University of Lille
Doctoral School of Engineering Sciences
Civil and geo-Environmental Engineering Laboratory

Manuscript submitted for the
Doctoral degree in Civil Engineering from Lille University

Thesis:

Use of the Artificial Intelligence methods for the detection and localization of leaks in the water distribution networks

By

Neda Mashhadi

22 July 2020

Jury members:

Shahrour Isam, Professor, University of Lille — Director
EL Khatabi Jamal, Professor, University of Lille — Co-Director
Hani Azedine, Professor, University of Badji Mokhtar Annaba — Reporter
BenHachmi Mohammed Karim, Professor, University of hassanII Casablanca — Reporter
Ashoori Negin, Engineer of Water Resources at Bay, United States — Member
Attoue Nivine, Engineer of Research,University of Lille — Member
Hage Chehade Fadi, Professor University of Libanaise — Member
Zaoui Ali, Professor, University of Lille — Member

Université de Lille

Ecole Doctorale Sciences pour l'Ingenieur

Laboratoire Génie Civil et Géo-Environnement

Thèse:

Utilisation des méthodes d'Intelligence Artificielle pour la détection et la localisation des fuites dans les réseaux de distribution d'eau

Spécialité: Génie Civil

Neda Mashhadi

22 Juillet 2020

Membres du jury :

Shahrour Isam, Professeur, Université de Lille	Directeur
EL Khatabi Jamal, Professeur, Université de Lille	Co-Directeur
Hani Azedine, Professeur, Université Badji Mokhtar Annaba	Rapporteur
BenHachmi Mohammed Karim, Professeur,Université HassanII de Casablanca	Rapporteur
Ashoori Negin,Ingénieur des ressources en eau, Bay Area Water Supply,États-Unis	Membre
Attoue Nivine, Ingénieur de recherche, Université de Lille	Membre
Hage Chehade Fadi,Professeur Université Libanaise	Membre
Zaoui Ali, Professeur, Université de Lille	Membre

Acknowledgement

First and foremost, praises and thanks to the God, the Almighty, for His showers of blessings throughout my research work to complete the research successfully.

I would like to express my deep and sincere gratitude to my research supervisor, Prof. Isam SHAHROUR, for giving me the opportunity to do research and providing invaluable guidance throughout this research. It was a great privilege and honor to work and study under his guidance.

My Special thanks go to Dr. Ammar ALJER for his helps, guidances and all supports that he provided me during my research.

I am extremely grateful to my parents for their love, prayers, caring and sacrifices for educating and preparing me for my future.

I am extending my heartfelt thanks to a special person in my life for supporting me, lifting me up, comforting me, bringing joy to my soul and putting a smile on my face.

Finally, my thanks go to all the people who have supported me to complete the research work directly or indirectly.

« Ce qui est le plus important dans la vie, c'est de donner à quelqu'un un peu de bonheur »

« Alice Parizeau »

Abstract

This manuscript presents the results of research about the use of the Artificial Intelligence moths to detect and localize leaks in the water distribution networks. The manuscript is organized in three chapters:

The first chapter includes a literature review about the leak in the water distribution networks. First, it presents first the origin of the water leak and its dramatic economic, social and environmental impact. Then, it presents the conventional methods used for the detection of the water leak including hardware-based and software-based methods. This chapter highlights the opportunities offered by the smart monitoring and the Artificial Intelligence methods for the detection of leaks in the water networks. It also shows a need to explore on the same example the capacity of the main AI methods to detect and localize leaks in complex water networks.

The second chapter presents the water network of the scientific campus of Lille University, which is used as a support for this research. It argues the selection of this campus by its representativity of a small town, the complexity of the water network and the availability of data about the water network asset and consumption. The chapter also presents the construction of a Lab pilot to investigate of the possibility to localize water leaks from the ratios of the water supply flow rates.

The third chapter presents a synthesis of the use of Machine Learning methods in leak localization. It also presents the use of the software EPANET for the generation of data including the impact of 215 individual and double leaks on the variation of the water supply flow rates and the pressure in five zones of the campus. These data are then used to investigate the capacity of five Machine Learning methods to localize leaks in the water distribution system. The chapter suggests some recommendations for the use of ML methods in water leak localization.

Resume

Ce manuscrit présente les résultats de recherches sur l'utilisation de l'Intelligence Artificielle pour détecter et localiser les fuites dans les réseaux de distribution d'eau. Le manuscrit est organisé en trois chapitres :

Le premier chapitre présente une analyse bibliographique sur les fuites dans les réseaux de distribution d'eau. Il présente d'abord l'origine des fuites d'eau et leurs impacts économiques, sociaux et environnementaux. Il présente aussi les méthodes conventionnelles utilisées pour la détection des fuites d'eau. Ce chapitre met en évidence les opportunités offertes par la smart technologie et les méthodes d'intelligence artificielle pour la détection des fuites dans les réseaux d'eau. Il montre également la nécessité d'explorer sur les mêmes exemples la capacité de ces méthodes à détecter et localiser les fuites dans les réseaux d'eau complexes.

Le deuxième chapitre présente le réseau d'eau de la cité scientifique de l'Université de Lille, qui sert de support à cette recherche. Il explique le choix de ce campus par sa représentativité d'une petite ville, la complexité de son réseau d'eau et la disponibilité de données sur le réseau d'eau et les consommations d'eau. Le chapitre présente également la construction en laboratoire d'un pilote pour étudier la possibilité de localiser les fuites d'eau à partir des débits d'alimentation en eau.

Le troisième chapitre présente une synthèse de l'utilisation des méthodes d'apprentissage automatique (Machine Learning) dans la localisation des fuites d'eau. Il présente l'utilisation du logiciel EPANET pour la génération de données incluant l'impact de 215 fuites individuelles et doubles sur la variation des débits d'alimentation en eau et la pression dans cinq zones du campus. Ces données sont ensuite utilisées pour étudier la capacité de cinq méthodes d'apprentissage automatique à localiser les fuites dans le système de distribution d'eau. Le chapitre suggère quelques recommandations pour l'utilisation des méthodes d'apprentissage automatique dans la localisation des fuites d'eau.

Table of contents

General Introduction

This research work concerns the detection and localization of leaks in urban water distribution networks. This issue is of major concern in the management of the water distribution systems, because leaks in water distribution system cause important water losses with significant economic, social and environmental impacts. It could also cause serious damages to the surrounding soils and infrastructures as well as for water resources.

A report of the American Water Works Association Research Foundation (AWWARF) estimates that water utilities in the Unites States suffer from 250,000 to 300,000 main breaks per year, causing about \$3 billion of annual damages (Thornton et al. 2008). Significant potable water losses from leakage and distribution mains breakages are also reported throughout the Middle East. In some countries of this region the water losses are about 50% of the water supply (Renzetti and Dupont 2013). In addition, failure in water distribution system could lead to water resources and soil pollution.

Water leak in the urban the water distribution system results from several factors, in particular infrastructures ageing, lack of maintenance, bad management, insufficient monitoring and lack of innovation. Water distribution systems are traditionally designed with redundancy to improve the network reliability against mechanical and hydraulic failure. They include many inter-connected closed loops, which make these networks very complex. This high redundancy also makes very difficult leak detection and localization.

Several methods were proposed for the detection and localization of leaks in the water distribution systems. These methods are generally classified into two categories: hardware-based methods and software-based methods. The hardware-based methods include different technologies such as the acoustic methods, the fiber optic sensing methods, the vapor or liquid sensing tubes, the liquid sensing cables and the soil monitoring. These methods are helpful in leak detection, but their use requires the mobilization of important resources and time consuming.

The conventional soft-based leak detection methods include the mass balance, the real-time transient modeling, the negative pressure wave, the pressure point analysis, the statistical methods and the digital signal processing. These methods are based on data collected from the water network monitoring as well as from users' alerts. They are largely used by the water companies for the detection of water leaks. They could provide some general information about the localization of the water leaks, but the exact position of leaks is generally determined using the hardware-based methods.

With the large development of real-time monitoring of the water pressure and flow, both professional and scholars have been developing data-based methods for the detection and localization of leakage in the water distribution networks. The principle key adopted in this area is to compare real-time data with the normal range of the system behavior with a focus on flow and pressure data (Salam et al. 2014; Wu & Liu 2017; Chan et al. 2018; Zhou et al. 2019). The Artificial Intelligence-based methods have been used forboth the detection and localization of the water leakage in the water distribution systems,including supervised methods (Mounce et al., 2011; Zhanget al., 2016; Soldevila et al., 2016; Soldevila et al., 2017; Ciupke, 2018; Shravani et al., 2019; van der Walt et al., 2018), unsupervised methods (Wu et al., 2015; Zhang et al., 2016), and deep learning methods (Caputo&Pelagagge, 2003; Salam et al., 2014; Mounce et al., 2015; van der Walt et al., 2018; Yalçin et al., 2018; Zhou et al., 2019; Shravani et al., 2019; Rojek&Studzinski, 2019).

Despite the important research in the use of the Artificial Intelligence methods for leak detection and localization, the literature is still missing a comparison of the capacities of the different categories of these methods for the localization of the water leakage on the same water distribution system. This research work proposes to fill this gap by comparing the capacities of these methods to localize leakage in a complex water distribution system, which is based on the water network of the scientific campus of Lille University in France. The of this water network for this research is motivated by its complexity and representatively of a water network of a small town with about 25 000 users. It is also motivated by the high degradation of this network due to ageing, which cause important water leak. The availability of data about the network asset and the water consumption helped in this research.

This research includes also the construction and use of Lab pilot of the campus water network to investigate of the possibility to localize water leak from the ratios of the water supply flows. This pilot will be used to conduct a series of individual or double leaks tests to determine the impacts of these leaks on the water supply flows. The results of these tests will help in answering the research question concerning the possibility to localize leaks from only water supply flow rates.

This research focuses on the capacity of Machine Learning methods to localize leaks in complex water distribution networks. The research includes two stages, which concern data generation and investigation of the capacity of ML methods to localize leaks in water distribution system. Data were generated by numerical simulation of leaks in the water distribution network of the campus. The water network was subdivided in five hydraulic zones. For each leak, the EPANET software is used for the determination of the variation on the water supply flow and the water pressure. These data are then used for the training and testing of (i) three supervised Machine Learning methods (Logistic Regression, Decision tree and Random Forest), (ii) two unsupervised Machine Learning methods (Hierarchical classification method and a combination of the Principal Component Analysis and K-means methods), and (iii) The Artificial Neural Network (ANN). The results of this research will be turned into recommendations for the use of Machine Learning methods for leak localization in water distribution system.

This manuscript presents the research of the PhD. It is organized in three chapters:
The first chapter includes a literature review about the leak in the water distribution networks. First, it presents first the origin of the water leak and its dramatic economic, social and environmental impact. Then, it presents the conventional methods used for the detection of the water leak including hardware-based and software-based methods. This chapter highlights the opportunities offered by the smart monitoring and the Artificial Intelligence methods for the detection of leaks in the water networks. It also shows a need to explore on the same example the capacity of the main AI methods to detect and localize leaks in complex water networks.

The second chapter presents the water network of the scientific campus of Lille University, which is used as a support for this research. It argues the selection of this campus by its representatively of a small town, the complexity of the water network and the availability of data about the water network asset and consumption. The chapter also presents the construction of a Lab pilot to investigate of the possibility to localize water leaks from the ratios of the water supply flow rates

The third chapter presents a synthesis of the use of Machine Learning methods in leak localization. It also presents the use of the software EPANET for the generation of data including the impact of 215 individual and double leaks on the variation of the water supply flow rates and the pressure in five zones of the campus. These data are then used to investigate the capacity of five Machine Learning methods to localize leaks in the water distribution system. The chapter suggests some recommendation for the use of ML methods in water leak localization.

Use of Artificial Intelligence in water leak detection and localization

1 Chapter 1- State of the Art

This chapter presents the state of the art about the researches in the field of the detection of leaks in the water distribution systems. This issue is very important, because it has high environmental, economic and social impact. The chapter presents successively the water distribution systems, water losses, conventional methods used in the water leak detection, and finally the use of the Artificial Intelligence (AI) in the water leak detection. The chapter concludes by a comparison of the conventional water leak detection methods and a discussion of the interest and perspective of the use of AI method in water leak detection.

1.1 Water Distribution System

The water Distribution Systems (WDS) are traditionally built with topological and energy redundancy to improve the network reliability against mechanical and hydraulic failure. This objective is achieved by designing the water networks with many inter-connected closed loops and with pipe diameters that are larger than that strictly necessary to fulfill the design pressure at the network nodes (Di Nardo et al. 2017).

Figure 1.1 shows an overview of the WDS. The function of the WDS is not limited to the water distribution through the water network. It also includes a water treatment plant (WTP) to purify the water coming from source to WTP and deliver it to various elevated reservoirs. Various metering systems are used to measure the water consumption using conventional methods (manual recording) or advanced technology such the smart metering. It also ensures operations related to consumption recording dispatching, billing, statements, bills' collection, survey and maintenance

The question of finance is also very important in the management of the water distribution system. It requires establishing an economic model for the water distribution with a balance between income (revenues, national and international support, taxes…) and expenses. The latter should cover expenses related to the maintenance of the water system as well as its renovation and extension (Kashid and Pardeshi 2014).

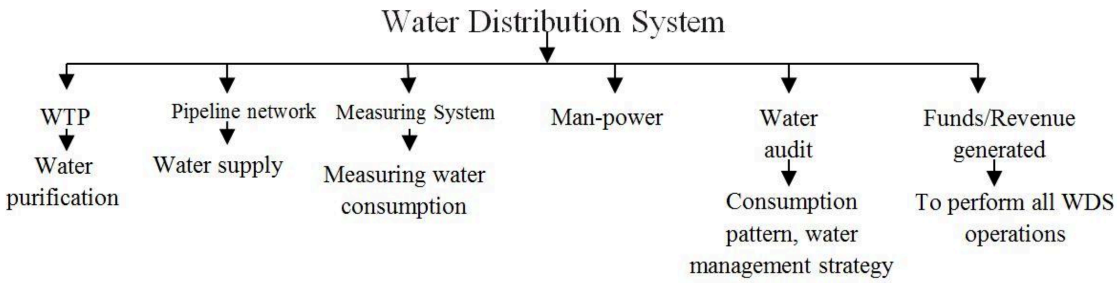


Figure 1.1: Water distribution system overview (Kashid and Pardeshi 2014)

1.2 Water losses

In recent years, water resources are subjected to an increasing stress due to the climate change, population increase, water infrastructures ageing and economic development. Currently, the water scarcity is recognized as a main threat particularly. Consequently, the water utilities should be highly efficient throughout the entire water supply process, to guarantee the water supply with sufficient quantities and good quality.

Water losses in the Water Distribution System (WDS) constitute a major challenge in the management of the WDS (Kanakoudis and Muhammetoglu 2014).Water losses include two components: real losses (RL) and apparent losses (AL). RLs refer to the annual volumes lost through all types of leaks and breaks on mains, service reservoirs (including overflows) and service connections, up to the point of customer metering. AL are the nonphysical losses that include customer meter under-registration, unauthorized use, meter reading and data handling errors (Mutikanga 2012).

A report of the American Water Works Association Research Foundation (AWWARF) estimates that water utilities in the Unites States suffer from 250,000 to 300,000 main breaks per year, causing about \$3 billion of annual damages (AADC, 2015). It is unknown how many small leaks occur, but annual leaks outnumber main breaks several times; likely resulting in 500,000 to 1,500,000 leaks per year (Thornton et al. 2008). Significant potable water losses from leakage and distribution mains breakages are reported throughout the Middle East. For example, water loss in Syria is about 45% of the water supply and it is about 50% in Jordan. In Bahrain, water loss by leak is about 20% of the water supply, while in Qatar it is about 30 %(Renzetti and Dupont 2013). In addition, failure in water distribution system could lead to water and soil pollution.

1.3 Leakage Management

1.3.1 Overview

Leakage management involves leak assessment, detection and control. According to Charalambous et al. (2014), water leak management should be conducted according to the integrated model in Figure 1.2. It includes 4 components:

- Continuity analysis and design optimization, including pipes selection, installation, maintenance and renewal.
- Leak and hydraulic zoning based on pressure management.
- Active leak detection.
- Asset management, in particular in repair management.

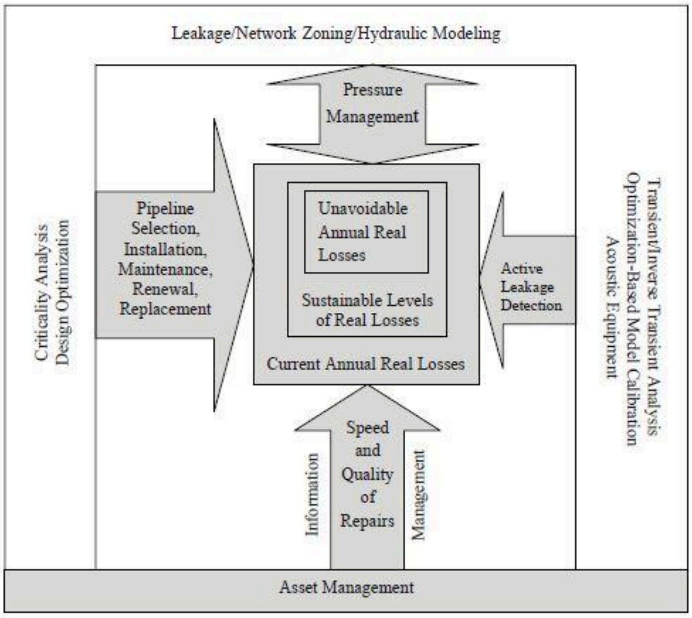


Figure 1.2 Integrated leakage management techniques (Charalambous et al. 2014)

1.3.2 Leak assessment methods

Leak assessment refers to tools and methods used to quantify the volume of leakage. The following three techniques have been widely used for leak assessment:

- Mass (or volume) balance methods (Water balance/audit).
- Network Hydraulic Modeling (NHM) simulations.
- Flow statistical analysis.

The mass (or volume) water loss methodology is based on the principle that the metered system input volume (V_{SIV}) must be equal to the sum of the water consumed (V_C), the change in water storage (D_V) and water leak (V_L):

$$V_{SIV} = V_C + \Delta V + VL \quad (Eq.\ 1.1)$$

This water balance methodology simplifies the complex task of tracking of the water supply in the WDSs.

- **Use of Network hydraulic modeling (NHM) in leak assessment**

WDSs are often very large and complex consisting of several kilometers of pipes of varying sizes and materials, storage reservoirs, pumps and various appurtenances. These systems are very difficult to understand and require large amounts of data for their analysis. Network Hydraulic Modeling (NHM) is largely used by engineers to understand and manage WDSs. NHM consists in using computers and mathematical models to predict the behavior of the WDS. It is used for operational investigations, planning tasks and network design purposes. Similar to mathematical modeling, the use of WDS models requires calibration on measured data. The calibration consists in adapting the model parameters to fit numerical simulations with measured data. Guidelines for WDS model calibration have been proposed (Savic et al. 2009).

Network simulation software provides the capability to mathematically replicate the non-linear dynamics of a WDS by solving the governing set of quasi-steady state hydraulic equations that include conservation of mass and energy within a loop. The software EPANET 2 is the most used in WDS modeling. The hydraulic solver of EPANET uses the gradient method with an open source code that allows extended modifications. For leakage management and control, the NHM can be applied for many purposes, including network zoning and rezoning, modeling leakage as pressure-dependent demand, pressure management planning, evaluating pipe renewal and replacement alternatives (Mutikanga 2012).

- **Use of statistical techniques for Leak assessment**

Statistical techniques have been used in leak assessment by various researchers. Palau et al. (2012) applied a multivariate statistical technique, called principle component analysis, for burst detection in urban WDSs. The advantage of this method is that it allows for a sensitive and quick analysis without use of computationally demanding mathematical algorithms. The technique can also be used to detect other abnormal flow conditions in the network such as illegal use of water.

1.4 Water Leak Detection Methods

1.4.1 Overview

Leaks can be classified as reported, unreported and background type (Adedeji et al. 2017).The former often appears on the ground and reported by the public or utility personnel. The unreported leak often does not appear on the ground, but in a similar manner to the reported one, they can be detected by leak detection method. For reported and an unreported leak, pressure reduction is usually noticed at the downstream of the pipes. The background type leak (such as flow through creeping joints) is not characterized by pressure drop and is difficult to detect. Since these leaks are difficult to detect, they constitute the largest threat to water utilities. Nonetheless, modeling a distribution water network should give a helpful breakthrough in detecting such kind of leakages (Adedeji et al., 2017).

There are two general ways for leak detection: hardware-based methods and software-based methods (Murvayet al. 2012). These two groups are named externally or internally based leak detection systems. Figure 1.3 illustrates the classification of leak detection methods.

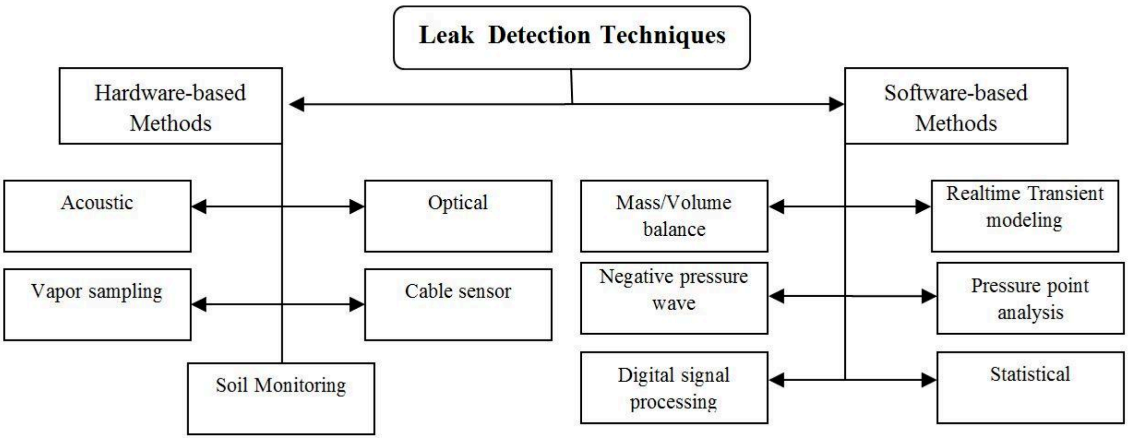


Figure 1.3 Classification of leak detection systems(Murvayet al. 2012)

1.4.2 Hardware Based Leak Detection

Hardware-based methods for leak detection and localization detect leaks from outside the pipeline by visual observation or by using appropriate equipment. These kinds of techniques are featured by a very good sensitivity to leaks and are precise in finding the leak location. However, they are expensive, and installation of their equipment is complex. As a result, their use is restricted to places with high potential of risk, likes near rivers or nature protection areas or in conditions which pipe is transferring a hazardous material (Murvayet al. 2012). Examples of this method are acoustic leak detection, fiber optical sensing cable, vapor sensing cable and liquid sensing cable-based systems.

1.4.2.1 Acoustic methods

The acoustic method is based on the principle that escaping liquid creates an acoustic signal as it passes through a perforation in the pipe (Wan et al. 2012). Acoustic sensors are used to track and detect acoustic signal caused by leakage (Figure 1.4).

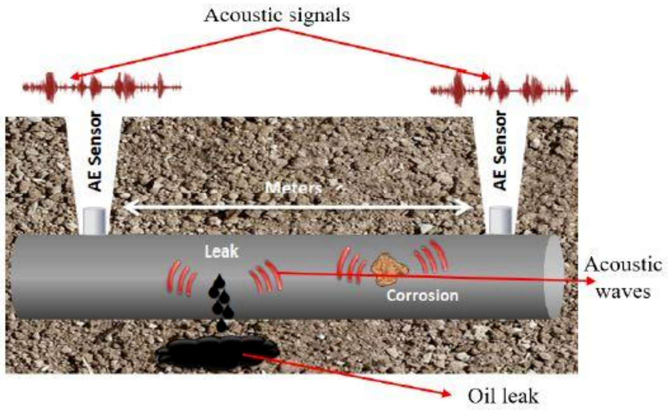


Figure 1.4: Acoustic emission method for leakage detection (Wan et al. 2012).

Since the received signal is higher in magnitude near the leak point; it gives an indication of the leak's location. However, signal characteristics as well as the variation in the environmental parameters surrounding the pipelines make it difficult to classify AE signals (Ahadi et al. 2010). Nevertheless, various signal analysis techniques have been applied to AE signals in order to obtain signal characteristics and locate leakage points, among which are correlation-based techniques (Gao et al. 2004). The application of these techniques depends on the pipe material. The correlation-based technique is effective in identifying leaks in metallic pipelines (Muggleton and Brennan 2004).

