Two main causes are suggested. First the non-optimal distribution of the sensors thus the methodology proposed in Section 4 is currently being applied in an on-going project in order to improve such results. On the other hand, the estimation of demands should be improved and an evaluation of the influence of the misfit of demand model on the methodology has been studied. First results have been published (Pérez et al., 2011).

An issue in the process is to recalculate the sensitivity matrix for each boundary condition using the simulation model because of the high dependence of it to global consumption. This approach is being currently developed using linear parameter varying (LPV) models that consider the consumption as a scheduling variable (Vento & Puig, 2009). Finally, a new approach is being studied that avoids the binarisation of the sensitivity matrix and it is based on correlation of model pressures with leakage and the measurements (Quevedo et al., 2011)

Acknowledgement

The authors wish to thank the support received by WATMAN ref. DPI2009-13744 of the Spanish Ministry of Education and PROFURED project funded by AGBAR-CETAQUA.

References

Bagajewicz, M. (2000). *Design and upgrade of process plant instrumentation*. Lancaster, PA: Technomic Publishers.

Bedjaoui, N., & Weyer, E. (2011). Algorithms for leak detection, estimation, isolation and localization in open water channels. *Control Engineering Practice*, *19*(6), 564–573.

Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2006). *Diagnosis and fault-tolerant control. 2nd ed.* Springer.

Brdys, M. A., & Ulanicki, B. (1994). *Operational control of water systems: Structures, algorithms and applications*. UK: Prentice Hall International.

Chen, J., & Patton, R. J. (1999). *Robust model-based fault diagnosis for dynamic systems*. Kluwer Academic Publishers.

Colombo, A. F., Lee, P., & Karney, B. W. (2009). A selective literature review of transient-based leak detection methods. *Journal of Hydro-environment Research*, *2*, 212–227.

Farley, M., & Trow, S. (2003). *Losses in water distribution networks*. UK: IWA Publishing.

Ferrante, M., & Brunone, B. (2003a). Pipe system diagnosis and leak detection by unsteady-state tests. *1 Harmonic Analysis, Advances in Water Resources*, *26*(1), 95–105.

Ferrante, M., & Brunone, B. (2003b). Pipe system diagnosis and leak detection by unsteady-state tests. *2 Wavelet Analysis, Advances in Water Resources*, *26*(1), 107–116.

Gertler, J. J. (1998). *Fault detection and diagnosis in engineering systems*. Marcel Dekker.

Lambert, A. (1994). Accounting for losses: the Bursa and background concept. *(BABE)IWEM Journal*, *8*(2), 205–214.

MacDonald G. (2005). *DMA design and implementation, a North American Context*. Leakage conference, IWA.

Misiunas, D., Lambert, M. F., Simpson, A. R., & Olsson, G. (2005). Burst detection and location in water distribution systems. *Water Science and Technology: Water Supply*, *Vol. 5*(3–4), 71–80.

Nejjari, F., Pérez, R., Escobet, T., & Traves, L. (2006). Fault diagnosability utilizing quasi-static and structural modelling. *Mathematical and computer modelling*, *Vol. 45*, 606–616.

Perez, R., de las Heras, S., Aguilar, J., Pascual, J., & Peralta, A. (2009a). District management areas characterisation in water network based on clustering. *Water Science Technology*, *9*(5), 591–600.

Pérez, R., Puig, V., Pascual, J., Peralta, A., Landeros, E., & Jordanas, Ll. (2009b). Pressure sensor distribution for leak detection in Barcelona water distribution network. *Water Science & Technology*, *9*(6), 715–721.

Pérez, R., Nejjari, F., Puig, V., Quevedo, J., Sanz, G., Cugueró, M., & Peralta, A. (2011). *Study of the isolability of leaks in a network depending on calibration of demands. Computing and control for the water industry 2011 "Urban water management—Challenges and opportunities"*. University of Exeter.

Puig, V., Quevedo, J., Escobet, T., Nejjari, F., & de las Heras, S. (2008). Passive robust fault detection of dynamic processes using interval models. *IEEE Transactions on Control Systems Technology*, *16*(5), 1083–1089.

Pudar, R. S., & Ligget, J. A. (1992). Leaks in pipe networks. *Journal of Hydraulic Engineering*, *118*(7), 1031–1046.

Quevedo, J., Puig, V., Cembrano, G., Blanch, J., Aguilar, J., Saporta, D., Benito, G., Hedo, M., & Molina, A. (2010). Validation and reconstruction of flow meter data in the Barcelona water distribution network. *Control Engineering Practice*, *18*(6), 640–651. doi:10.1016/j.conengprac.2010.03.003.

Quevedo, J., Cuguero, M. À., Perez, R., Nejjari, F., Puig, V., & Mirats, J. M. Leakage location in water distribution networks based on correlation measurement of pressure sensors. *WATERMATEX*, submitted for publication, San Sebastián.

Ragot, J., & Maquin, D. (2006). Fault measurement detection in an urban water supply network. *Journal of Process Control*, *16*(9), 887–902.

Sarrate, R., Puig, V., Escobet, T., Rosich, A. (2007). Optimal sensor placement for model-based fault detection and isolation. In *Proceedings of the 46th IEEE conference on decision and control*. New Orleans, USA.

Sezer, M. E., & Siljak, V. (1986). Nested epsilon-decomposition and clustering of complex systems. *Automatica*, *22*(3), 321–331.

Song, Z., Chen, Y., Sastry, C. R., & Tas, N. C. (2009). *Optimal observation for cyber-physical systems*. Hanover: Springer.

Vento, J., Puig, V. (2009). Leak detection and isolation in pressurized water networks using interval LPV models. In *Proceedings of the 7th workshop on advanced control and diagnosis (ACD'09)*. Zielona-Gora, Poland.

Verde, C., Visairo, N., & Gentil, S. (2007). *Two leaks isolation in a pipeline by transient response. Advances in water resources*, Vol. 30. Elsevier.