1.4.2.2 Fiber optic sensing methods

This method involves the installation of a fiber optic cable to measure the temperature over the pipeline. Conventionally, leakage introduces local temperature anomalies in the vicinity of the pipeline, by scanning the entire length of the fiber in short intervals, the temperature profile along the fiber is obtained and the leakage point can be detected. The cost of this system is quite high. Further developments lead to the use of distributed fiber optic technologies such as Raman distributed temperature sensor (RDTS), Brilloun optical time domain reflecto metry (BOTDR) and Fiber Bragg Gratting (FBG) for pipeline health monitoring (Rajeev et al. 2013).

1.4.2.3 Vapor or liquid sensing tubes

The vapor or liquid sensing tube-based leak detection method involves the installation of a tube along the entire length of the pipeline. If a leak happens, the content of pipe gets in touch of tube. The tube is full of air in atmospheric pressure. Once the leak occurs, the leaking substance penetrates into the tube. First of all, to assess the concentration distribution in the sensor tube, a column of air with constant speed is forced into the tube. There are gas sensors at the end of sensor tube. Every increase in gas concentration leads to a peak in gas concentration whose size is an indication of the size of the leak (Golmohamadi 2015).

1.4.2.4 Liquid sensing cables

Liquid sensing cables are placed near to a pipeline. Their main function is to track changes in transmitted energy pulses that have happened due to impedance differentials. Safe energy pulses are continually sent through the cable. As these energy pulses travel down the cable, reflections are returned to the monitoring unit and a "map" of the reflected energy from the cable is stored in memory. The presence of liquids on the sensor cable, in sufficient quantities to "wet" the cable, will alter its electrical properties. This alteration will cause a change of the reflection at that location. The alteration is then used to determine the location of a potential leak. For localization time delay between input pulse and reflected pulse are used, this method works well for multiple leak detection and localization for short pipelines (Golmohamadi 2015).

1.4.2.5 Soil monitoring

Soil monitoring technique exploits an inexpensive and non-hazardous gaseous tracer to be guided into the pipeline. This tracer is featured as a volatile gas, which escapes from the pipeline at the location of the leak. By analyzing the soil above the pipeline, the presence of leak and its location could be estimated. Producing low false alarms along with detect ability of very small leaks are the advantages of this method. Nevertheless, this is very expensive because the tracer should be injected into the pipe unceasingly in the detection process. It also is not feasible in uncovered pipelines (Golmohamadi 2015).

1.4.3 Software Based Systems

The software-based Systems are based on the monitoring of internal pipeline parameters (pressure, flow and temperature).Generally, the effectiveness of these methods depends on the uncertainties associated with the system's characteristics, operating conditions and collected data.

1.4.3.1 Mass-Volume balance

The mass balance method is based on the principle of conservation of mass, which states that a fluid enters the pipe section either remains in the pipe section or leaves the pipe section. In standard pipeline networks the flow entering and leaving the pipes can be metered. A leak can be identified if the difference between upstream and downstream flows changes by more than established threshold value (Murvayet al. 2012).

1.4.3.2 Real time transient modeling

This leak detection technique is based onpipe flow models which are constructed using equations of conservation of mass, conservation of momentum and conservation of energy. The difference between the measured value and the estimated value of the flow is used to determine the presence of leaks. For building this model flow, pressure and temperature measurements at both ends of the pipeline are necessary. Furthermore, to design a reliable system with minimum false alarm, the noise level should be continuously inspected to modify the model (Murvayet al. 2012).

1.4.3.3 Negative pressure wave

When leak occurs, the fluid pressure drops suddenly at the position of the leak and generates negative pressure wave, which propagates with a certain speed towards both upstream and downstream of the pipeline. Two pressure sensors are installed at the beginning station and the end station of the pipeline respectively. The negative pressure wave received by the two sensors can identify pipeline leak and furthermore locate the leak by calculating the time difference between the arrival times of the negative wave at each end (Hou and Zhang 2013).

1.4.3.4 Pressure point analysis

This method detects the occurrence of leaks by comparing the current pressure signal with a running statistical trend taken overa period of time along the pipeline by pressure monitoring and flow monitoring devices. The principle of this method is based on the fact of pressure drop as a result of leak occurrence. Using an appropriate statistical analysis of most recent pressure measurements, a sudden change in statistic properties of pressure measurement such as their mean value is detected.

1.4.3.5 Statistical methods

A statistical leak detection system uses statistical techniques to analyze the flow rate, pressure and temperature measurements of the water network pipes. This method is appropriate for complex pipe system as it can be monitored continuously for continual changes in the line and flow/pressure instruments. In addition, this technique could be used for leak localization. According to Murvay et al. (2012), the use of statistical analysis is easy and applicable into different pipeline systems. The main objective of this system is to minimize the rate of false alarms. It is also suitable for real-time application. It has been successfully tested in oil pipeline systems (Ghazali and Staszewski 2012).

1.4.3.6 Digital signal processing

Another method for leak detection is based on using digital signal processing techniques. The procedure of this method is that the response of the pipeline to a known input is measured over a period of time. Afterwards, this response is compared with the later measurements. Based on comparison of their signal’s features like frequency response or wavelet transform coefficients a leak alarm could be generated. Similar to statistical methods, this technique does not need a pipeline model (Murvayet al. 2012).

1.4.4 Leak detection systems

In-line systems specifically designed for large diameter pipes. They are able to discriminate between multiple leaks in a single length of pipeline. Pipelines can be inspected while under pressure and in service. Leaks are accurately located. There are two types of in-line systems: tethered and free swimming. In both cases, a sensor passes directly beside leaks, meaning that neither the pipe material nor the type of leak is relevant.

1.4.4.1 Tethered systems

Tethered systems operate by deploying a hydrophone into the pipeline to be inspected. The hydrophone is connected to a signal processing and display unit via an umbilical cable. The sensor travels inside the pipe pulled by the water flow. As the sensor passes any leak on the pipeline, it will detect the noise generated by the leak. The position of the sensor can be determined using a locating system mounted in the sensor head (Hamilton and Charalambous 2017).

1.4.4.2 Free swimming systems

Free-swimming data acquisition systems are used for leak detection. These systems are inserted into the pipeline and pushed by the water flow. These systems include following components:

- An internal camera, acoustic sensor, tracking sensors, acoustic transponder, data processor, memory device and batteries.
- Above-ground tracking devices (which are used to track the progress of the sensor through the pipe).
- Insertion equipment/launch tube
- Retrieval equipment

At the end of the inspection, a net or similar capture device is used to catch and extract the system from the pipeline, or they are discharged into an open catchment area such as a reservoir and recovered. During inspection, the system has the capability of capturing CCTV footage and acoustic noise. The maximum length of pipeline that can be surveyed is determined by the flow rate in the line. For instance, with a flow rate of 1 m/s and a maximum operating life of 12 hours, the system can survey 43 km from a single insertion point. The system can traverse around tight bends and through inline valves (Hamilton and Charalambous 2017).

Figure 1.6: Free swimming device (Kumar et al. 2017)

1.5 Use of Artificial Intelligence (AI) in leak detection

1.5.1 Overview

The Artificial Intelligence (AI) approach is a well-known data driven model. It mimics the human perception, learning and reasoning to solve complex problems through data description and analysis process. Many AI techniques exist in the literature: case-based reasoning, rule-based systems, artificial neural networks, genetic algorithms, cellular automata, fuzzy models, multi-agent systems, swarm intelligence, reinforcement learning, hybrid systems, Bayesian networks and data mining. The main objective of these techniques is the extraction of significant information from data for the semantic analysis and interpretation purposes (Lidia 2017).

Recent developments in data-driven models helped to solve various problems in the water management domain. The most reliable AI models used in water resource estimation are the Artificial Neural Network,(ANN),the Artificial Network-based Fuzzy Interface System (ANFIS), Genetic Algorithm (GA) and Support Vector Machine (SVM).

1.5.2 Use of AI in water management

The Artificial Neural Network (ANN) was one of the first AI methods used in water resource studies providing better performance in nonlinear modeling and more accurate forecasting. It is considered as one of the most efficient and popular methods in several hydrological applications (Afan et al. 2014). This technique is based on the ability of the human brain to predict patterns based on learning and recalling processes. The artificial neuron constitutes the processing unit. Each single neuron is connected to other neurons of a previous layer through adaptable synaptic weights. The network is formed by an input layer, an output layer and hidden layers.AI applications have been used for improving the classification and forecasting the flood at ungauged basins in data-driven models (Seckin et al. 2013). This application includes forecasting and modeling ground water level (Mohanty et al. 2015).

The adaptive neuro-fuzzy inference system (ANFIS) is a combination of an adaptive neural network (AN) and a fuzzy inference system (FIS). Fuzzy systems are used to deal with incomplete data through the application of fuzzy sets in function approximation, classification/clustering and control and prediction. Fuzzy models can describe vague statements as in natural language since it takes any value between 0 and 1. This technique was used for predicting ground water level fluctuation (Mirzavand et al. 2015), in water quality assessment (Soltani et al. 2010), water resource estimation(El-Shafie et al. 2011; Talei et al. 2010; Valizadeh and El-Shafie 2013) and water evaporation (Cobaner 2011).

The genetic algorithm (GA) is an evolutionary algorithm based on biological evolution inspired by Darwinian theories of natural selection and survival of the fittest. This technique is based on a search method mimicking natural selection. The algorithm repeatedly modifies a population of individual solutions until it satisfactorily solves the problem. It was used for optimization and forecasting problems.

The support vector machine (SVM) is based on the identifying of a hyper-plane that separates two classes in classification. It can be used for classification, regression and other tasks. This technique was used for water quality application (Tinelli and Juran 2019), regulation water level and discharge capacity (Shiri et al. 2019) and forecasting ground water level (Mukherjee and Ramachandran 2018).

1.5.3 Use of AI for leak detection

Artificial Intelligence was also used for leak detection in complex pipe system. Fuzzy logic was used by Sanz et al. (2012) to detect leak in WDS of Barcelona. Leak localization was determined through installing pressure sensors for pilot test and analysis of simulated data. Moczulski et al. (2016) also used the AI methods for leak detection in the water network of Rybnik city in Poland. Both Pressure and flow rate data were used for burst detection. Sousa et al. (2015) used the genetic algorithm method with pressure data to detect leakage. Satisfactory results were obtained at the macro scale.

Yalçin et al. (2018) applied a learning algorithm based on ANFIS method to detect water leak. An open-air experimental setup was installed. The method is proposed for industrial processes. Mounce et al. (2011) developed a hybrid ANN and FIS (fuzzy interference system) to detect leaks in a district meter area (DMA). They modified and extended the AI system to an automated online application, which executes self-learning and updating. It was shown that the proposed method constituted an effective tool for online leak detection.

1.6 Conclusion

This chapter presented a review of researches concerning leak in the water distribution system. This issue is of major concern for both public authorities and managers of water distribution system. Leak in the water distribution system cause water losses with significant economic and environmental impact. It could also cause damage to surrounding soils and infrastructures. In some cities water leak could account for 50% of the water supply. Water leak is due to several factors, in particular infrastructures ageing, lack of maintenance, bad management and insufficient monitoring.

Several methods were proposed for the detection of water leaks. These methods could be classified into two categories: hardware-based methods and software-basedmethods.The hardware-based methods include different technologies such as the acoustic methods, the fiber optic sensing methods, the vapor or liquid sensing tubes, the liquid sensing cables and the soil monitoring.

The soft-based methods include the mass balance, the real-time transient modeling, the negative pressure wave, the pressure point analysis, the statistical methods, the digital signal processing and the Artificial Intelligence methods.

The increase development in smart monitoring provides large capacities to collect pressure and flow data in large distribution water networks. These data are precious for the use of Artificial Intelligence methods for water leak detection and localization. The literature review showed an important concern in the use of these methods. However, the majority of the applications of the Intelligent Artificial methods remain at the research stage. The literature review revealed also a lack of a comprehensive use of these methods. The following chapter will present a comparison of the performance of different methods of the AI on the localization of leaks through an application on the scientific campus of Lille University, which stands for a small town with around 25 000 users. The aim of this application is to suggest good practices for the use of AI methods for leak detection and localization.

Chapter 2 will present the water distribution system of the scientific campus, while the chapter 3 will present an investigation of the use of AI for leak localization.

References:

Adedeji, K., Y. Hamam, B. Abe, and A. M. Abu-Mahfouz. 2017. "Leakage Detection Algorithm Integrating Water Distribution Networks Hydraulic Model." in SimHydro: Choosing the right model in applied hydraulics.

Afan, Haitham Abdulmohsin, Ahmed El-Shafie, Zaher Mundher Yaseen, Mohammed Majeed Hameed, Wan Hanna Melini Wan Mohtar, and Aini Hussain. 2014. "ANN Based Sediment Prediction Model Utilizing Different Input Scenarios." Water Resources Management. DOI: 10.1007/s11269-014-0870-1

Ahadi, Majid and Mehrdad Sharif Bakhtiar. 2010. "Leak Detection in Water-Filled Plastic Pipes through the Application of Tuned Wavelet Transforms to Acoustic Emission Signals." Applied Acoustics. DOI: 10.1016/j.apacoust.2010.02.006

Charalambous, B., D. Foufeas, and N. Petroulias. 2014. "Leak Detection and Water Loss Management." Water Utility Journal.

Cobaner, Murat. 2011. "Evapotranspiration Estimation by Two Different Neuro-Fuzzy Inference Systems." Journal of Hydrology. DOI: 10.1016/j.jhydrol.2010.12.030

Di Nardo, A., M. Di Natale, C. Giudicianni, D. Musmarra, J. M. Rodrigue. Varela, G. F. Santonastaso, A. Simone, and V. Tzatchkov (2017). "Redundancy Features of Water Distribution Systems." in Procedia Engineering. DOI: 10.1016/j.proeng.2017.03.244

El-Shafie, A., O. Jaafer, and A. Seyed. 2011. "Adaptive Neuro-Fuzzy Inference System Based Model for Rainfall Forecasting in Klang River, Malaysia." International Journal of Physical Sciences.

Gao, Y., M. J. Brennan, P. F. Joseph, J. M. Muggleton, and O. Hunaidi. 2004. "A Model of the Correlation Function of Leak Noise in Buried Plastic Pipes." Journal of Sound and Vibration .DOI: 10.1016/j.jsv.2003.08.045

Golmohamadi, Marcia. 2015. "Pipeline Leak Detection." MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY. Electronic OCLC # 913485554

Hamilton, Stuart and Bambos Charalambous. 2017. Leak Detection: Technology and Implementation. DOI: 10.26530/oapen_578133

Hou, Chun Xia and Er Hua Zhang. 2013. "Pipeline Leak Detection Based on Double Sensor Negative Pressure Wave." in Applied Mechanics and Materials. DOI: 10.4028/www.scientific.net/AMM.313-314.1225

Kanakoudis, Vasilis and Habib Muhammetoglu. 2014. "Urban Water Pipe Networks Management towards Non-Revenue Water Reduction: Two Case Studies from Greece and Turkey." Clean - Soil, Air, Water. DOI: 10.1002/clen.201300138

Kashid, Santosh Gautam and Sanjay A. Pardeshi. 2014. "A Survey of Water Distribution System and New Approach to Intelligent Water Distribution System." in 1st International Conference on Networks and Soft Computing, ICNSC 2014 - Proceedings.

Kumar, Dileep, Dezhan Tu, Naifu Zhu, Reehan Ali Shah, Dibo Hou, and Hongjian Zhang. 2017. "The Free-Swimming Device Leakage Detection in Plastic Water-Filled Pipes through Tuning the Wavelet Transform to the Underwater Acoustic Signals." Water (Switzerland). DOI: 10.3390/w9100731

Lidia Ogiela. 2017. Cognitive Information Systems in Management Sciences 1st Edition. Academic Press. ISBN:0128038039 9780128038031

Mirzavand, Mohammad, Benyamin Khoshnevisan, ShahaboddinShamshirband, Ozgur Kisi, Rodina Ahmad, and ShatirahAkib. 2015. “Evaluating Groundwater Level Fluctuation by Support Vector Regression and Neuro-Fuzzy Methods: A Comparative Study.” Natural Hazards. DOI: 10.1007/s11069-015-1602-4

Mohanty, S., Madan K. Jha, S. K. Raul, R. K. Panda, and K. P. Sudheer. 2015. “Using Artificial Neural Network Approach for Simultaneous Forecasting of Weekly Groundwater Levels at Multiple Sites.” Water Resources Management. DOI: 10.1007/s11269-015-1132-6

Muggleton, J. M. and M. J. Brennan. 2004. “Leak Noise Propagation and Attenuation in Submerged Plastic Water Pipes.” Journal of Sound and Vibration. DOI: 10.1016/j.jsv.2003.10.052.

Mukherjee, Amritendu and Parthasarathy Ramachandran. 2018. “Prediction of GWL with the Help of GRACE TWS for Unevenly Spaced Time Series Data in India : Analysis of Comparative Performances of SVR, ANN and LRM.” Journal of Hydrology. DOI: 10.1016/j.jhydrol.2018.02.005

Murvay, Pal Stefan and IoanSilea. 2012. “A Survey on Gas Leak Detection and Localization Techniques.” Journal of Loss Prevention in the Process Industries. DOI: 10.1016/j.jlp.2012.05.010

Mutikanga, harrison 2012. “WATER LOSS MANAGEMENT TOOLS AND METHODS FOR DEVELOPING COUNTRIES.” Delft- Netherland.ISBN: 9780415636483

Rajeev, Pathmanathan, Jayantha Kodikara, Wing Kong Chiu, and Thomas Kuen. 2013. “Distributed Optical Fibre Sensors and Their Applications in Pipeline Monitoring.” Key Engineering Materials. DOI: 10.4028/www.scientific.net/KEM.558.424

Renzetti, Steven and Diane Dupont. 2013. “Buried Treasure: The Economics of Leak Detection and Water Loss Prevention in Ontario.” Environmental Sustainability Research Centre (ESRC) Working Paper Series. http://hdl.handle.net/10464/4279

Sanz, Gerard, Ramon Perez, and Antoni Escobet. 2012. “Leakage Localization in Water Networks Using Fuzzy Logic.” in 2012 20th Mediterranean Conference on Control and Automation, MED 2012 - Conference Proceedings. DOI: 10.1109/MED.2012.6265711

Savic, Dragan A., Zoran S. Kapelan, and Philip M. R. Jonkergouw. 2009. “Quo Vadis Water Distribution Model Calibration?” Urban Water Journal. DOI: 10.1080/15730620802613380

Seckin, Neslihan, Murat Cobaner, Recep Yurtal, and TefarukHaktanir. 2013. “Comparison of Artificial Neural Network Methods with L-Moments for Estimating Flood Flow at Ungauged Sites: The Case of East Mediterranean River Basin, Turkey.” Water Resources Management. DOI: 10.1007/s11269-013-0278-3

Shiri, Jalal, Mohammad ZounematKermani, Ozgur Kisi, and Sahar Mohsenzadeh Karimi. 2019. “Comprehensive Assessment of 12 Soft Computing Approaches for Modeling Reference Evapotranspiration in Humid Locations.” Meteorological Applications. DOI: doi.org/10.1002/met.1841

Soltani, Foad, Reza Kerachian, and Ehsan Shirangi. 2010. “Developing Operating Rules for Reservoirs Considering the Water Quality Issues: Application of ANFIS-Based Surrogate Models.” Expert Systems with Applications. DOI: 10.1016/j.eswa.2010.03.057

Sousa, Joaquim, Luísa Ribeiro, João Muranho, and Alfeu Sá Marques. 2015. "Locating Leaks in Water Distribution Networks with Simulated Annealing and Graph Theory." in Procedia Engineering. DOI: 10.1016/j.proeng.2015.08.854

Thornton, J., R. Sturm, and George Kunkel. 2008. Water Loss Control. Second. McGraw Hill. ISBN: 9780071499187

Tinelli, Silvia and IlanJuran. 2019. "Artificial Intelligence-Based Monitoring System of Water Quality Parameters for Early Detection of Non-Specific Bio-Contamination in Water Distribution Systems." Water Science and Technology: Water Supply. DOI: 10.2166/ws.2019.057

Valizadeh, Nariman and Ahmed El-Shafie. 2013. "Forecasting the Level of Reservoirs Using Multiple Input Fuzzification in ANFIS." Water Resources Management. DOI: 10.1007/s11269-013-0349-5

Wan, Jiangwen, Yang Yu, Yinfeng Wu, Renjian Feng, and Ning Yu. 2012. "Hierarchical Leak Detection and Localization Method in Natural Gas Pipeline Monitoring Sensor Networks." Sensors. DOI: 10.3390/s120100189

Yalçin, Baris Can, Cihan Demir, Murat Gökçe, and Ahmet Koyun. 2018. "Water Leakage Detection for Complex Pipe Systems Using Hybrid Learning Algorithm Based on ANFIS Method." Journal of Computing and Information Science in Engineering. DOI: 10.1115/1.4040130

2 Chapter 2: Presentation of the Water Distribution Network of the Scientific Campus of Lille University

2.1 Introduction

This chapter presents the water distribution network of the scientific campus of Lille University, which is used as a support for this research. This network was selected for this research because of its important dimension, complexity and the availability of data about both the network asset and water consumption. This chapter presents successively, the scientific campus, the water distribution network and the investigation of the water leak using a lab pilot of the campus water distribution network

2.2 Presentation of the scientific campus

The scientific campus of Lille University is located in the city of Villeneuve d'Ascq, near the city of Lille in the North of France. The campus was constructed between 1964 and 1966. It stands for a small town of about 25,000 users. It includes 145 buildings with a total construction area of about 325,000 m^2 (Figure 2.1). Buildings are used for research, teaching, administration, students' residence and sports. The campus is disserved by 100 km of urban networks including drinking water, sewage, electrical grid, public lighting and district heating (Figure 2.2).

The campus was used by the SunRise team of the Civil Engineering Laboratory as a pilot of the smart city (Shahrour et al. 2017). It was also used within the European project SmartWater4Europe as ademonstrator of the Smart water system (Farah 2016).

The water distribution system was also used as a support for the PhD of Elias Farah (Farah 2016)"Detection of water leak using a smart monitoring system combined with the DMA and minimum night flow methods" and the PhD of Christine Saab (2018) "Real-time control of the water quality and detection of accidental water pollution".

Figure 2.1. Scientific Campus of Lille University

Figure 2.2. Urban networks of the scientific campus

2.3 Water distribution system of the scientific campus

2.3.1 Water network architecture

The water distribution network of the scientific campus is around 55 years old. It is composed of 15 km of grey cast iron pipes. Figure 2.3 shows the details of this network. It can be observed that this network is highly meshed, which makes this network very complex concerning the hydraulic characteristics.

Figure 2.4 shows the EPANET model of the water distribution system. It includes 260 pipes, 244 junctions and 3 tanks.

The water network is connected to the network of the city of Villeneuve d'Ascq at the following locations (Figure 2.5):

- CitéScientifique in the North of the Campus.
- 4 Cantons in the South of the Campus.
- ECL in the South-West of the Campus.
- Bachelard in the West of the Campus.
- M5 in the West of the Campus.

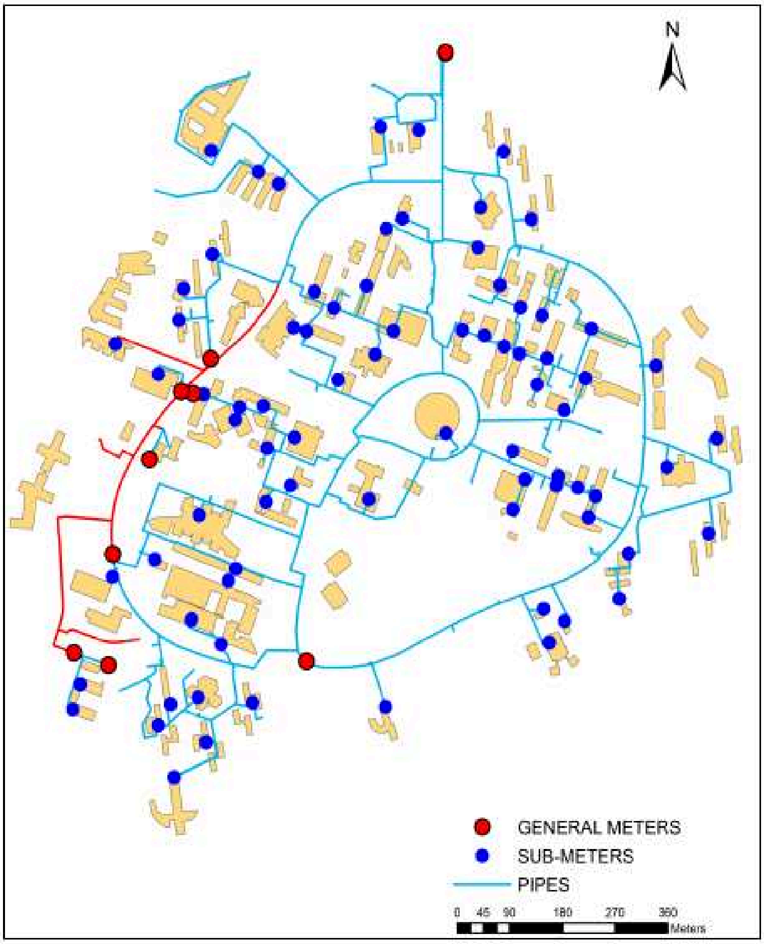


Figure 2.3. Water network of the Scientific Campus (Farah, 2016)

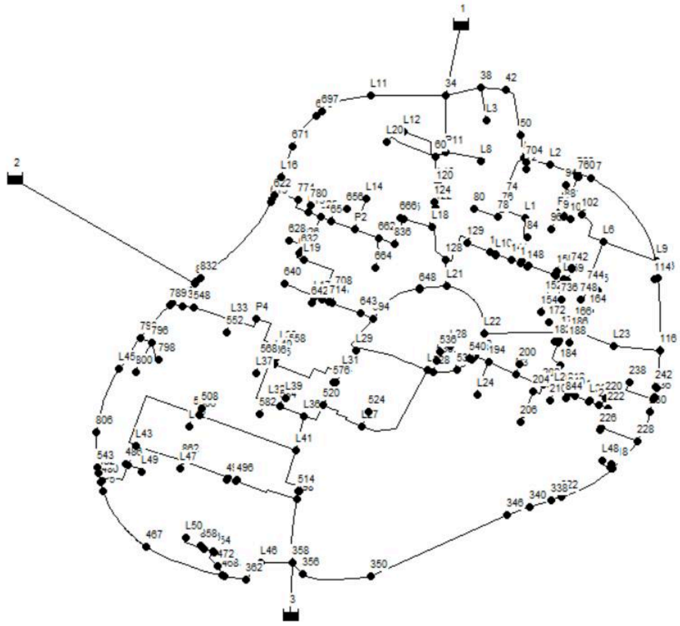


Figure 2.4. EPANET model of the water distribution system of the campus (260 Pipes, 244 Junctions)

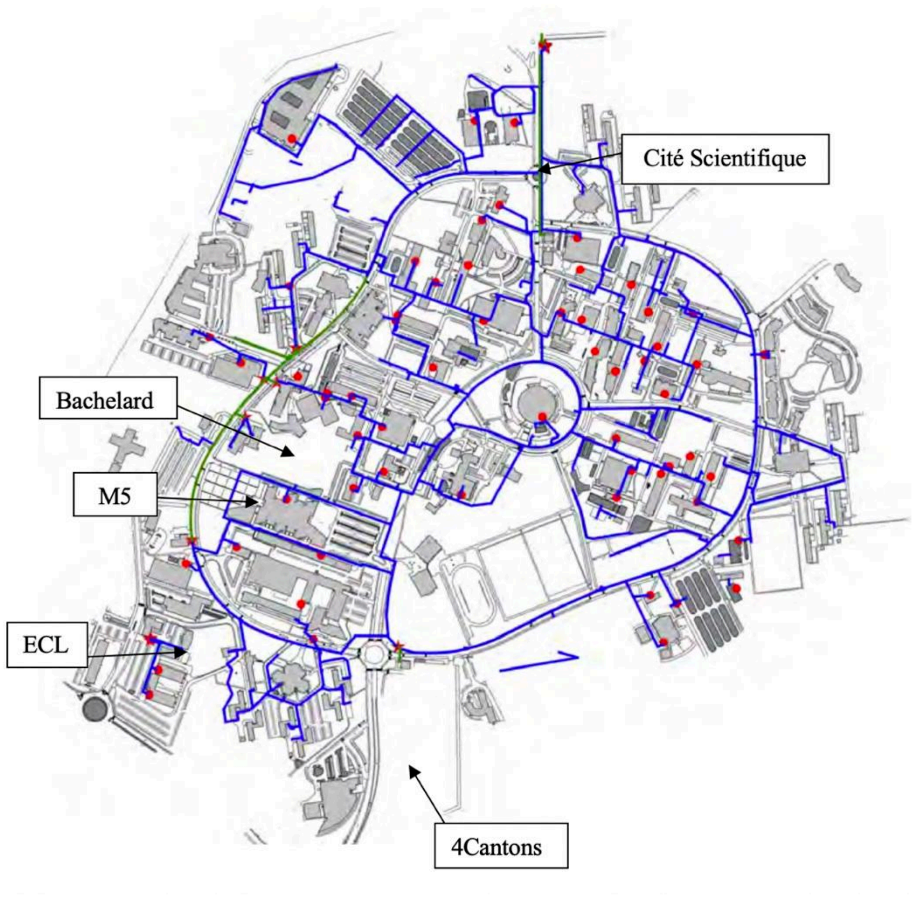


Figure 2.5 Water supply network of the Scientific Campus (Farah, 2016)

2.3.2 Water pipes and valves

The pipes of the water network are composed of three materials: Cast iron (90.2%), ductile iron (7.9%) and PVC (1.9%). They are buried under the pavements or under the vegetal lands. A part of the pipes is located in technical galleries that connect some buildings in the campus.

The diameter of the pipes varies between 20 to 300 mm (Figure 2.6). 58% of the pipes have a diameter of 100mm, 23% have a diameter superior to100mm, and 19% of less than 100mm.

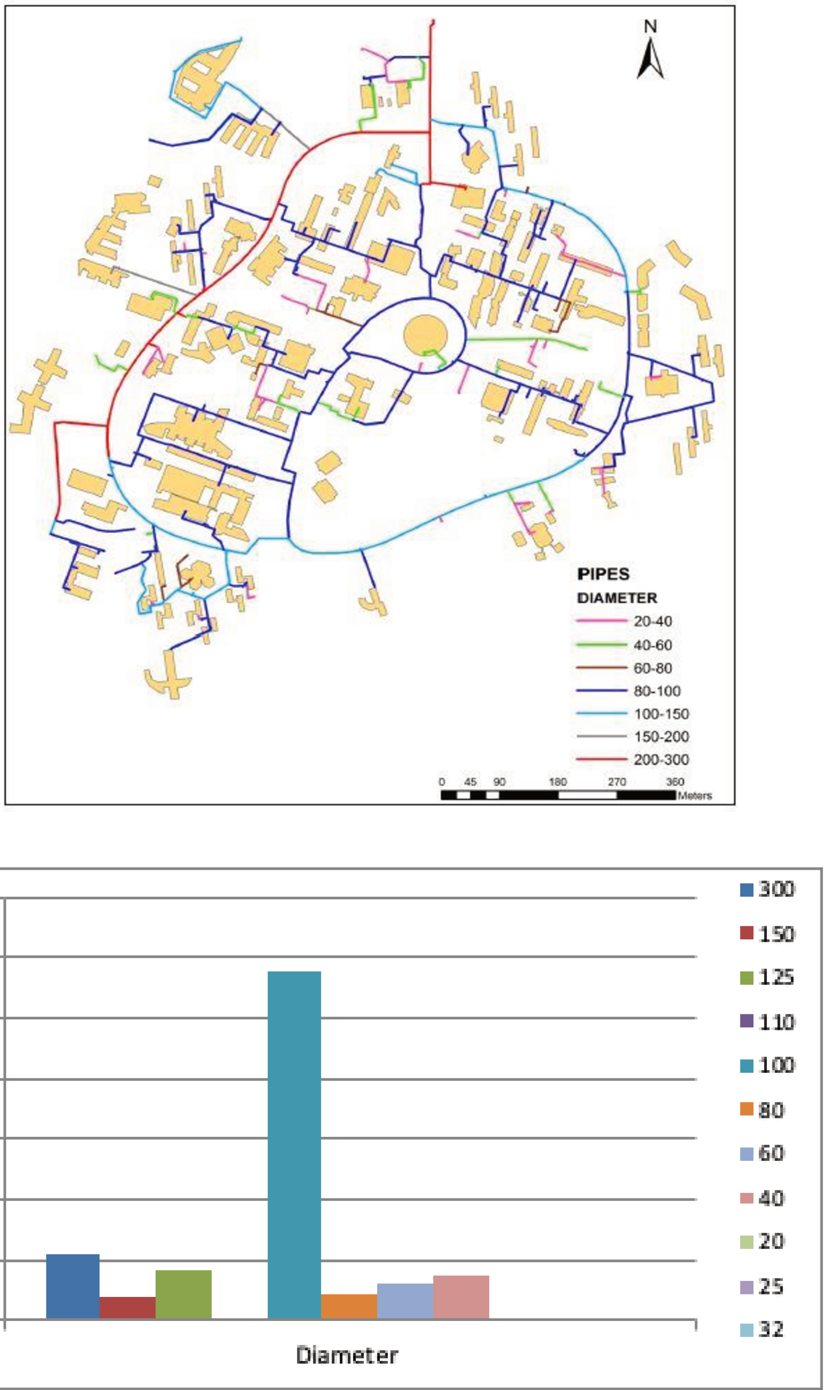


Figure 2.6. Pipes distribution according to the diameter (Farah, 2016)

The water distribution network includes about 250 valves (Figure 2.7). Most of them are old, corroded and suffer from leaks. Valves ensure the following tasks: (i) flow and/or pressure regulation, and (ii) isolation of parts of the network for maintenance or other operations.

In addition to the isolation valves, the network includes a set of manual and automatic air valves that help to release air from pipelines, which prevents reduction of the conveying capacity.

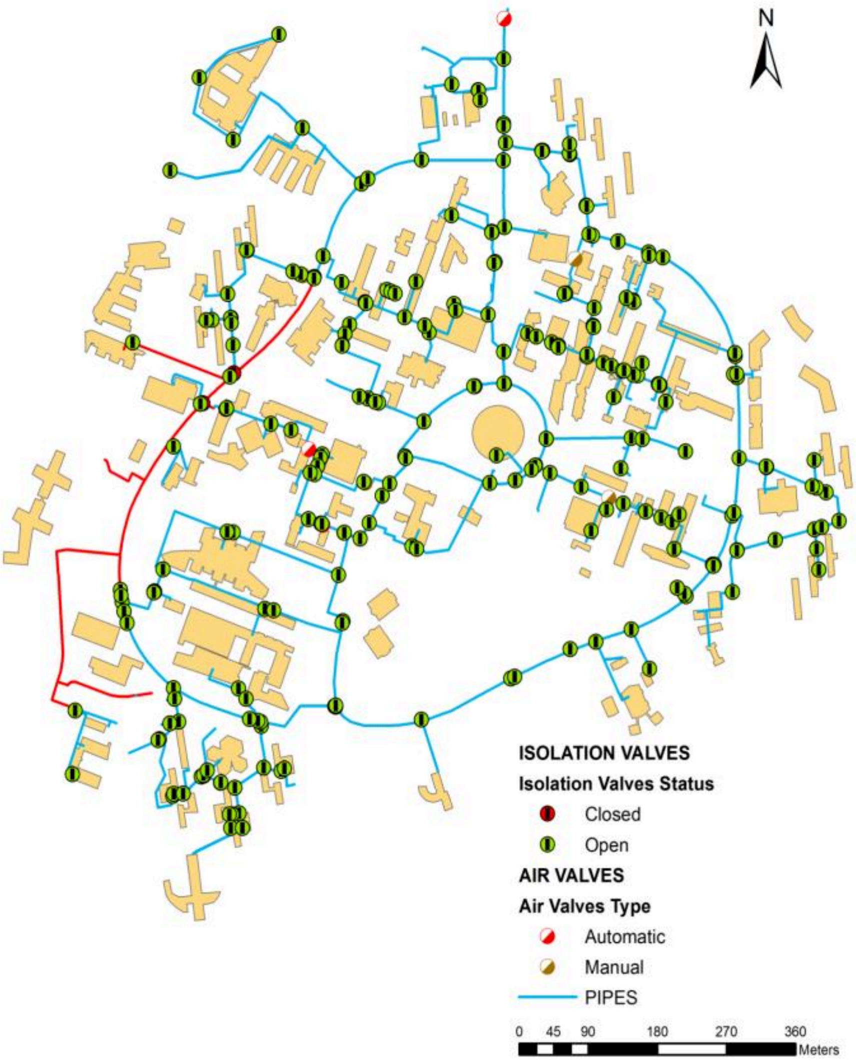


Figure 2.7 Distribution of valves in the Campus

Water pipes suffer from aging, soil aggressively and lack of renovation and maintenance. These factors caused frequent water leaks as demonstrated by Farah (2016). Figure 2.8 shows that water leak concerned the totality of the campus.

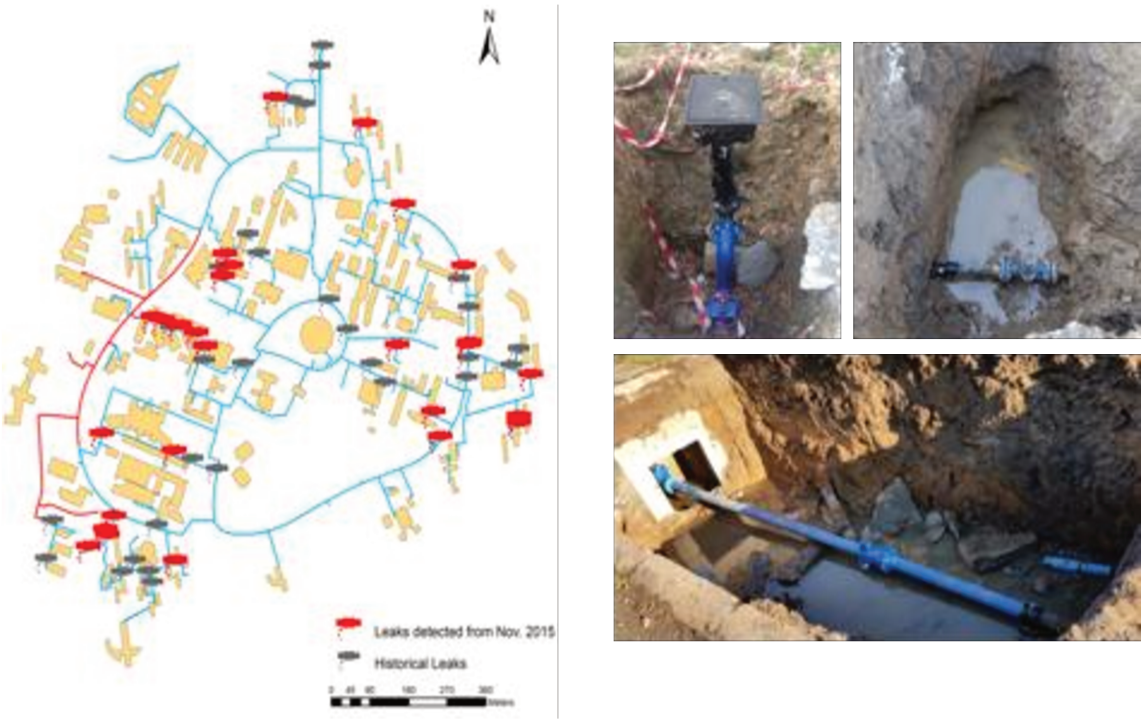


Figure 2.8. Distribution of water leaks in the Campus (Farah, 2016)

2.3.3 Water consumption monitoring

2.3.3.1 Automatic Reading Meters

The water network was equipped by 90 Automatic Reading Meters (AMR) (Figure 2.9), which record at one-hour interval the water supply of the campus as well as the water consumption in the main buildings of the campus. The consumption records are then transmitted via radio-transmission system to a server. Data are then analyzed using different methods to assess the integrity of the water network and to detect water leak.

Figure 2.10 shows the localization of the AMR in the campus. They include the following 13 supply AMR: (i) 4 AMRs at the water meter of 4 CANTONS, ECL, BACHELARD, LML, (ii) 5 AMR at the water meters of the CitéScientifique and (iii) 4 AMR at buildings M5, Hall Vallin, ICARE and CUEEP.

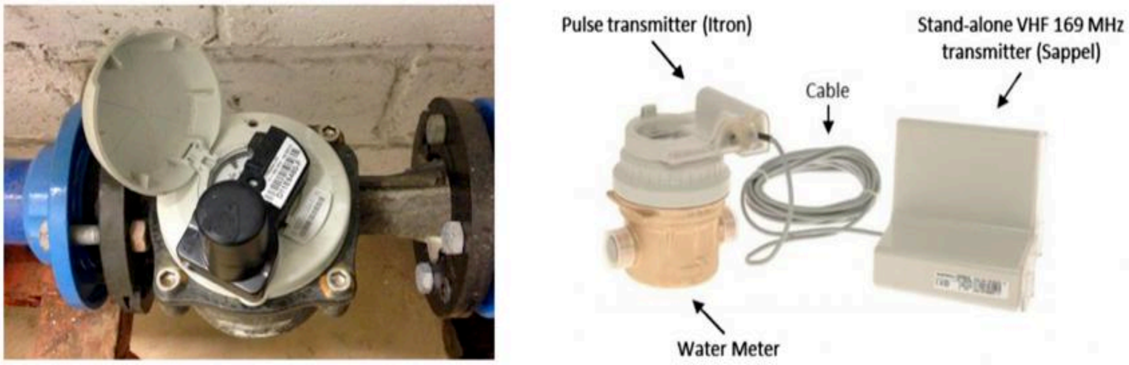


Figure 2.9 Automatic Reading Meters (AMR) used in the scientific campus

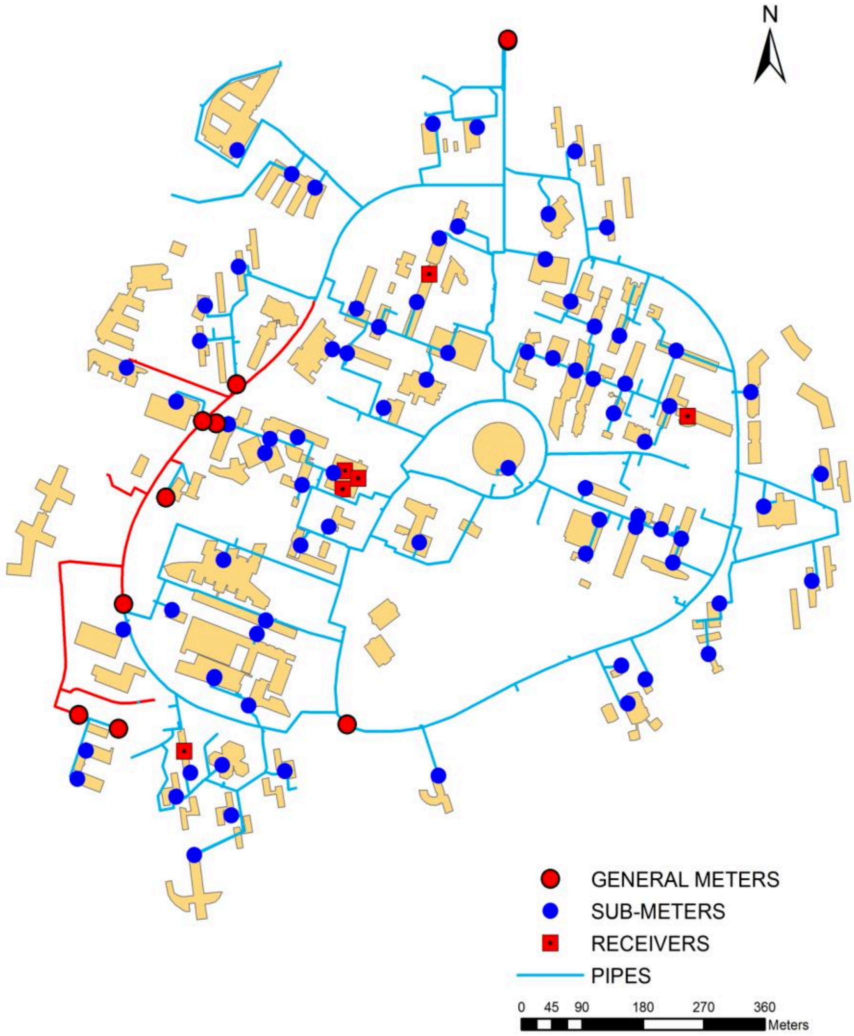


Figure 2.10: Distribution of the Automatic Reading Meters (AMR) in the campus. The red color designates the supply meters, while the blue color designates the consumption meters (Farah, 2016).

2.3.3.2 Data transmission

Data transmission from the AMR to the server is accomplished using a system composed of 4 data collectors, which are placed on the roof of 4 buildings (M1, P5, R - CAMUS and C7) (Figure 2.11).

Figure 2.12 shows a data collector. It consists of an antenna, a VHF card and a collector.

Data collectors are connected to the AMRs via radiofrequency transmission protocol at 169 MHz and to the server via GPRS. Data Transmission is operated as follows (2.13):

- The AMR measures the data consumption everyone hour, store the consumption values for 24 hours and then transmit the 24 consumption values to the nearest data collector (antenna) using radiofrequency protocol.
- The data collectors transmit the data collected from the AMR to the server every 24 hours using GPRS connection.
- The server stores the raw data and operates some verification about data integrity in particular about missing data and abnormal consumptions.
- The server transmits via email and SMS data and eventual awareness messages to the managers of the water distribution network.

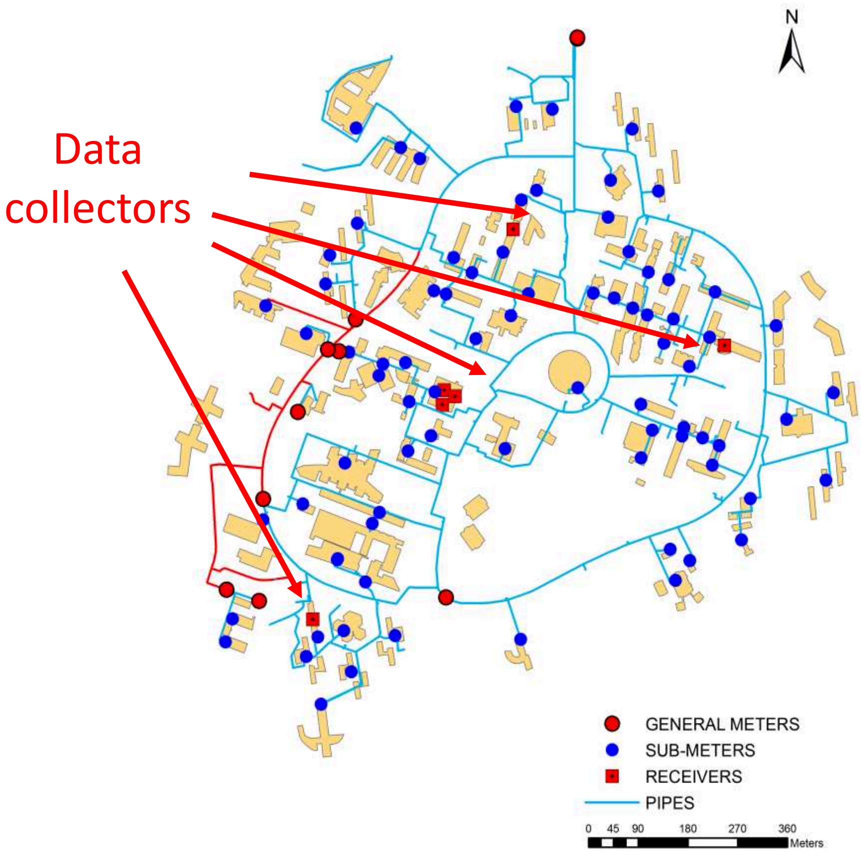


Figure 2.11: Distribution of data collectors in the campus

Figure 2.12 Data collector - Water distribution system of the campus

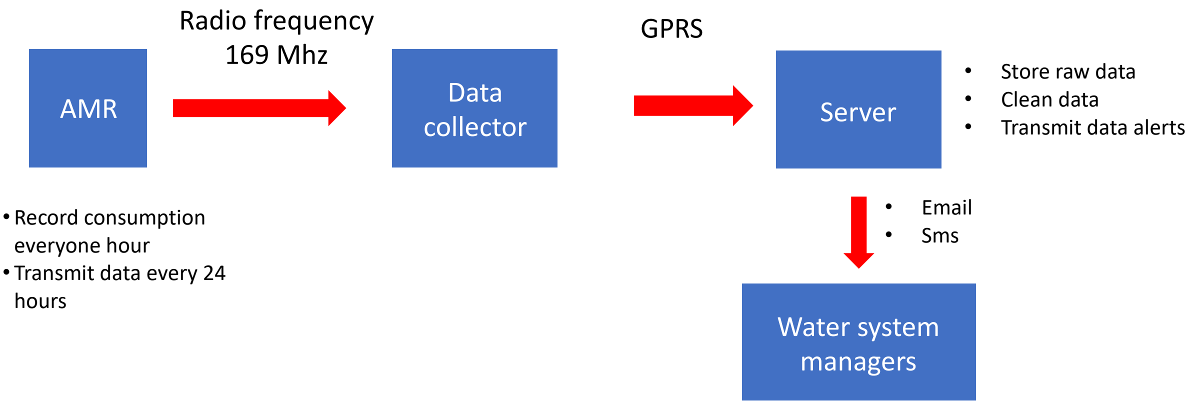


Figure 2.13Data transmission from the AMR to the mangers

2.3.4 Water consumption of the campus

Figures 2.14 and 2.15 show the box plot and violin graphs of the hourly water consumption of the campus for the period (2012 – 2016). The blue color refers to the working days, while the orange color refers to the weekend. Tables 2.1 and 2.2 summarize the statistical analysis of the water consumption for the working days and the weekend. We observe an important increase in the hourly working days consumption between 2012 and 2015. The average hourly water consumption increases from 26.2m^3/h in 2012 to 43.6 m^3/h in 2015. This increase is about 66%. For the weekend, the increase in the hourly water consumption between 2012 and 2015 is equal to 82%. In 2016, we observe a decrease in the water consumption, which comes back to the water consumption level in 2014. This variation in the water consumption could result from some leaks in the campus.

Figures 2.14 and 2.15 shows the presence of high values of the water consumption, mainly in 2015. These high values are related to water leaks. The detection of these leaks will be discussed in chapter 3.

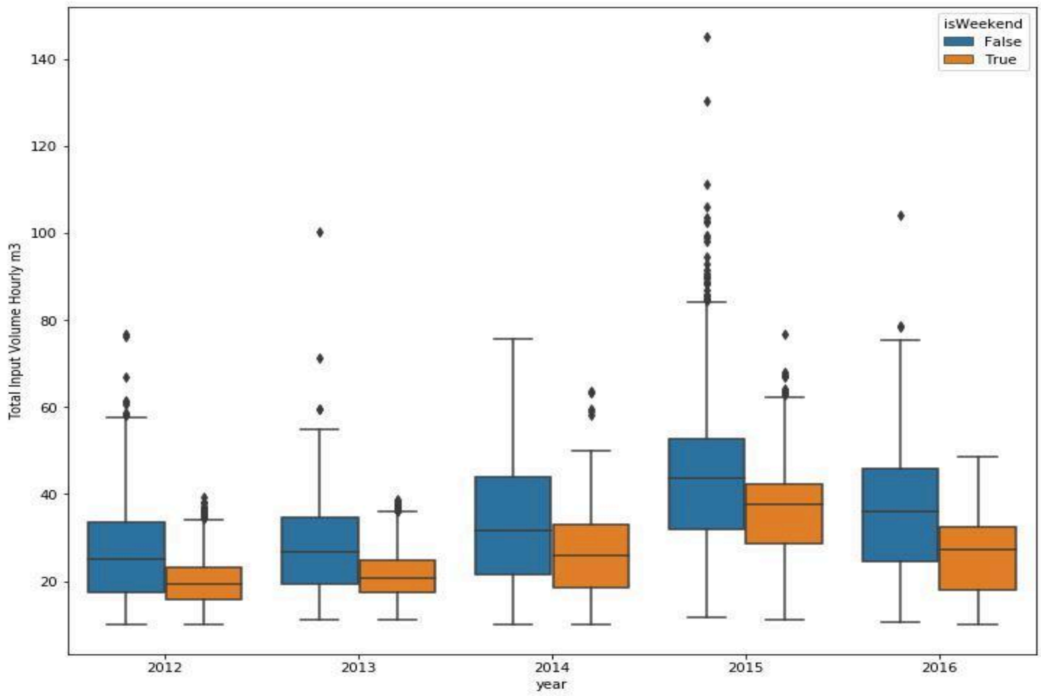


Figure 2.14: Hourly water consumption of the campus during working days (blue color) and weekend (orange color) (2012-2016)

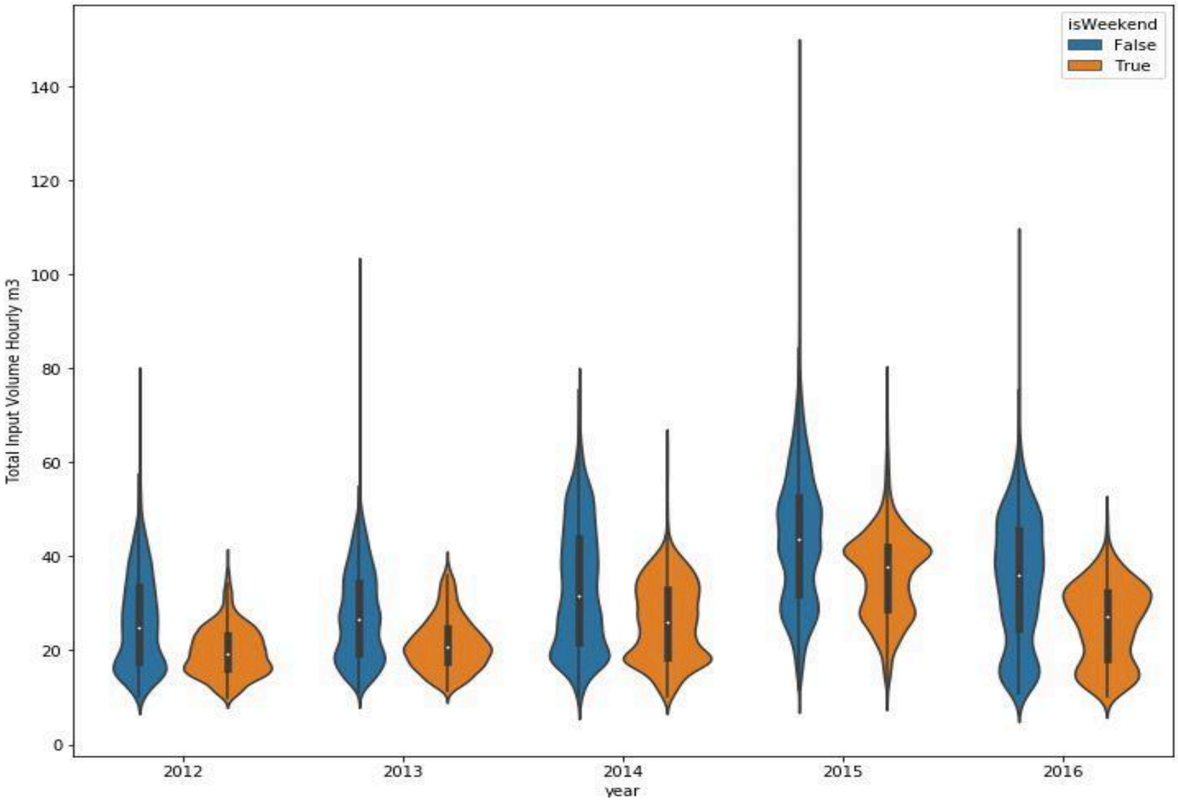


Figure 2.15 Violin plot of the hourly water consumption during working days (blue color) and weekend (orange color) (2012-2016)

Table 2.1: Descriptive statistics of the hourly water consumption of the campus (working days)

Year	2012	2013	2014	2015	2016
Mean	26.2	27.46	33.36	43.65	34.92
Min	10.02	11.12	10.02	11.67	10.68
Max	76.68	100.19	75.51	145.12	103.98
25%	17.437	19.32	21.57	31.81	24.54
50%	24.94	26.58	31.7	43.56	36.01
75%	33.53	34.61	43.82	52.77	45.79

Table 2.2: Descriptive statistics of the hourly water consumption of the campus (weekend)

	2012	2013	2014	2015	2016
Mean	19.79	21.54	26.2	36.07	25.75
Min	10.00	11.25	10.05	11.16	10.05
Max	39.38	38.72	63.56	76.68	48.58
25%	15.89	17.36	18.46	28.6	17.98
50%	19.35	20.82	26.02	37.74	27.18
75%	23.266	24.79	33.06	42.23	32.54

Figure 2.16 illustrates the variation of the hourly water consumption in 2015 (Farah, 2016). It shows some missing data, low consumption in summer vacations and some peaks, which are related to water leaks.

Figure 2.17 illustrates the variation of the weekly water consumption in 2015 (Farah 2016). It also shows low consumptions during holidays and some peaks, which are related to water leaks.

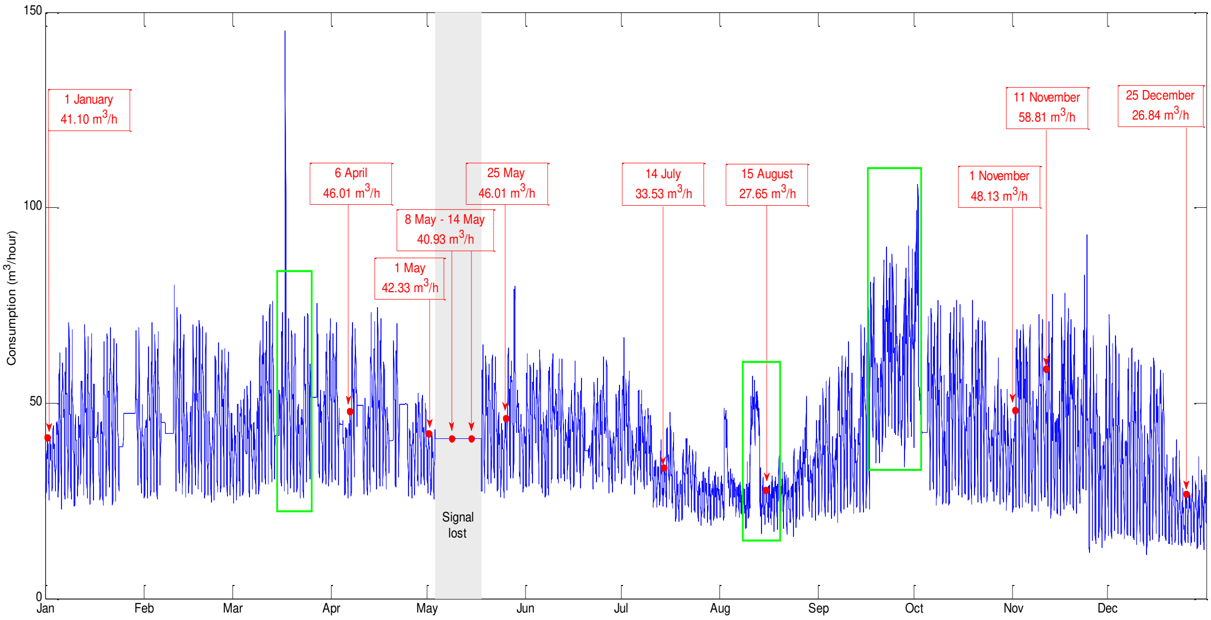


Figure 2.16:Hourly water consumption in 2015 (Farah, 2016)

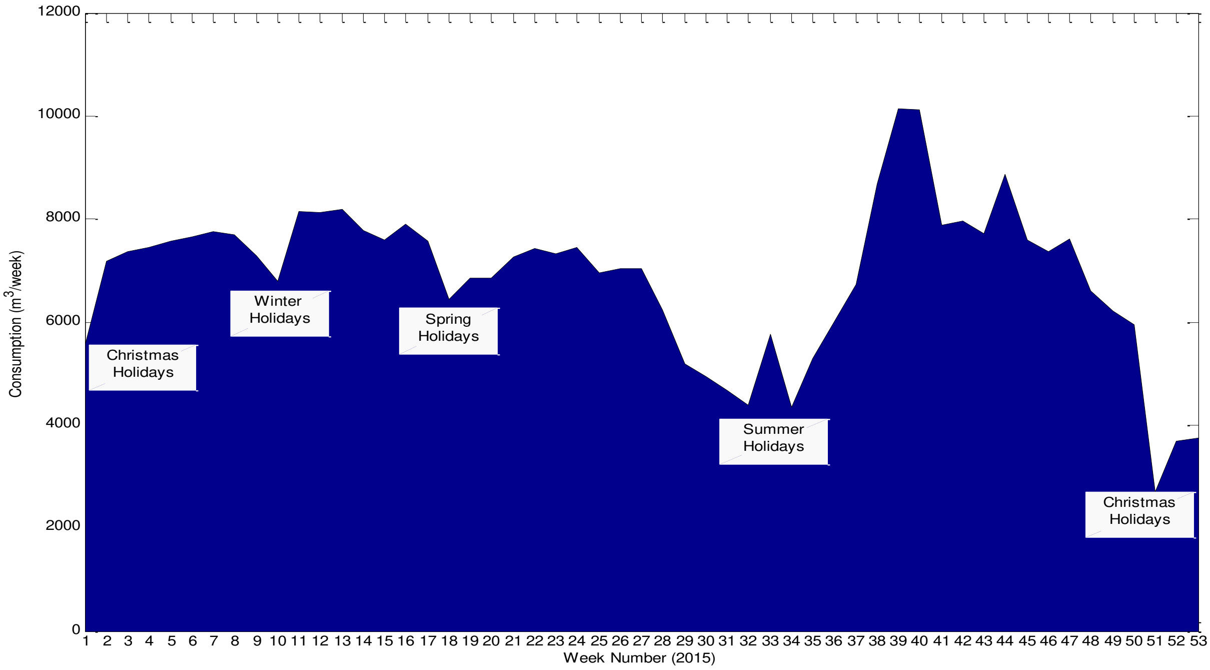


Figure 2.17: Weekly water consumption in 2015 (Farah, 2016)

2.4 Investigation of the water leak using a Lab pilot model

2.4.1 Overview

This research aims at investigating the possibility to localize the water leak in the water distribution network from the variation in the water supply flow. Indeed, for an area with multiple water supply such as the scientific campus, it is expected that the location of the water leak could impact the repartition of the water supply of this area. In order to explore this possibility a pilot of the water distribution system of the campus was constructed. The pilot provides the possibility to simulate one or more leaks in different locations of the campus and to measure the water flow from variation in the water supply sections of the campus.

In the following, we present successively the pilot and then its use to explore the possibility to localize the water leak from the variation in the water supply flow.

2.4.2 Presentation of the campus water distribution pilot

Figure 2.18 shows the pilot of the campus water distribution pilot. Its size is equal to 2m x 2m. It includes a large loop and a small loop connected by a set of pipes. Rigid multilayer pipes are used. The inner diameter of the pips is equal to 8.8 mm.

The water network is connected to three water supply sections, located in the North (S1), in the West (S2) and in the South (S3). A water tank is used for the water supply of the pilot in closed circuit. A pump ensures water transfer from the tank to the pilot. The pump has the following characteristics (Figure 2.19): Hmax = 36m and Qmax = 2.4 m^3/h. The water supply could be controlled through a series of valves.

A pressure-reducing valve is used to reduce and stabilize the pressure of the water supply. It can reduce the pressure down to 1 bar.

Flow meters are used to measure the water flow from the three water supply sections. They have the following characteristics:

- Accuracy ±1%
- Flow-rate range 10-120 L/min
- Max working pressure 20 bars

Six valves are used in the pilot for the simulation of the water leak in the campus. Figure 2.18 shows the locations of these valves: L1 in the North-East, L2 in the west, L3 in the North, L4 in the South, L5 in the South-East and L6 in the East. Valves are connected to a flow meter, which measure the water leak from the pilot.

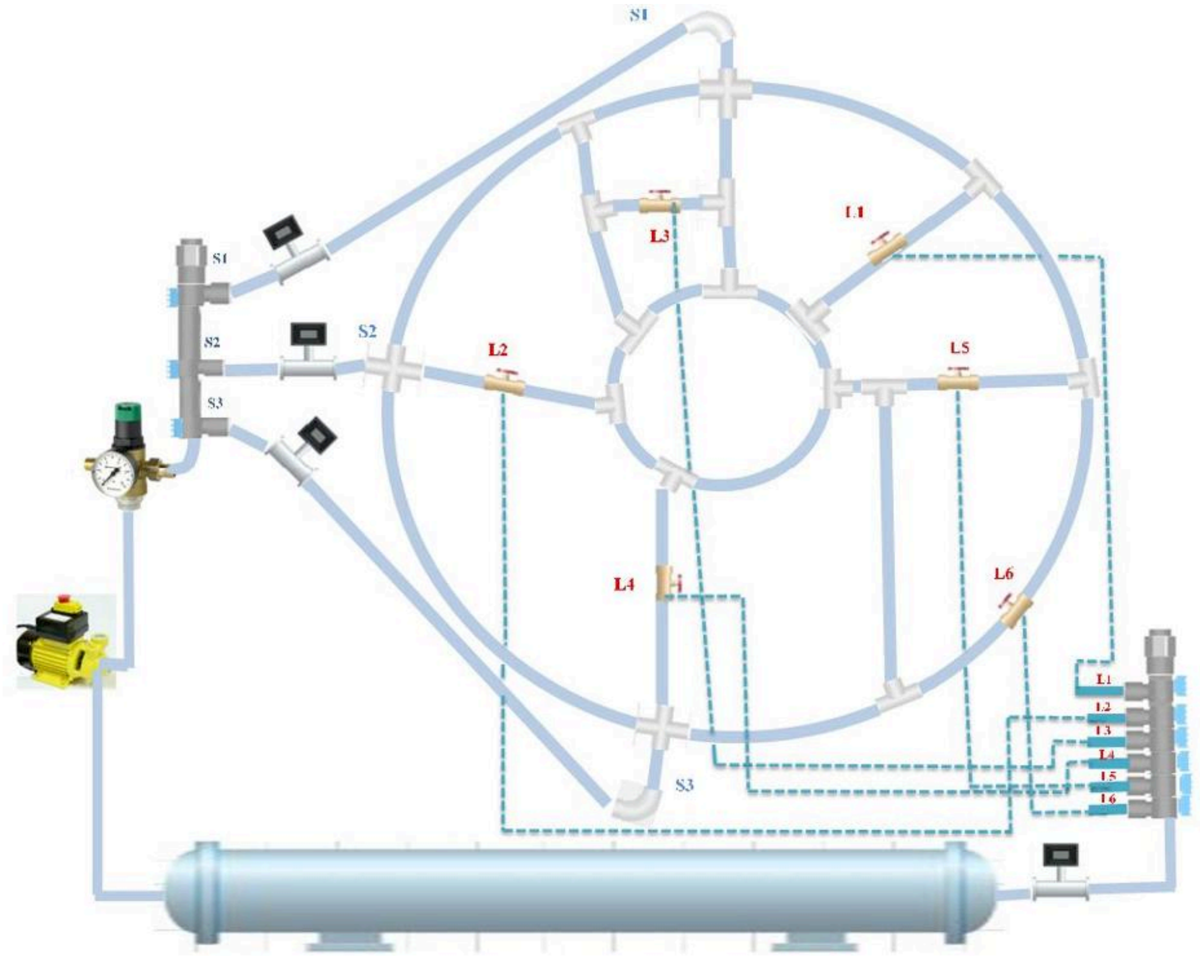


Figure 2.18: Pilot of the water distribution system of the scientific campus

Valve

Flow meter

Water pump

Figure 2.19: Pilot equipment and monitoring

Table 2.3 Characteristics of the pilot equipment and monitoring

Item	**Characteristics**
Pressure-Reducing Valve	-Medium water -Inlet Pressure max. 25 bar -Outlet pressure 1.5- 5.5 bar (present to 3 bar) -Min pressure drop 1 bar
Flow meter	-outlet & inlet size 1inch -measuring accuracy ±1% -flow-rate range 10-120L/MIN -measurement unit L, GAL, PTS, QTS -Max working pressure 20Bar
Pump	Hmax: 36m Qmax: 2400 L/h IPX4 Max: ≤35 ˚C
Multilayer pipe	Diameter: 12mm Inner diameter: 8,8mm Max. working pressure: 6 bars

2.4.3 Operating mode

Tests are conducted as follows:

- The pump as well as the valves for the water supply and water leak is opened.
- Valves are adjusted to control the water flow in the circuit.
- Flow and pressure data are recorded
- The test continues until the stabilization of the water flow and pressure. Stabilization is generally observed after 15 minutes.
- Stabilized flow and pressure data are recorded.
- The pump is stopped, and the valves are closed.

The reliability of the pilot and the operating mode were investigated through repeating the same test several times. Table 2.4 and figure 2.20 summarizes the results obtained in repeating four times the leak test from the valve L1. It shows good results with a maximum variation in the tests of 3.5% for the water supply S1.

Table 2.4: Results of 4 repeating tests with leak in the valve L1.

Supply rate (%)	Test1	Test2	Test3	Test4	Precision (Max-Min)/Mean (%)
S1	34,68764	34,31943	35,09487	35,55909	3,5
S2	33,07442	33,66007	32,99719	32,7903	2,62
S3	32,03936	31,93919	31,93693	31,71598	1

2.4.4 Results and discussions

Table 2.5 and figure 2.20 summarize the results of 6 experiments conducted with leaks L1, to L6, successively. They show the impact of each leak in the flow rates S1, S2 and S3. It could be observed that despite the low impact of the position of leaks in the water flow rate, we can distinguish the following:

- The highest flow rates for the source S1 are obtained with leaks L1, L3, and L5. This result is coherent with the proximity of these leaks from the source S1.
- The highest flow rate for the source S2 is associated with leak L2, which is close to the source S2.
- The highest flow rate for the source S3 is associated with leak L4, which is close to the source S3.
- The impact of leak L6 is the same for the three sources. This result is not coherent with the position of this leak, which is close to the source S3.

These results show a clear impact of leaks L2, L3, and L4 on the closest water source supply. The impact of other leaks (L1, L5 and L6) remains low.

Table 2.5 Impact of one leak (L1 to L6) on the flow rates S1, S2 and S3

Water Supply (%)	L1	L2	L3	L4	L5	L6
S1	34,9	30,8	36,1	27,7	34,9	34,5
S2	33,2	36,1	33,5	33,3	32,8	33,2
S3	31,8	33,1	30,4	39	32,3	32,2

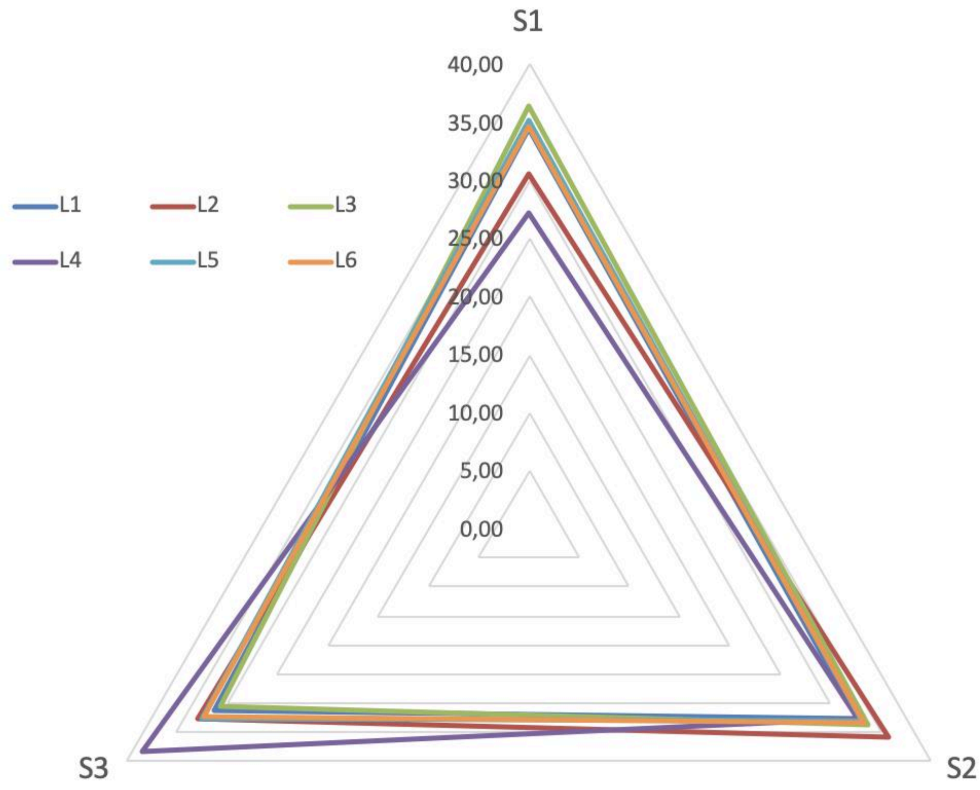


Figure 2.20 Impacts of one leak (L1 to L6) on the flow rates S1, S2 and S3

Table 2.6 and figure 2.21 summarize the impact of two leaks (L1 with successively L2, L3, L4 and L5) on the flow rates S1, S2 and S3.

It could be observed that:

- The highest flow rates for the source S1 are associated to leaks L1-L3, L1-L5 and L1-L6. This result is coherent of the proximity of these leaks from the source S1.
- The impacts of other leaks (L1 - L2 and L1 - L4) are not well distinguished.

Table 2.6: Impact of two leaks (L1 with successively L2, L3, L4 and L5) on the flow rates S1, S2 and S3

Water Supply (%)	L1-L2	L1-L3-	L1-L4	L1-L5	L1-L6
S1	34,8	35,9	33,1	35,2	35,4
S2	34	32,6	33,6	33,1	32,8
S3	31,2	31,5	33,3	31,7	31,8

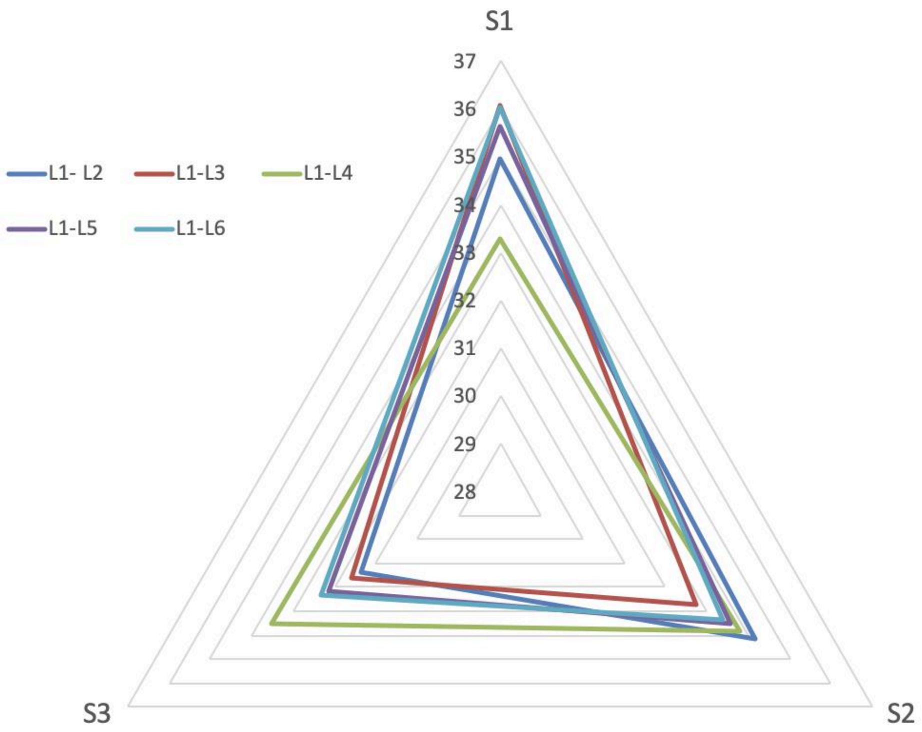


Figure 2.21 Impact of two leaks (L1 with successively L2, L3, L4 and L5) on the flow rates S1, S2 and S3

Table 2.7 and figure 2.22 summarize the impact of two leaks (L2 with successively L3, L4 and L5) on the flow rates S1, S2 and S3.

It could be observed that:

- The leak L2-L3 causes a high flow rate from the source S1.
- Leaks L2-L4, L2-L5 and L2-L6 have similar high impact of the flow rate S2 and S3 and low impact of the flow S1.

Table 2.7: Impact of two leaks (L2 with successively L3, L4, L5 and L6) on the flow rates S1, S2 and S3

Water Supply (%)	**L2-L3**	**L2-L4**	**L2-L5**	**L2-L6**
S1	36,5	29,1	29,8	31,4
S2	33,1	35	35	35,1
S3	30,4	35,9	35,2	33,5

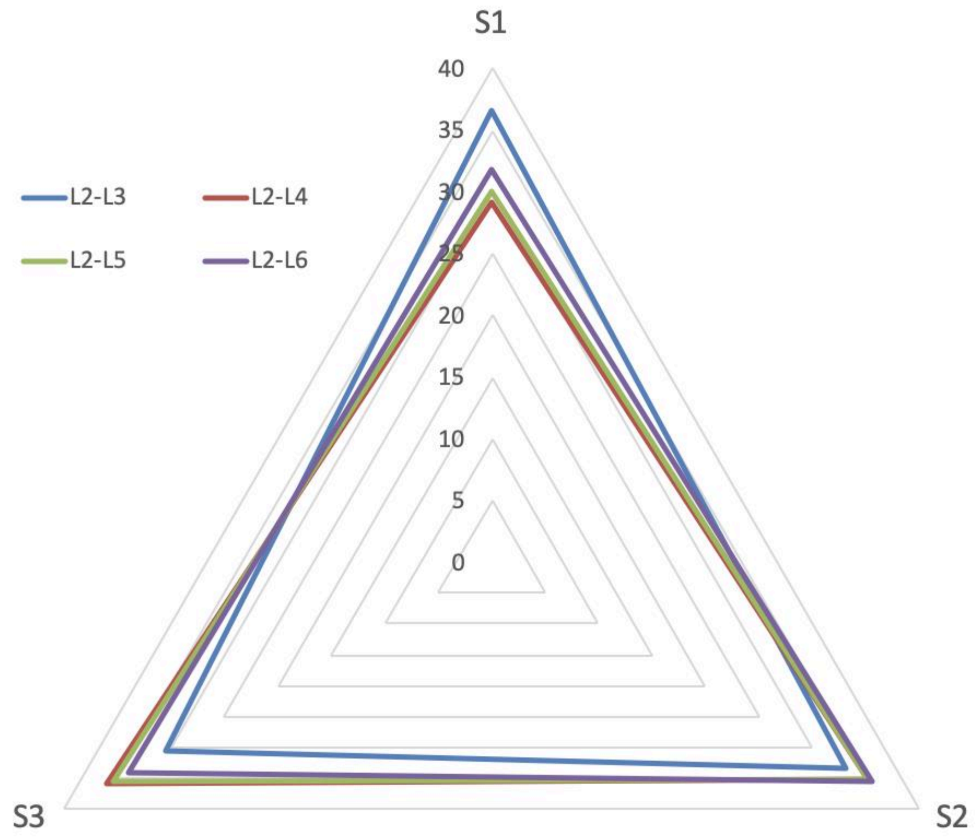


Figure 2.22: Impacts of two leaks (L2 with successively L3, L4, L5 and L6) on the flow rates S1, S2 and S3

Table 2.8 and figure 2.23 summarize the impact of two leaks (L3-L4, L3-L5 and L3-L6)on the flow rates S1, S2 and S3.It could be observed these leaks have high impact on the water flow rate S1. These results are coherent with the proximity of these leaks to the source S1.

Table 2.8: Impact of two leaks (L3-L4, L3-L5 and L3-L6) on the flow rates S1, S2 and S3

Water Supply (%)	L3-L4	L3-L5	L3-L6
S1	35,1	34,9	35,6
S2	32,8	33	32,6
S3	32,1	32,1	31,8

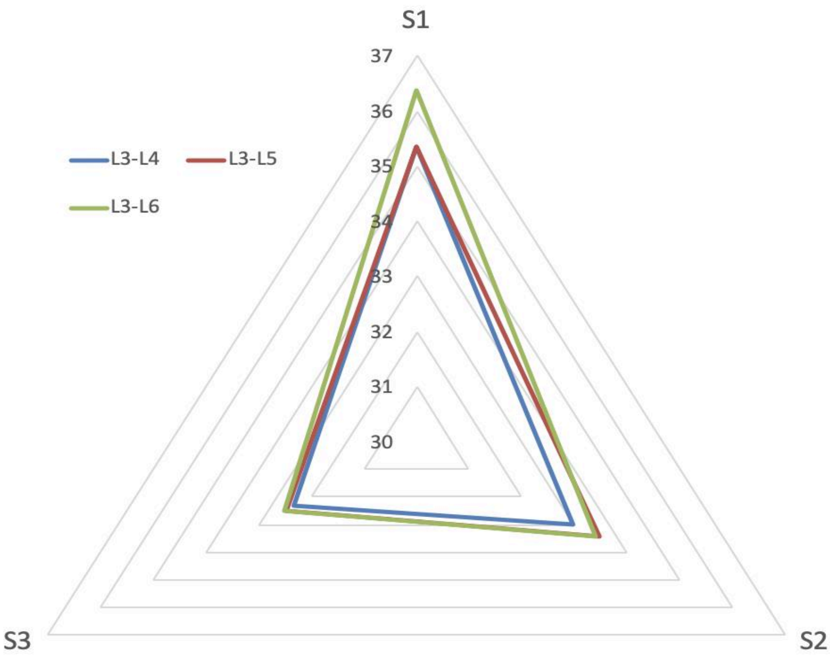


Figure 2.23 Impact of two leaks (L3-L4, L3-L5 and L3-L6) on the flow rates S1, S2 and S3

2.4.5 Partial conclusion

This section presented the construction of a Lab pilot of the water distribution network to investigate the impact of the position of a leak in the water flow from the water supply sources. The pilot did not include pressure monitoring. Results showed that a clear impact of the position leak is obtained when the leak position is close to the water source supply. For other locations, the impact is not clear, which means that the leak position could not be systematically determined from only the supply flow rates. Results obtained with two leaks confirm those obtained with one leak.

In the future, it could be interested to add pressure cells to record the pressure variation during a leak. The pressure data could be helpful in the determination of the leak position.

2.5 Conclusion

This chapter presented the water distribution system of the scientific campus of Lille University, which is used a support for this research. The of this water network for this research is motivated by its complexity and representatively of a water network of a small town with about 25 000 users. It is also motivated by the high degradation of this network due to ageing, which cause important water leak. The availability of data about the network asset and water consumption helped in conducted this research. Indeed, the water network is monitored by 93 Automated MeterReading (AMR) that record the water supply and consumption in the main buildings at hourly-time interval.

A rapid analysis of the campus water supply showed important leaks. Analysis of these leaks will be discussed in chapter 3.

This chapter presented also the construction of a Lab pilot of the campus water network to investigate the impact of leaks position on the water supply flow from. Tests conducted with a series of individual or double leaks showed a clear impact of leaksposition leak, when leaks' positions are close to the water source supply. For other locations, the leak impact is not clear, which means that leaks' position could not be systematically determined from only the supply flow rates. The performance of this pilot could be improved in the future by adding a system to measure the water pressure variation.

2.6 References

Farah E., A. Abdallah & I. Shahrour (2017). SunRise: LARGE SCALE DEMONSTRATOR OF THE SMART WATER SYSTEM. Int. J. Sus. Dev. Plann. Vol. 12, No. 1 (2017) 112–121

Farah E. (2016) « Detection of water leakage using innovative smart water system – Application to SunRise Smart City Demonstrator» Thèsesoutenue le 2 novembre2016; Université de Lille.

Farah F., I. Shahrour (2017) Smart Water for Leakage Detection: Feedback about the Use of Automated Meter Reading Technology. Sensors Networks Smart and Emerging Technologies (SENSET), Beirut, Lebanon, 2017, IEEE Xplore, DOI: 10.1109/SENSET.2017.8125061.

Farah, E. & I. Shahrour (2017) Leakage Detection Using Smart Water System: Combination of Water Balance and Automated Minimum Night Flow. Water Resources Management, 31, 4821-4833.

Saab C., I. Shahrour, F. Hage-Chehade (2017) Smart Technology for Water Quality Control: Feedback about use of water quality sensors. Sensors Networks Smart and Emerging Technologies (SENSET), Beirut, Lebanon, 2017, IEEE Xplore, DOI: 10.1109/SENSET.2017.8125060.

Saab Ch. (2018) "Smart system for water quality control: feedback from large-scale experimentation", Thèsesoutenue le 2 juillet 2018 ; Université de Lille, UniversitéLibanaise.

Shahrour I., O. Abbas, A. Abdallah, Y. AbouRjeily, A. Afaneh, A. Aljer, B. Ayari, E. Farrah, D. Sakr, F. Al Masri (2017). Lessons from a Large-Scale Demonstrator of the Smart and Sustainable City. In Happy City Book, A. Brdulak and H. Brdulak (eds.), Springer International Publishing AG 2017, DOI 10.1007/978-3-319-49899-7_11

3 Chapter 3: Use of the Artificial Intelligence techniques for the localization of leakage in water distribution system

3.1 Introduction

Water distribution constitutes a vital service for both citizens and economic activity. In many existing cities, the water distribution system suffers from different factors such ageing, lack in maintenance, high increase in the population, and recently climate change. The latter cause an increase in the frequency and the magnitude of adverse climate events such as flood, storms and variation in the temperature. Due to these factors, water the distribution system becomes more vulnerable to events related to natural, technological and manmade stresses. Water leak constitutes an important issue in the water system vulnerability, because leak in urban area causes interruption in water services with consequences on the quality of life for citizens as well as on the economic activity. It could also cause important economic losses due to the damage of water distribution system and the surrounding environment including other utilities, infrastructures and buildings. It could also lead to flood or/and environment contamination. As a consequence, the early detection of the water leakage becomes a top priority of cities and managers of water distribution systems.

Figure 3.1 shows the variation in the number of Web of Sciences indexed papers about water leakage during the period 1996 - 2019. It indicates an important and an increasing scientific concern for the water leakage. In 2019, the number of papers concerning the water leakage attained 1599. Figure 3.2 shows the variation of the number of papers about detection of water leakage. It also shows an important concern for this issue. The number of papers in 2019 is equal to 215.

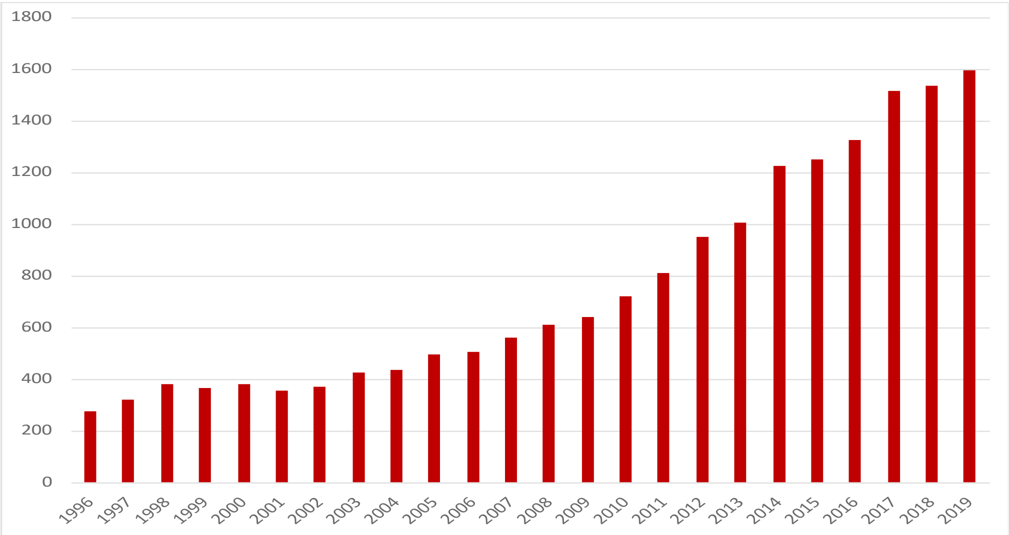


Figure 3.1: Web of Sciences indexed papers concerning water leakage (Source Web of Sciences) (period 1996 – 2019, total number 18,088)

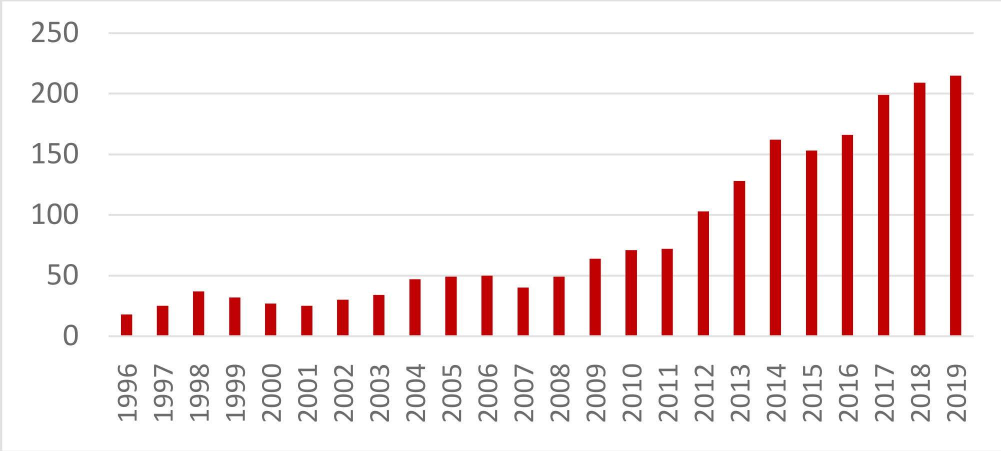


Figure 3.2: Web of Sciences indexed papers concerning detection of water leakage (Source Web of Sciences) (period 1996 – 2019, total number 2,005)

With the development of real-time monitoring of the water pressure and flow, data are being widely used for the detection and localization of leakage in the water distribution networks. The principle key adopted in the literature for leak detection is to compare real-time data with the normal range of the system behavior with a focus on flow and pressure data (Salam et al. 2014; Wu & Liu 2017; Chan et al. 2018; Zhou et al. 2019).

The Artificial Intelligence-based methods have been widely used for both the detection and localization of the water leakage in the water distribution systems, including supervised methods (Mounce et al., 2011; Zhanget al., 2016; Soldevila et al., 2016; Soldevila et al., 2017; Ciupke, 2018; Shravani et al., 2019; van der Walt et al., 2018), unsupervised methods (Wu et al., 2015; Zhang et al., 2016), and deep learning methods (Caputo&Pelagagge, 2003; Salam et al., 2014; Mounce et al., 2015; van der Walt et al., 2018; Yalçin et al., 2018; Zhou et al., 2019; Shravani et al., 2019; Rojek&Studzinski, 2019).

Section 2 of this chapter will present a literature review about the recent researches on the use of the Machine Learning techniques for the localization of the water leakage in the water distribution systems. This literature review shows that intensive researches have been conducted for the use of the Machine Learning methods for leak localization. However, the literature is still missing a comparison of the capacities of the different categories of the ML techniques for the localization of the water leakage on the same water distribution system. This chapter proposes to fill this gap by comparing the capacities of these methods to localize leakage in a complex water distribution system, which is based on the water network of the scientific campus of Lille University in France.

The chapter is organized as follows. The first part presents a literature review about the use of Machine Learning methods for leak detection. The second part presents the methodology of this research including data generation and use of the Machine Learning methods. The third part presents the results of the application of three categories of the Machine Learning methods: supervised, unsupervised and deep learning. The last presents synthesis of this research and recommendations about the use of ML for the localization of leakage in water distribution systems.

3.2 Literature review

Caputo & Pelagagge (2003) used the Artificial Neural Networks (ANN) for the detection and localization of the water leak in the water distribution systems. Data were generated using a hydraulic model of the water network for various operating conditions of the water network as well as for cases with different locations and amount of water leak. Two-level architecture was used, which concern the localization of the water leak and the determination of the amount of the leak, respectively. The method was tested on a small water network. Test showed that the method detected correctly the leaking branch as well as the leak flow rate. Salam et al. (2014) used the Radial Basis Function Neural Network method for leak detection. The hydraulic software EPANET was used for data generation. Data included pressure variation at the water network junctions resulting from leaks created in the water network. The pressure variations in the water network were used as input data for the ANN model, while the leak intensity and locations constituted the output parameters. The authors showed that the method was able to detect the magnitude and the location of leakage with 98 % accuracy.

Mounce et al. (2015) used the ANN method in a pattern matching-based approach to identify anomalies in the water distribution time series data. This method is based on the similarity research between new events and profiles established from past events. This research allowed the classification of the new events and consequently to identify abnormal events, which could be related to leak.

Zhang et al. (2016) used the multiclass Support-Vector Machine method (SVM) for leakage detection in a large-scale water distribution network. The method K-means clustering was used to subdivide the water network into leakage zones. Data with leak events were generated using the Monte Carlo method together with hydraulic model simulation to train the SVM clustering. Authors showed that the trained multiclass SVM was able to identify leakage zone using values of both water flow and pressure. Chan et al. (2018) reported that this method faced a big challenge concerning the determination of the number of clusters and the high impact of the random determination of the first cluster on the clustering process.

Soldevila et al. (2016) used the K-nearest neighbors for a classification of data generated by the hydraulic model EPANET from the simulation of leak events at the totality of the nodes of the water distribution network. Nodes that have similar effects on the pressure variation were grouped in the same class. These data were then used to train the K-nearest neighbors' method with the objective to localize the leak area. The good performance of this method in the localization of one water leak was assessed on three examples. Soldevila et al. (2017) used the Bayesian classifier for leak localization in a water distribution network. The method included two stages. In the first stage, a hydraulic model was used to generate data for different potential leaks in considering uncertainty conditions. A probability density functions was then calibrated on these data. In the second stage, the Bayesian classifier was used for the analysis of the pressure recorded data for the determination of the probability of potential leaks. This analysis allowed the localization of leaks in the water distribution system. The performance of the method was illustrated on the Hanaoi and Nova Icaria water distribution networks.

Han et al. (2017) presented a comprehensive system, called Aqua SCALE, for water leakage detection. This system is based on collecting data from different information sources including IoT sensing data, geophysical data, human data, and hydraulic modeling software. Data were then analyzed using the Machine Learning techniques for the creation of normal operating profiles. The comparison of new data with existing profiles allowed to detect anomalous events including leak. The system was evaluated on real water networks under different failure scenarios. Results showed that the system was able to locate simultaneous pipe failures with high level of accuracy.

Ciupke (2018) used the Regression Tree method to detect water leakage. This method was used for water demand modeling taking into account the demand changes due to holidays and seasonal variation. Alerts were established when the water flow exceeded normal water flow range. The method was tested on real examples and gave very good results, even for the detection of small leaks.

Van der Walt et al. (2018) analyzed the capacity of the Bayesian Probabilistic Analysis, the Support Vector Machine, and the Artificial Neural Network to detect and localize water leakage in water distribution networks. Both pressure and water flow data were used for training and testing these methods. The methods were compared on data generated from numerical modeling as well as on data recorded in laboratory tests. Since analysis showed that the performances of these methods depend on the complexity of the water network and amount of available data, authors did not propose general recommendations for the use of Machine Learning methods for leak detection.

Rojek&Studzinski (2019) used the ANN method for the detection and localization of water leak in water distribution systems. The method is based on the use of a (i) water network monitoring, (ii) a calibrated hydraulic model of the water network, and (iii) an ANN classifier. Water network monitoring together with the hydraulic model allowed leak detection, while the ANN method was used for leak localization. The method was tested on real off-line data. Tests showed that the ANN method correctly identified the localization of simulated leaks.

3.3 Research Methodology and materials

3.3.1 Research Methodology

This research aims at investigating the capacity of the Machine Learning techniques to localize the position of leaks in urban water distribution systems using flow and/or water pressure data. Considering the complexity of the urban water distribution networks, this research focuses on the first step of the leak localization, which concerns the identification of the sub-areas concerned by the leak. Each urban network is decomposed in subareas, which are called "zones" in this chapter. The monitoring system includes only a pressure cell for each zone and flow meters at the sub areas supply sections. This choice is related to the fact, that the installation of the pressure cells in the existing water networks in less complex and less expensive than the installation of the flow meters. The research includes two phases. The first phase concerns the generation of data, while the second is related to the investigation of the performance of the Machine Learning techniques on the generated data.

Data were generated using the software EPANET, which is presented below.

3.3.2 Presentation of EPANET

EPANET software was developed by the Water Supply and Water Resources Division (Formerly the Drinking Water Research Division) of the U.S. Environmental Protection Agency - National Risk Management Research Laboratory. It is public domain software that could be freely copied and distributed. EPANET has capabilities to perform extended period simulations and determine water quality behavior within pressurized pipe networks. It is largely used by researchers due to its free availability and performance.

EPANET provides an integrated environment for hydraulic and water quality network modeling, editing input data, running hydraulic and water quality simulations, and viewing the results. In EPANET, a water network is composed of pipes, nodes (pipe junctions), pumps, valves and reservoirs. The program EPANET tracks the flow of water in pipes, the pressure at nodes, the height of the water in reservoirs, and the concentration of a chemical species in the water network. The program is designed as a research tool for water networks. It can be used for different kinds of applications: sampling program design, model calibration, chlorine analysis and consumer assessment. EPANET could help in assessing management strategies, such as (i) altering source utilization within multiple source systems, (ii) altering pumping and tank filling, (iii) use of satellite treatment, such as re-chlorination at storage tanks and (iv) targeted pipe cleaning and replacement.

EPANET computes junction heads and link flows for a fixed set of reservoir levels and water demands over a succession of points in time. From onetime step to the next reservoir levels and junction demands are updated according to their prescribed time patterns while tank levels are updated using the current flow solution. The solution for heads and flows at a particular point in time involves solving simultaneously the conservation of flow equation for each junction and the head loss relationship across each link in the network. This process requires a use of an iterative technique to solve the nonlinear equations involved. The "Gradient Algorithm" is used in EPANET. For a good use of this program, engineers must understand the basic hydraulics of the system to be able to interpret the results properly.
The solution of the problem must satisfy three basic requirements:

- Continuity must be satisfied: the flow into a junction of the network must equal the flow out of the junction.
- The head loss between any two junctions must be the same regardless of the path in the series of pipes taken to get from one junction point to the other. This requirement is related to the requirement the water pressure continuity throughout the network.
- The flow and head loss must be consistent with the appropriate velocity-head loss equation.

3.3.3 Data generation

The water distribution network of the scientific campus of Lille University was selected as a support of this research. The campus is representative to a small town of 110 hectares, 150 buildings and about 25 000 users including students, faculty members and technical and administrative staffs (Shahrour et al., 2017).Figure 3.1 illustrates the water distribution network of the campus (Farah et al. 2017, Farah and Shahrour 2017). The water network is composed of 15 km of strongly meshed pipes. The water company supplies the campus with water at three sections, which are located in the North, West and South of the campus.

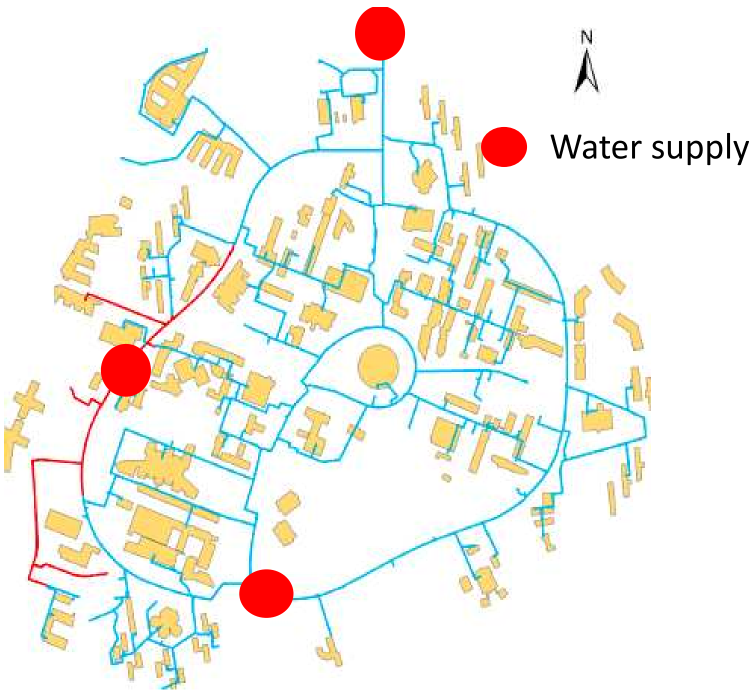


Figure 3.1: Water distribution system of the Scientific Campus (Farah and Shahrour 2017)

EPANET software (Haxton et al., 2019) was used for the generation of data related to the water leak in the water network. For the purpose of this research, the water distribution network was simplified into a network composed of 260 pipes and 244 junctions. This network was divided into5 zones, which are illustrated in Figure 3.2.Data were generated by modeling the water leak according to 215 leak scenarios (Table 3.2). Zone 1 is the largest and most complex zone. It included 67 leak scenarios. Zones 4, 5, 2 and 3, include 47, 44, 35 and 33 leak scenarios, respectively.

For each leak scenario, EPANET was used for the determination of the water supply (flow) form the three tanks (FL1, FL2 and FL3) as well as the values of the pressure at the five observation nodes, which are given in table3.2 and figure 3.3.

Each leak scenario was modeled under two conditions. The first condition concerns a constant pressure at the water supply sections, which were considered as tanks with a constant water height (H = 40 meters). The second condition concerns the water leak, which was considered by the following condition between the pressure (P) and water the flow (Q):

$$Q = C{*}P^{\alpha} \qquad \text{(Eq. 3.1)}$$

The parameters C and α characterize the water leak. Simulations were conducted with the following values: $\alpha = 0.5$ and C = 1.

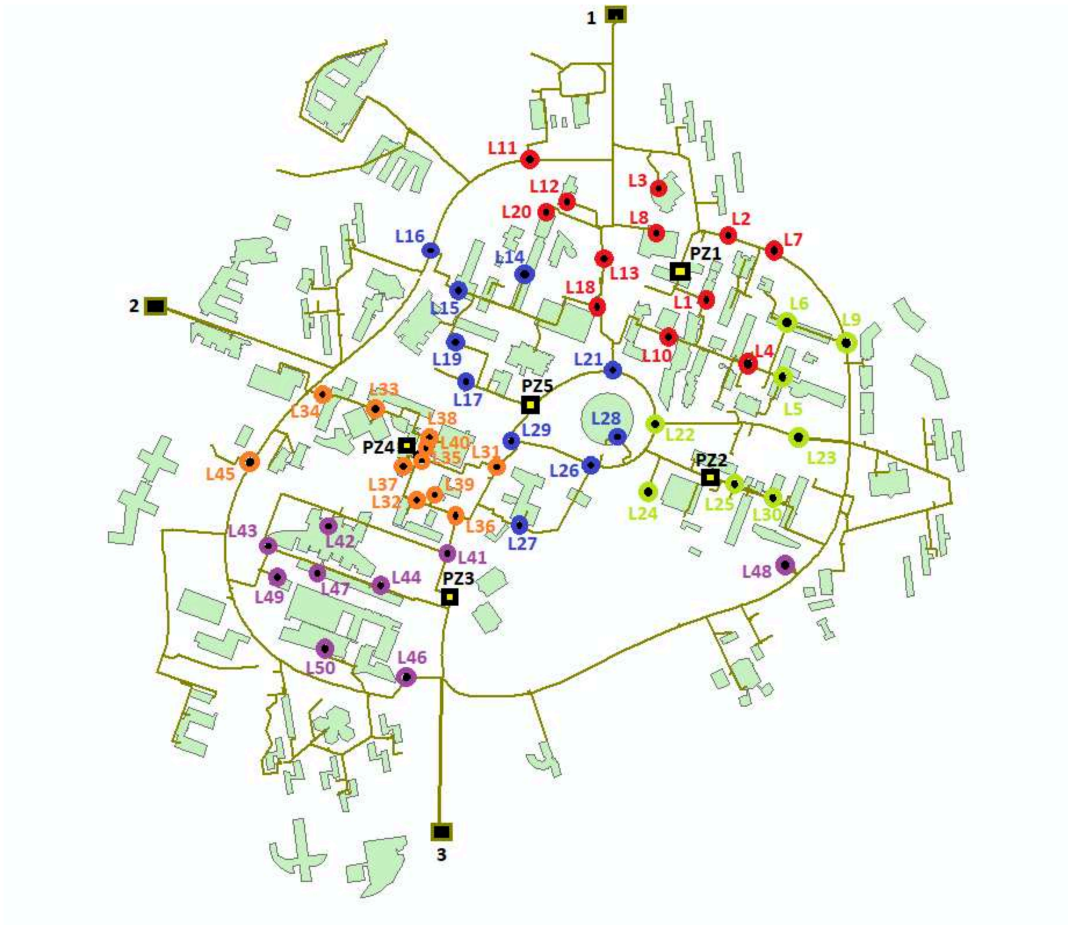


Figure 3.2: Simplified water network of the scientific campus (260 Pipes, 244 Junctions)

Table3.1:Leak scenarios used for the generation of leak dataset (leak nodes are given in figure 3.2)

Zone	Position of water leak
1 (67 leak scenarios)	L1, L2, L3, L4, L7, L8, L10, L11, L12, L13, L15, L18, L20 L10 + L1; L10 + L12; L10 + L13; L10 + L18; L10 + L2; L10 + L20; L10 + L3; L10 + L4; L10 + L7; L10 + L8; L11 + L12; L11 + L13; L11 + L18; L11 + L2; L11 + L20; L11 + L3: L11 + L4; L11 + L7; L11 + L8; L12 + L13; L12 + L18; L12 + L2; L12 + L20; L12 + L4; L12 + L7; L12 + L8; L13 + L18; L13 + L2; L13 + L20; L13 + L3; L13 + L4; L13 + L7; L13 + L8; L18 + L2; L18 + L20; L18 + L3; L18 + L4; L18 + L7; L18 + L8; L2 + L20; L2 + L4; L2 + L7; L2+ L8; L1+L10; L1+L11; L1+L12; L1+L13; L1+L18; L1+L2; L1+L20; L1+L3; L1+L4; L1+L7; L1+L8
2 (35 leak scenarios)	L5, L6, L9, L22, L23, L24, L25, L30 L23+L22; L24+L22; L24+L23; L25+L22; L25+L23; L25+L24; L30+L22; L30+L23; L30+L24; L30+L25; L5+L22; L5+L23; L5+L24; L5+L25; L5+L30; L6+L22; L6+L24; L6+L25; L6+L30; L6+L5; L9+L22; L9+L23; L9+L24; L9+L25; L9+L30; L9+L5; L9+L6:
3 (33 leak scenarios)	L41, L42, L43, L44, L46, L47, L48, L49, L50 L41+L44; L41+L46; L41+L47; L41+L48; L41+L50; L42+L44; L42+L46; L42+L47; L42+L48; L42+L50; L43+L44; L43+L46; L43+L47; L43+L48; L43+L50; L44+L46; L44+L47; L44+L48; L44+L50; L46+L47; L47+L48; L47+L50; L48+L49; L50+L49
4 (47 leak scenarios)	L31, L32, L33, L34, L35, L36, L37, L38, L39, L40, L45 L31+L36; L31+L37; L31+L38; L31+L39; L31+L45; L32+L36; L32+L37; L32+L38; L32+L39; L32+L45; L33+L36; L33+L37; L33+L38; L33+L39; L33+L45; L34+L36; L34+L37; L34+L38; L34+L39; L34+L45; L35+L36; L35+L37; L35+L38; L35+L39; L35+L45; L36+L37; L36+L38; L36+L39; L36+L45;L37+L38; L37+L39; L37+L45; L38+L39; L38+L45; L39+L45; L40+L45
5 (44 leak scenarios)	L14, L16, L17, L19, L21, L26, L27, L28, L29 L14+L21; L14+L26; L14+L27; L14+L28; L14+L29; L15+L21; L15+L26; L15+L27; L15+L28; L15+L29; L16+L21; L16+L26; L16+L27; L16+L28; L16+L29; L17+L21; L17+L26; L17+L27; L17+L28; L17+L29; L19+L21; L19+L26; L19+L27; L19+L28; L19+L29; L21+L26; L21+L27; L21+L28; L21+L29;L26+L27; L26+L28; L26+L29; L27+L28; L27+L29; L28+L29

Table3.2: Pressure observation nodes (Positions are given in Figure 3.2)

Zone	Pressure observation node
1	PZ1 (P12)
2	PZ2 (P23)
3	PZ3 (P31)
4	PZ4 (P41)
5	PZ5 (P52)

Table 3.3 provides a statistical analysis of the dataset. It shows that tank 1 provides the highest rate of campus water supply (Supply flow rate = 0.41), followed by tank 2 (flow rate = 0.35). It means that the water supply of the campus is mainly provided from the North and West of the Campus, where the construction density is higher than that in the South of the Campus. The highest average pressure is observed in zone 3, which is located in the South of the Campus, (average pressure around 35 m), followed by zones 5 and 4 (average pressure around 30 m). The average pressure in zones 1 and 2 is around 28 m.

Table 3.3: Statistical descriptive parameters of the pressure and flow rates values

	Min	**Max**	**Average**	**Standard Deviation**
FL1 (%)	0.13	0.71	0.4	0.15
FL2 (%)	0.23	0.62	0.35	0.82
FL3 (%)	0.60	0.59	0.23	0.11
PZ1 (m)	2.0	39.7	28.4	9.4
PZ2 (m)	1.4	39.4	27.4	9.3
PZ3 (m)	11.0	39.8	35.6	4.5
PZ4 (m))	1.8	39.2	29.2	10.2
PZ5 (m)	4.9	39.4	30.0	7.4

Figure 3.3 illustrates the impact of the leak position on the flow rate ratios FL1, FL2, and FL3. It shows that leaks in zones 1 and 2 cause high flow rate from tank 1 (FL1), medium flow from tank 2 (FL2) and low flow from tank 3 (FL3). Leaks in zone 3 cause high flow rate from tank 3 (FL3), medium flow from tank 2 (FL2) and low flow from tank 1 (FL1).Leaks in zone 4 cause high flow rate from tank 2 (FL2) and low to medium flow from tank 3 (FL3). Finally, leaks in zone 5 cause high flow rate from tank 1 and tank 2 (FL1 and FL2), and flow from tank 3 (FL3). Table 3.4 summarizes the impact of leak position on the water flow rate from the three tanks. It could be observed that a high flow rate from tank 2 (FL2) could be attributed to a leak in zone 4, and a high flow rate from tank 3 (FL3) could be attributed to a leak in zone 3. A high flow rate from tank 1 (FL1) could be attributed to zone 1. Medium flow rate from tanks 1 and 2 could be related to leaks zones 2 and 5.

Figure 3.4 illustrates the impact of the leak position on the pressures PZ1 to PZ5. It shows that leaks in each zone cause significant drop in the pressure in the leak zone. It also shows an impact of some leaks in a zone on the pressure on other zones, such as the impact of (i) leaks in zone 1 on the pressure in zone 2 (ii) leaks in zone 2 on the pressure in zone 1, (iii) leaks in zone 3 on the pressure in zone 4, and (iv) leaks in zone 5 on the pressure in zone 2.

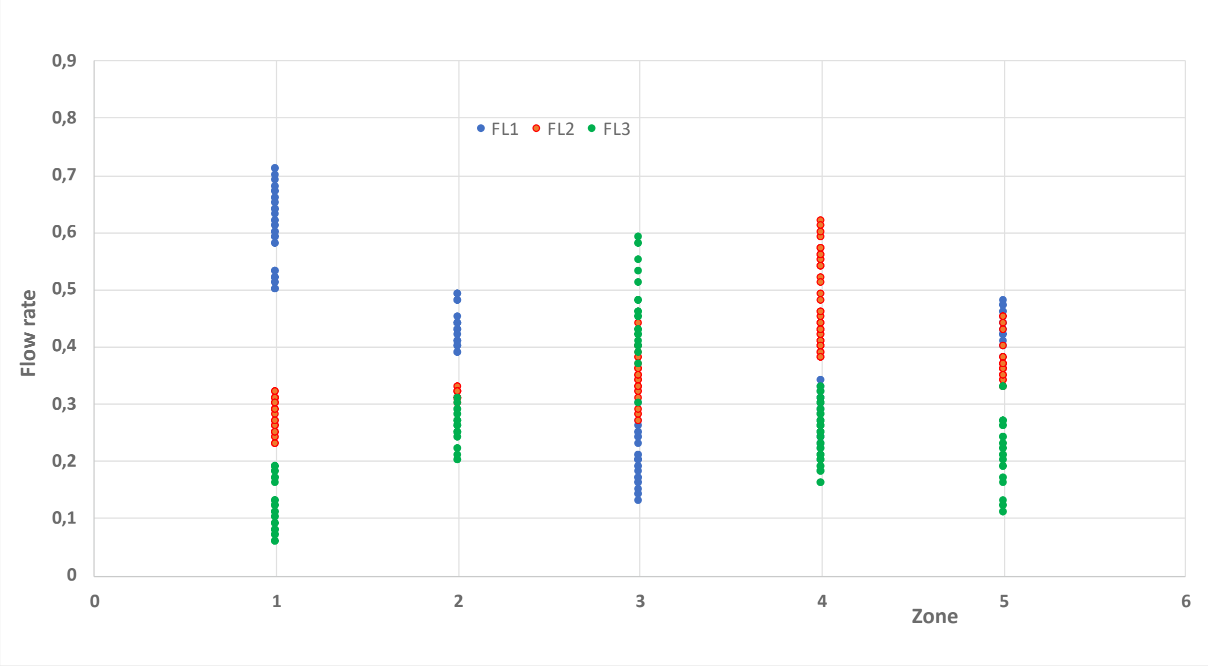


Figure 3.3: Impact of the leak localization on the flow rate values

Table3.4: Impact of the leak position on the water supply rate from (FL1, FL2, and FL3)

Leak zone	FL1	FL2	FL3
1	Strong	Medium	Low
2	Medium	Medium	Low to medium
3	Low	Medium	Strong
4	Low	Strong	Low to medium
5	Medium	Medium	Low to medium

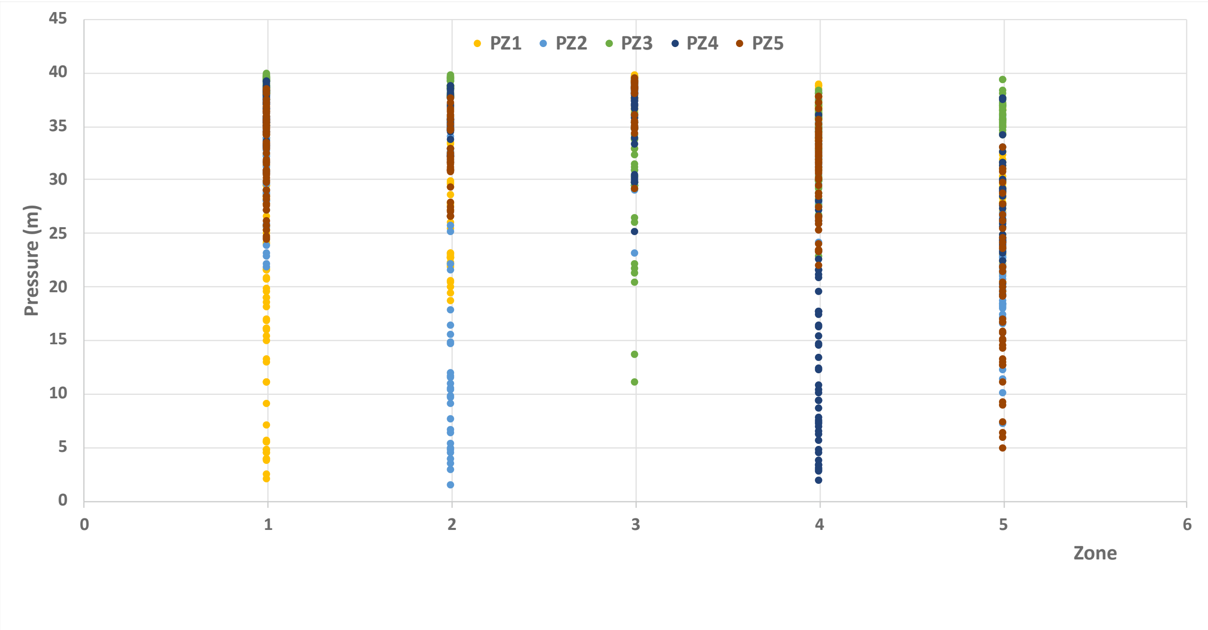


Figure 3.4: Impact of the leak localization on the pressure values at the observation points

3.4 Application of the Machine Learning Methods

This section presents the use of the three Machine Learning techniques (supervised, unsupervised, and Artificial Neural Network) for the localization of leaks in the water distribution network of the scientific campus. The localization is based on the use of the water supply rate from the three tanks (FL1, FL2, and FL3) and the pressure values in the observation pressure nodes in the 5 zones of the campus (PZ1 to PZ5).

Analyses were conducted using Kaggle platform[1]. The following sections present briefly the methods used in this research.

3.4.1 Supervised methods

Analyses were conducted with three supervised methods: Logistic regression, Decision Tree, and Random Forest.

Logistic Regression

The Logistic Regressions used for binary classification. The method used in this work is based on the function:

$$h_\theta(x) = g(\theta^T x) \text{(Eq. 3.2)}$$

$$g(z) = \frac{1}{1-e^{-z}} \text{ (Eq. 3.3)}$$

Where x is the input data, θ^{T} is the parameter, which is determined by the minimization of the cost function.

Decision Tree

The Decision Tree method is based on the application of a series of questions for the determination of the model response. The model generates a flowchart (tree), where each internal node (represented by a question) tests some features and guides down through the branches (the result of the splitting) with a 'gini' coefficient, which is defined as follows (Zhou et al. 2019):

$$G = \sum_{i=1}^{c} p(i)x(1-p(i)) \text{ (Eq. 3.4)}$$

The parameter c designates the number of total classes; p(i) is the probability of picking a datapoint with class i.

[1]https://www.kaggle.com

Random Forest

The Random Forest uses a combination of decision trees to aggregate the answer. In this method, a random training set is selected for every decision tree and a bootstrap set is chosen too (da Cruz et al. 2020). This algorithm can overcome the drawbacks of the regular decision tree.

3.4.2 Unsupervised methods

The unsupervised techniques include the Hierarchical classification and a combination of the Principal Component Analysis (PCA) method and the K-means method.

Hierarchical classification

The Hierarchical classification method is used for clustering unlabeled data. It is based on the variance minimization algorithm. The Ward method is used in this analysis.

PCA and k-means

The PCA method is used to reduce the dimension of the input data by focusing on the principle components. The K-means algorithm aims at partitioning n observations into k clusters. Initially, K initial means are randomly generated. Then, K clusters are created by associating each observation with the nearest centroid. The objective function, sum of the distance, is optimized until the best cluster centers candidates are found.

3.4.3 Artificial neural network (ANN)

The Artificial Neural Network (ANN) is inspired from the human brain functioning (Soldevilaet al. 2016, Romano et al. 2014). It transforms thee input data (input layer) through a series of neural layers (hidden layers)to output data (output layer). The transformation is based on the use of weights, which are adjusted by the optimization of prediction of a training data set. The Sigmoid function is used in data transformation.

3.5 Results and discussion

3.5.1 Supervised methods

The training phase of the supervised methods is conducted with 80% of the data, while 20% of data are used for the testing phase. The performance of the classification methods is investigated using the parameters Accuracy, Precision, Recall and F1-score, which are determined from the confusion matrix (Table 3.5) as follows:

$$Precision = \frac{Truepositive}{Truepositive + Falsepositive} \qquad \text{(Eq. 3.5)}$$

$$Recall = \frac{Truepositive}{Truepositive + Falsenegative} \quad \text{(Eq. 3.6)}$$

$$F1 - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad \text{(Eq. 3.7)}$$

$$Accuracy = \frac{Truepositive + Truenegative}{Truepositive + Falsepositive + Truenegative + Falsenegative} \quad \text{(Eq. 3.8)}$$

Table 3.5: Confusion matrix

	Prediction		
		Positive	**Negative**
Actual	**Positive**	True Positive	False Negative
	Negative	False positive	True Negative

3.5.1.1 Use of the water supply data

The Logistic Regression, Decision Tree and Random Forest methods were first used with the water supply data. Table3.6 summarizes the classifications of the use of these methods. It could be observed that both the Logistic Regression and Random Forest give excellent results with Accuracy =1.0, Precision =1.0, Recall= 1.0. F1-score = 1. The Decision Tree method also gives very good results with Accuracy = 0.95, Precision = 0.96; Recall= 0.95, and F1-score = 0.95.

Figure 3.5 shows the confusion matrix of the Decision Tree method. It indicates excellent performances for all zones, except for the zone 3 (Precision = 0.78) and zone 4 (Recall = 0.75).

Figure 3.6 and 3.7 show the architectures of the Decision Tree and the Random Forest methods, respectively.

Table 3.6: Classification report for the supervised methods - Flow data

Method	Accuracy	Precision	Recall	F1-score
Logistic Regression	1.0	1.0	1.0	1.0
Decision Tree	0.95	0.96	0.95	0.95
Random Forest	1.0	1.0	1.0	1.0

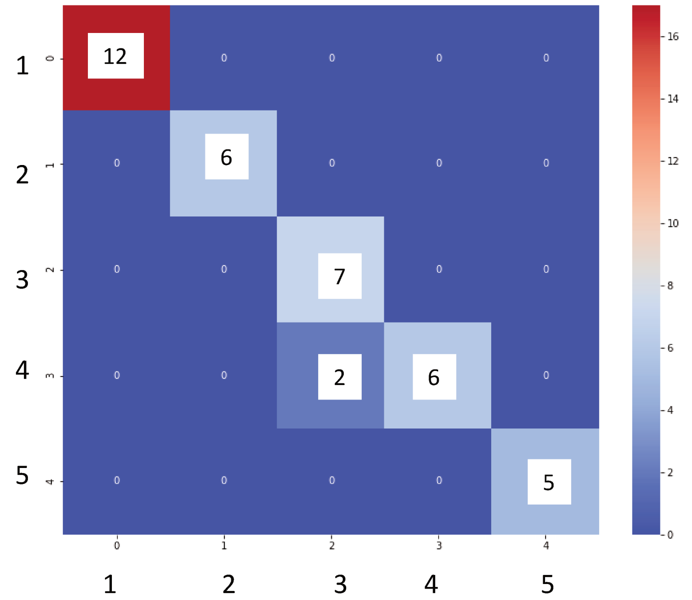


Figure3.5: Confusion matrix for the Decision Tree method - Flow Data

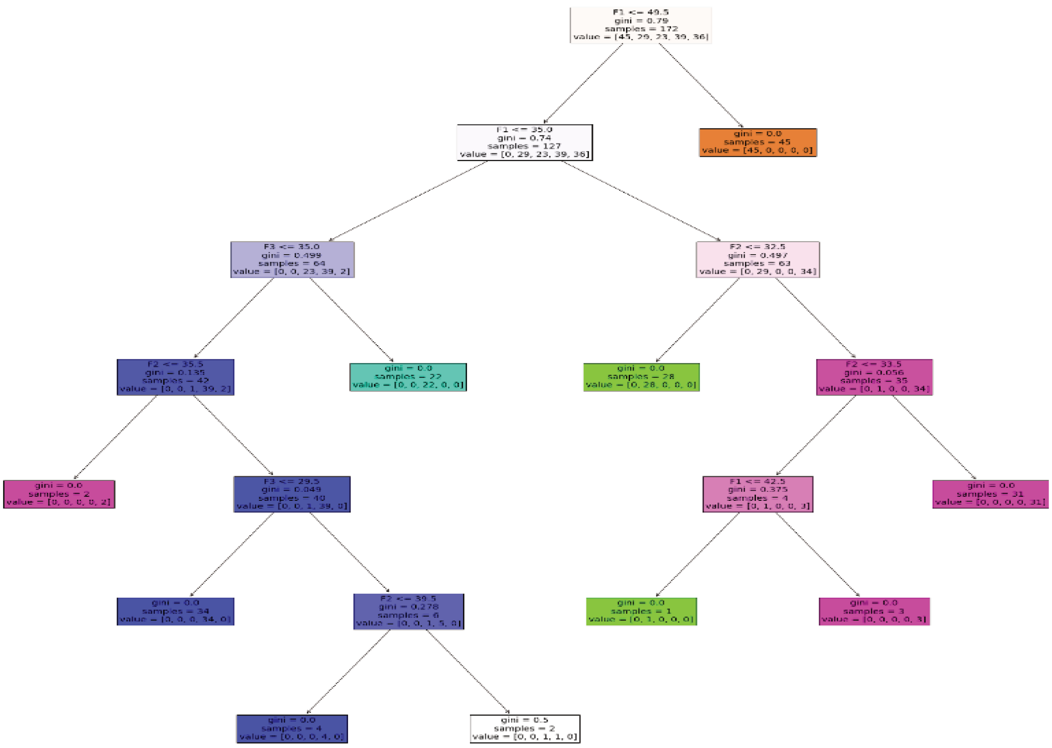


Figure 3.6 : Decision tree architecture - Flow data

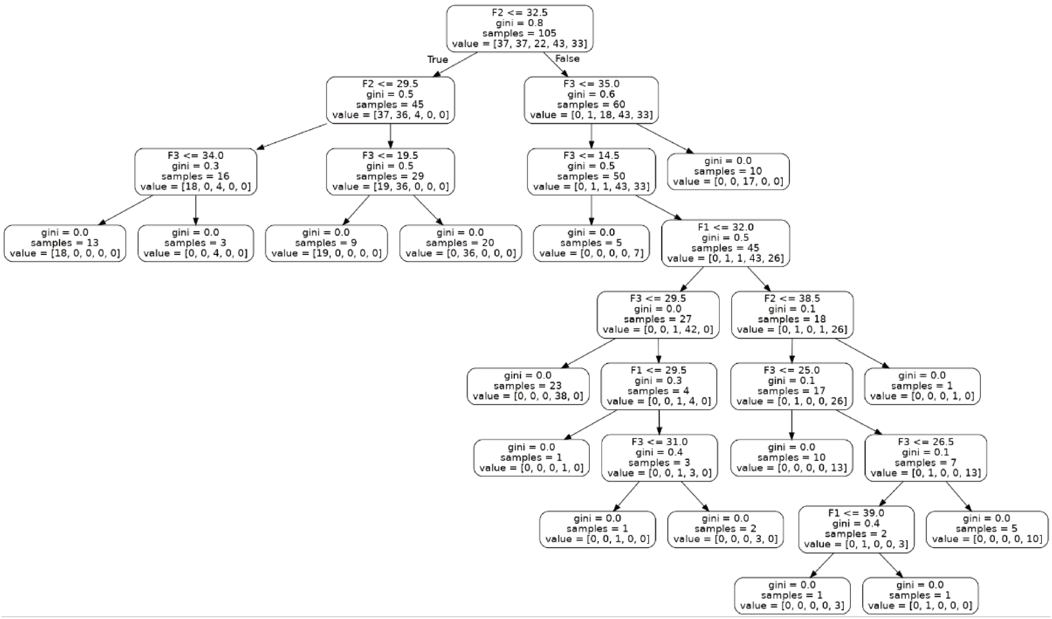


Figure3.7 : Random Forest architecture - Flow data

3.5.1.2 Use of the pressure data

Table 3.7 summarizes the classifications reports for the calibration of the supervised methods on the pressure data. It could be observed that both the Logistic Regression and Random Forest give excellent resultswith Accuracy = 1.0, Precision = 1.0, Recall =1.0. The Decision Tree method gives good results with Accuracy = 0.88, Precision = 0.91, Recall= 0.94, and F1-score = 0.91. Figure 3.8 shows the confusion matrix for the Decision Tree method. It indicates excellent performances for all the zones, except for the zone 1 (Recall = 0.70) and zone 2 (Precision = 0.54). Figure 3.9 and 3.10 show the architectures of the Decision Tree and the Random Forest methods, respectively.

Table 3.7: Classification report for the Supervised methods - Pressure data

Method	Accuracy	Precision	Recall	F1-score
Logistic Regression	1.0	1.0	1.0	1.0
Decision Tree	0.88	0.91	0.94	0.91
Random Forest	1.0	1.0	1.0	1.0

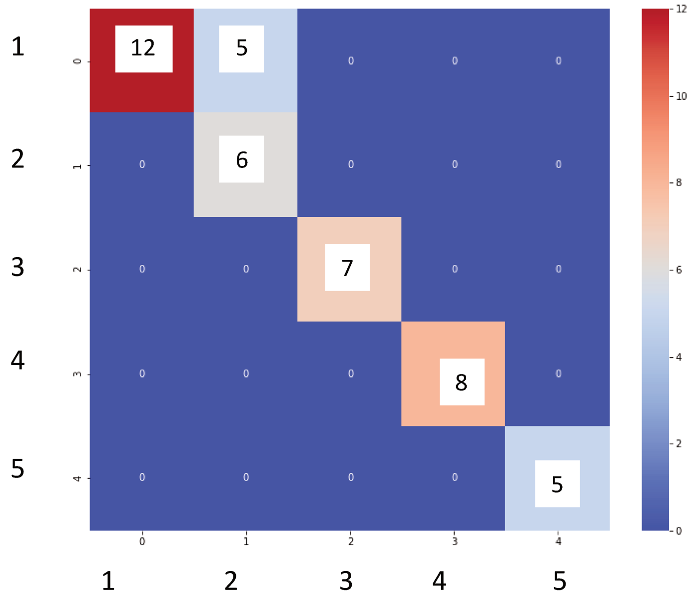


Figure 3.8: Confusion matrix for the Decision Tree method - Pressure Data

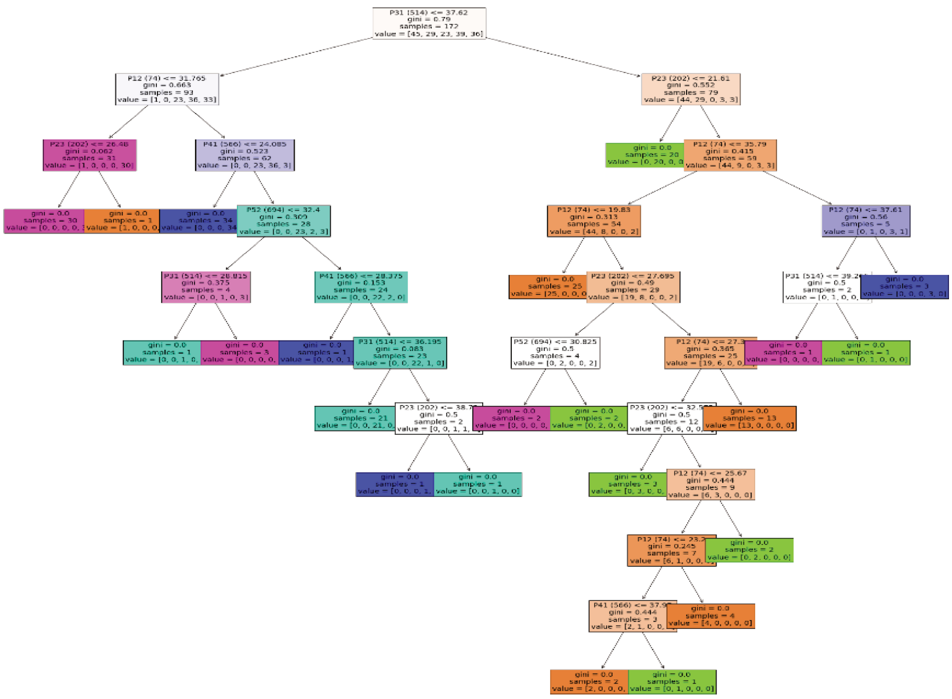


Figure 3.9 : Decision tree architecture - Pressure data

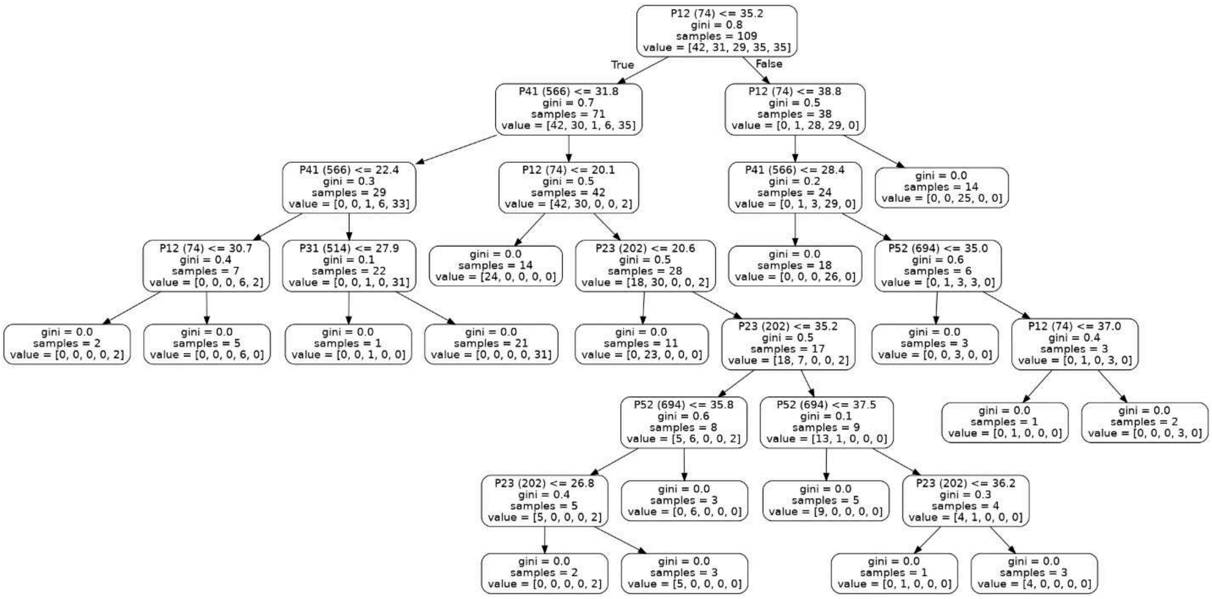


Figure 3.10 : Random Forest architecture - Pressure data

3.5.1.3 Use of the pressure and flow data

Pressure and flow data were used with the Decision Tree method, because the Logistic Regression and Random Forest methods gave already excellent results with either the flow data or the pressure data. Table3.8 provides the classifications report for the Decision Tree. It shows that this method gives excellent results with an accuracy of 0.98, precision of 0.97, a Recall of 0.97, and F1-score of 0.96. It could be observed that the performance obtained with the flow and pressure data is better than that obtained with the flow data (Table 3.6) and pressure data (Table 3.7).

Figure 3.11 shows the confusion matrix for the Decision Tree Method. It indicates excellent performances for all the zones, except for the zone 2 (Recall = 0.83) and zone 5 (Precision = 0.83). Figure 3.12 shows the architectures of the Decision Tree method.

Table 3.8: Classification report for Decision Tree Method - Flow and pressure data

Method	Accuracy	Precision	Recall	F1-score
Decision Tree	0.98	0.97	0.97	0.96

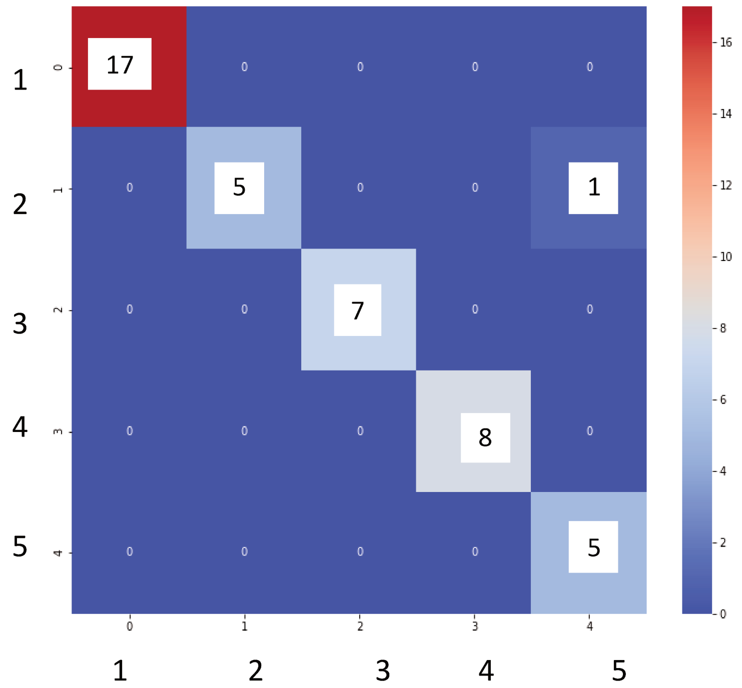


Figure 3.11: Confusion matrix for the Decision Tree method - Flow and pressure Data

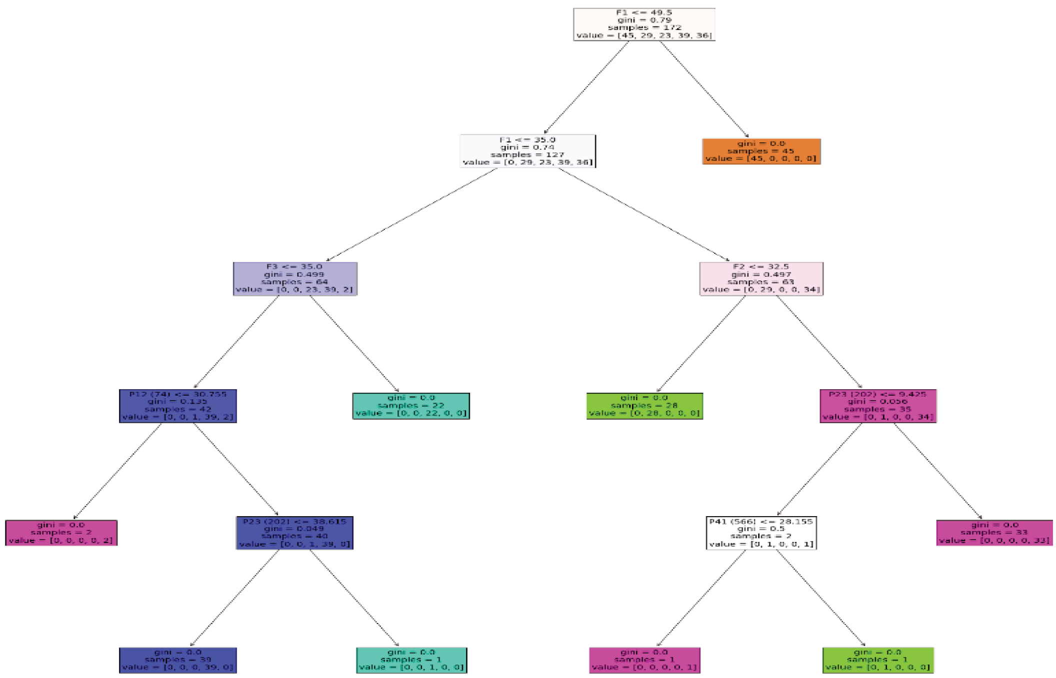


Figure 3.12 : Decision Tree architecture - Flow and pressure data

3.5.2 *Unsupervised methods*

This section presents the use of two methods of the unsupervised ML methods: The Hierarchical Classification method and a combination of the PCA and K-means classification methods.

3.5.2.1 *Hierarchical Classification Method*

Figure 3.13 shows the results obtained with the Hierarchical Classification method with the pressure data. It shows the existence of three groups: The first group G1 is composed of the pressures in zones 3 and 4, the second group G2 concerns the pressure in zone 1, and the third group G3 includes the pressures in zones 2 and 5.

Figure 3.14 illustrates the classification with flow rate and pressure data. It shows the presence of 5 groups: G1 with the pressure in zone 4; G2 with the flow rate FL1 and the pressure in zone 3, G3 with the pressures in zones 2 and 5; G4 with the flow rate FL2, and G5 with the flow rate FL3 and the pressure in zone 1.

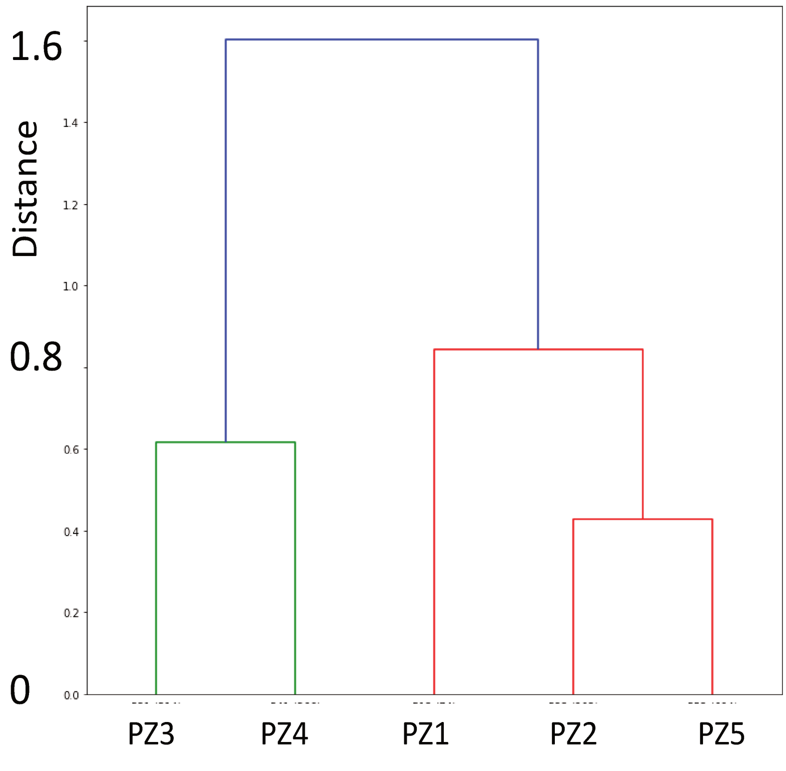


Figure 3.13: Hierarchical Classification method - results with the pressure data

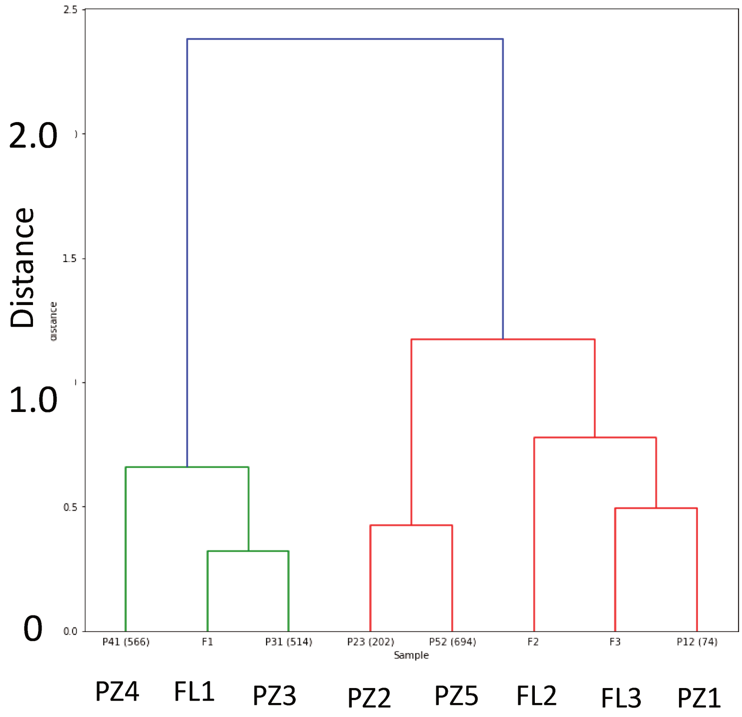


Figure 3.14: Hierarchical Classification method - results with the flow rate and pressure data

3.5.2.2 PCA and K-means

Figure 3.14 illustrates the results obtained with the application of the of the PCA and K-means methods on flow rate data for k = 5 clusters. The component PC1 shows three clusters and two partially overlapping clusters. The component PC2 indicates two clusters and three partially overlapping clusters. In the (PC1, PC2) plan, the 5 clusters could be well distinguished.

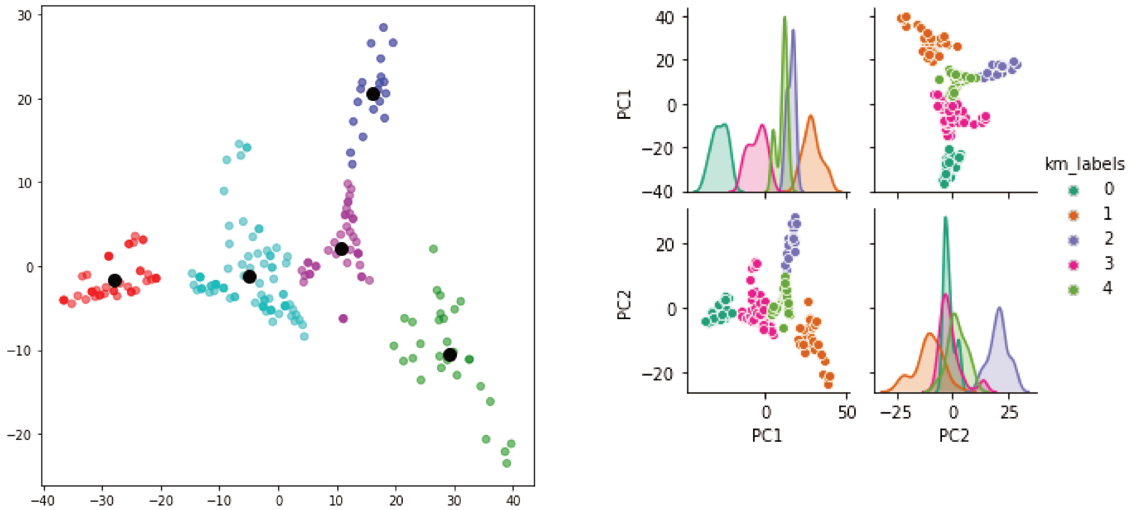


Figure 3.14: PCA and K-means clustering - Flow rate data

Figure 3.15 illustrates the results obtained with the pressure data for k = 5 clusters. Both PC1 and PC2 show important clusters' overlapping. In the (PC1, PC2) plan, the 5 clusters are not well distinguished.

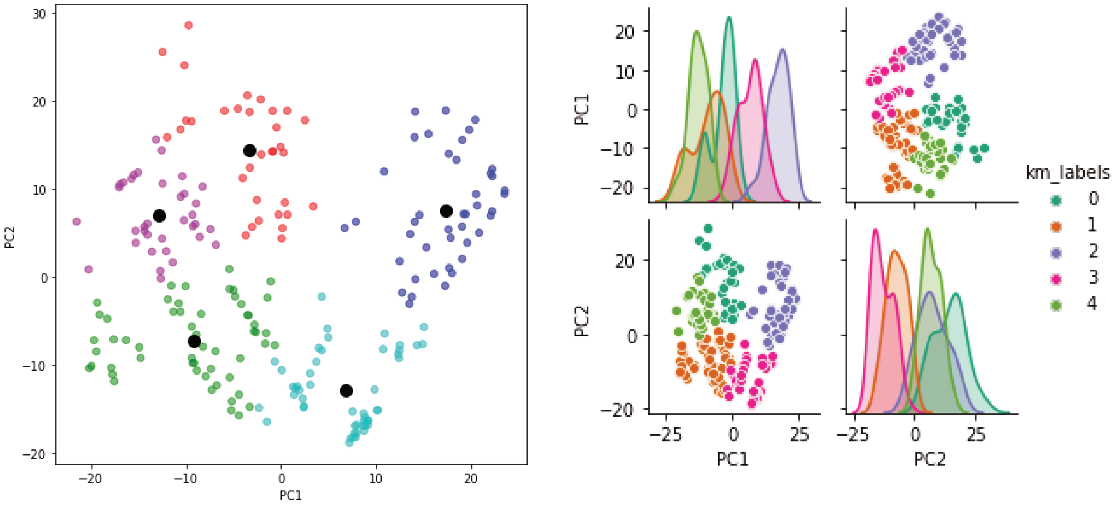


Figure 3.15: PCA and K-means clustering - Pressure data

Figure 3.16 illustrates the results obtained with flow rate and pressure data for k = 5 clusters. Both PC1 and PC2 show clusters' overlapping. In the (PC1, PC2) plan; only 3 clusters could be well distinguished.

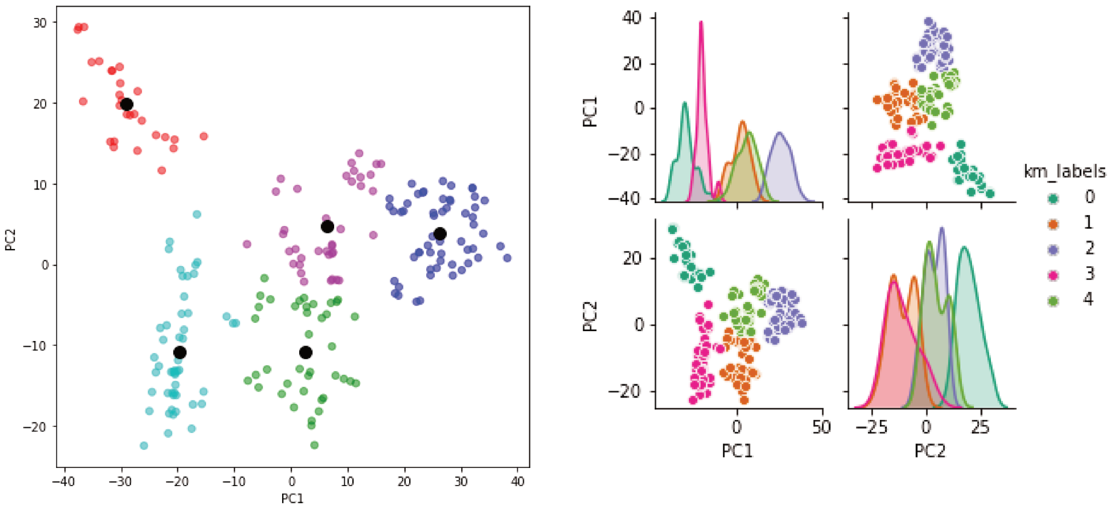
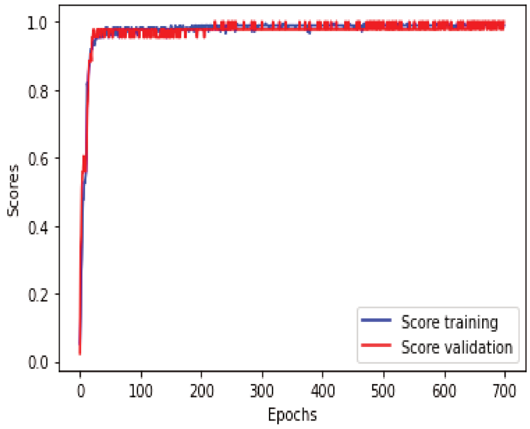


Figure 3.16: PCA and K-means clustering - Flow rate and pressure data

3.5.3 Artificial Neural Network

Figure 3.17 shows the results of the application of the ANN method with water supply data. It indicates a rapid convergence of the ANN model. Indeed, a good convergence is observed with around 20 epochs. The ANN model gives excellent results with Accuracy = 1.0, Precision = 1.0, Recall = 1.0, and F1-Score = 1.0.



Accuracy = 1.0

Precision = 1.0

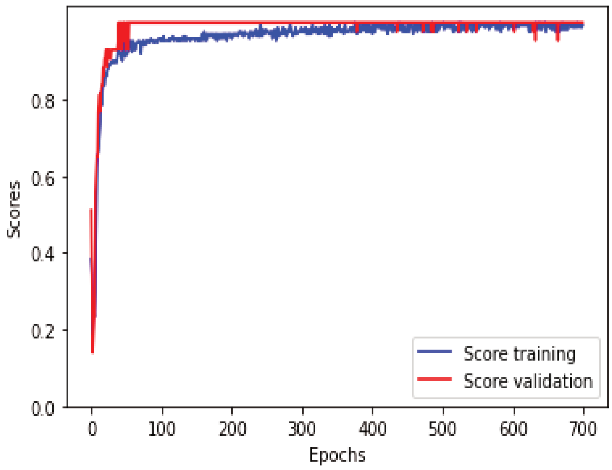
Recall = 1.0

F1-Score = 1.0

Figure 3.17 : Application of the ANN method with flow rate data

Figure 3.18 illustrates the results of the application of the ANN method with the pressure data. The convergence of the training phase is obtained with around150 epochs, while the convergence of the validation phase is achieved with around 50 epochs. The model gives excellent results with an Accuracy = 1.0, Precision = 1.0, Recall = 1.0, and F1-Score = 1.0.

`



Accuracy = 1.0

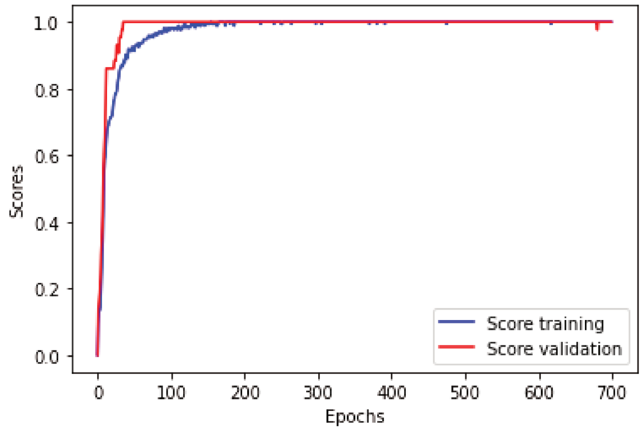
Precision = 1.0

Recall = 1.0

F1-Score = 1.0

Figure 3.18: Application of the ANN method with pressure data

Figure 3.19 illustrates the results of the application of the ANN method with water supply and pressure data. It indicates a convergence of the training phase with around 100 epochs anda convergence of the validation phase with around 50 epochs. The model gives excellent results with Accuracy = 1.0, Precision = 1.0, Recall = 1.0, and F1-Score = 1.0.



Accuracy = 1.0

Precision = 1.0

Recall = 1.0

F1-Score = 1.0

Figure 3.19: Application of the ANN method with flow pressure data

3.6 Analysis of the water leak in the scientific campus of Lille University

This section presents analysis of the water consumption and leak in the scientific campus of Lille University. Analysis is based on data collected in 2015at hourly time interval from the three supply sections: FL1 in the North, FL2 in the west and FL3 in the South of the Campus. The following section present successively (i) analysis of the daily water consumption, (ii) analysis of the hourly water consumption (Qh), (iii) comparison of water consumption in the working days and the weekend, (iv) analysis of the night water consumption and determination of permanent leaks, and (v) leak analysis.

3.6.1 Analysis of the daily water consumption (Qd)

Figure 3.20 and 3.21 illustrate the variation of the daily water consumption of the campus (Qd). They indicate missing data in the period May 3 to May 27. This period will not be considered in the analysis. These figures show an important variation in Qd. The minimum daily consumption is equal to 414 m^3, while the maximum is equal to 1680 m^3, and the average is equal to 890 m^3. Low consumption values could be attributed to the vacation periods, while the high consumption values could be associated to water leaks in the water network of the campus.

Figure 3.22 and 3.23illustrate the repartition of the daily water supply among the three supply sections. They show that the daily water supply from the North (F1D) is higher than those from the west and the South of the campus. It also has the largest variation (Table 3.9): the minimum daily supply is equal to 100 m^3, while the maximum is equal to 772 m^3 and the average is equal to 442 m^3, to be compared with F2D (Resp F3D): minimum = 197 m^3 (Resp. 51 m^3), maximum = 772 m^3 (Resp. 454 m^3) and average = 251 m^3(Resp. 197 m^3). The water supply F1D accounts for 50% of the total water supply, while F2 accounts for 28% and F3 for 22% of the water supply of the campus.

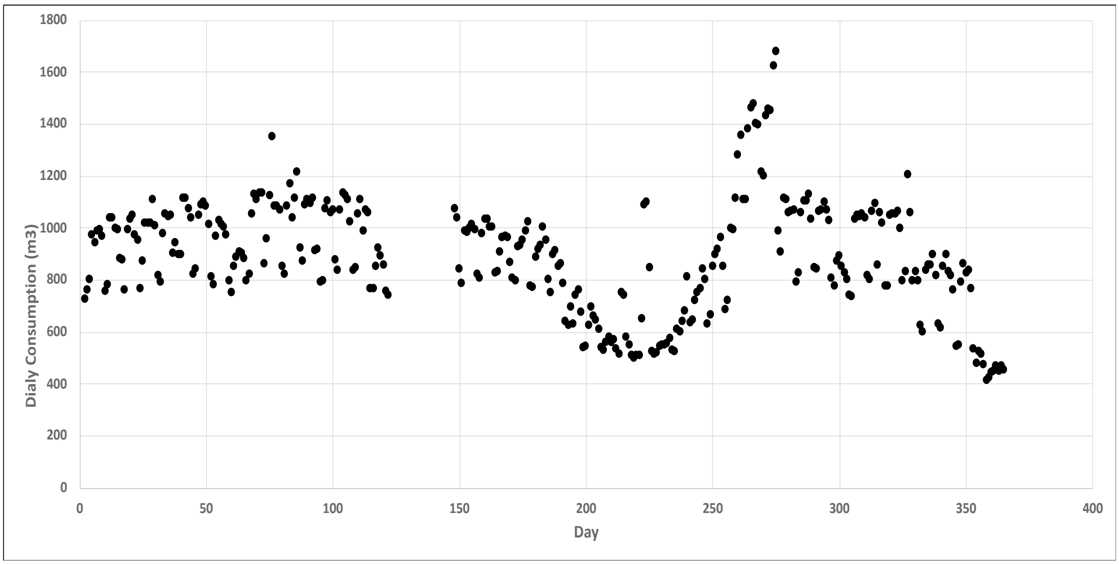


Figure 3.20: Daily water consumption (Qd) of the Scientific Campus in 2015

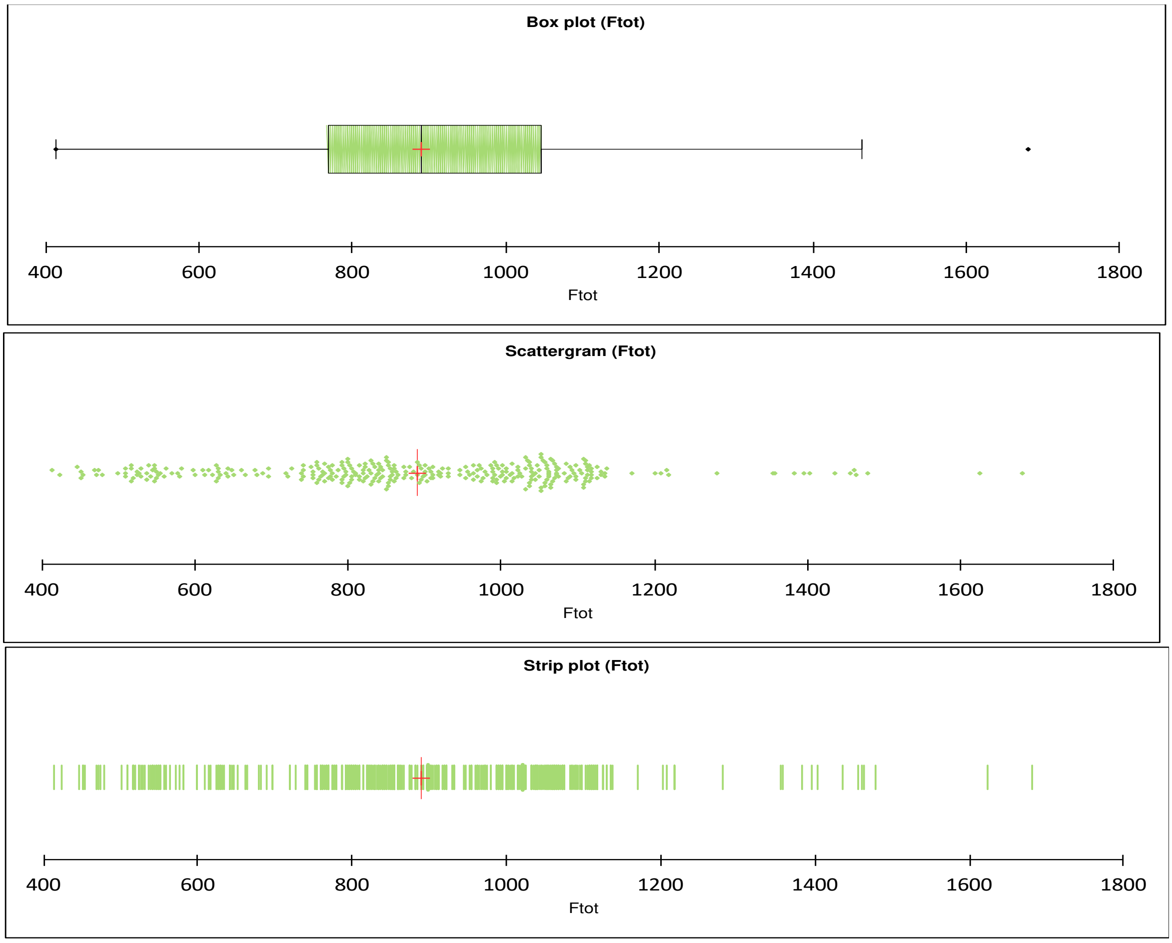


Figure 3.21: Distribution of the daily water consumption of the scientific campus in 2015.

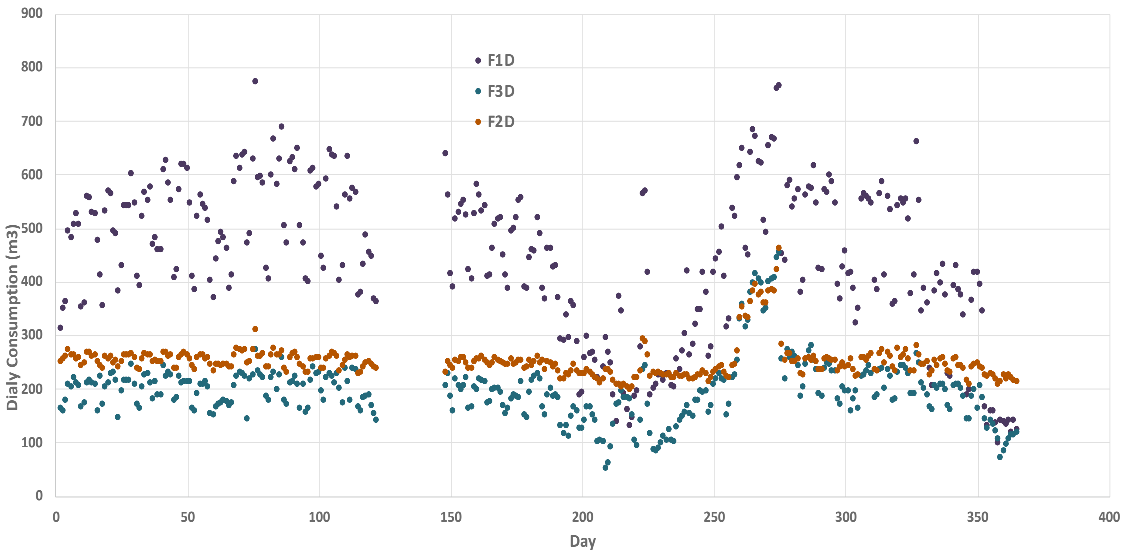


Figure 3.22: Repartition of the daily water supply of the scientific campus in 2015

(F1: North, F2: west, F3: South)

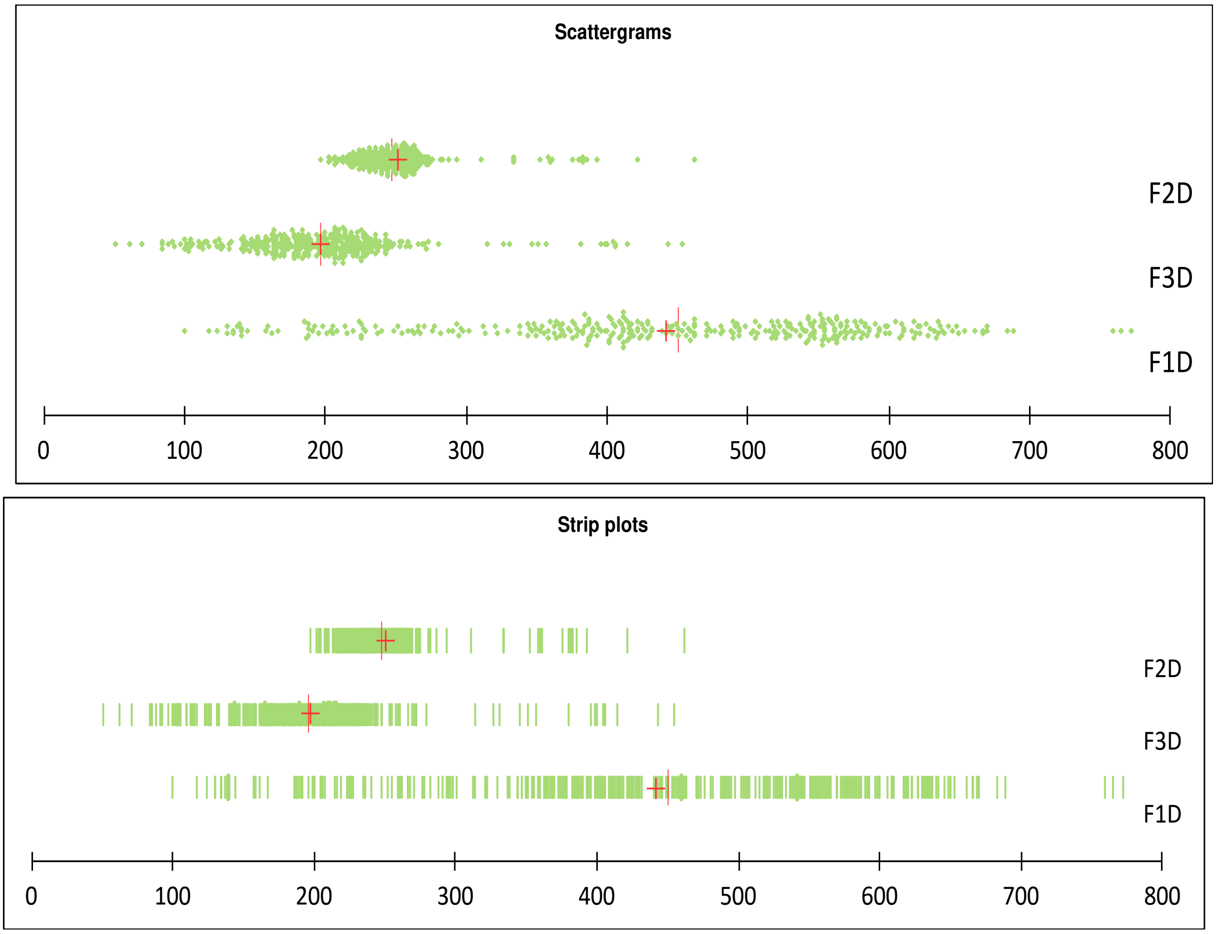


Figure 3.23: Repartition of the daily water supply of the scientific campus in 2015.

(F1: North, F2: west, F3: South)

Table 3.9: Statistical descriptive analysis of the daily water supply of the campus.

	F1D m^3/day	F2D m^3/day	F3D m^3/day	Total (Qd) m^3/day
Minimum	100	197	51	414
Maximum	772	462	454	1680
Average	442	251	197	890
Standard deviation	143	33	59	219

3.6.2 Analysis of the hourly water consumption (Qh)

Figure 3.24and table 3.10 show the distribution of the hourly consumption (Qh) in the campus. They indicate an important variation. F1 varies between zero and 81 m^3, with an average value of 18 m^3. F2 varies between zero and 30 m^3 with an average value of 10 m^3. F3 varies between zero and 31 m^3 with an average value of 8 m^3.

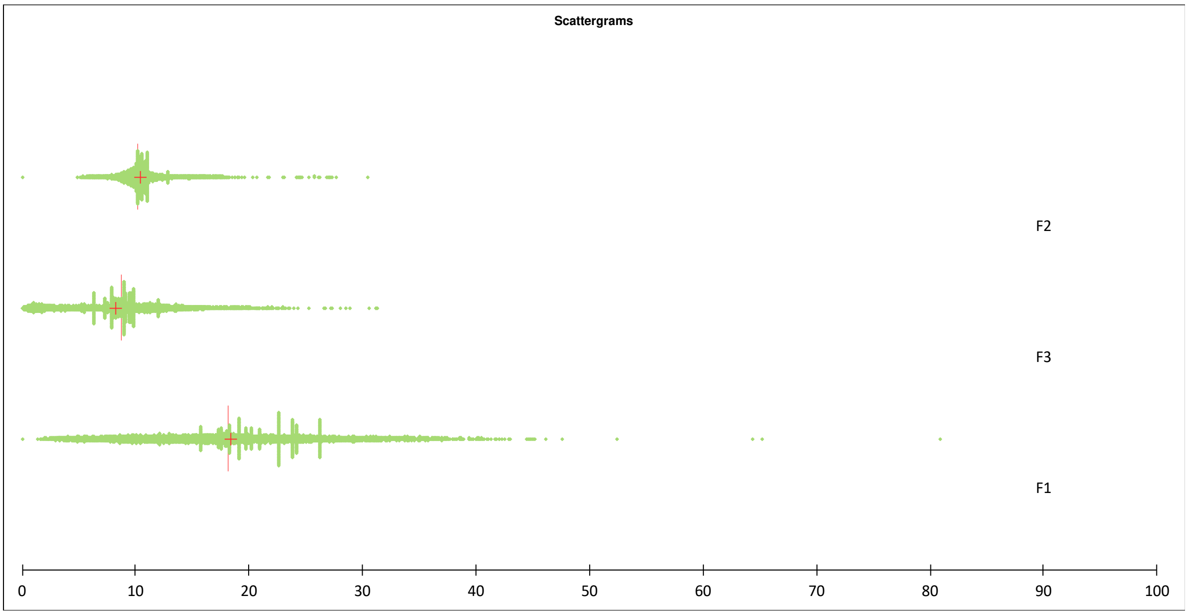


Figure 3.24 Repartition of the hourly consumption of the scientific campus in 2015.

(F1: North, F2west, F3: South)

Table 3.10 : Statistical analysis of the hourly water consumption (Qh)

	F1 m^3/hour	F2 m^3/hour	F3 m^3/hour
Minimum	0.00	0.00	0.00
Maximum	81.4	30.5	31,3
Average	18.4	10.5	8.2
Standard Deviation	59.7	3.4	20.7

3.6.3 Water consumptions in the working days and the weekend

Analysis was conducted to explore the variation of the water consumption between the working days and the weekend. Figure 3.25 shows a comparison of the mean values of the consumption in the working days of each week to that of the weekend of the same week. We observe that the water consumption in the working days is about 20% higher than that in the weekend. This result is confirmed in figure 3.26, which shows the variation of the ratio between the water consumption in the working days to that in the weekend. The high values of this ratio in weeks 36 and 37 are related to the water leak during these weeks.

Table 3.11 shows a statistical analysis of the ratio of the water supply during the working days to those in the weekend. It shows that the average value of the ratio related to the total supply is equal to 1.22, while that related to F1 (North), F2 (west), and F3 (South) are equal to 1.33, 1.07 and 1.23. These results show that the water supply from the west of the campus is not impacted by the campus activity during the working days.

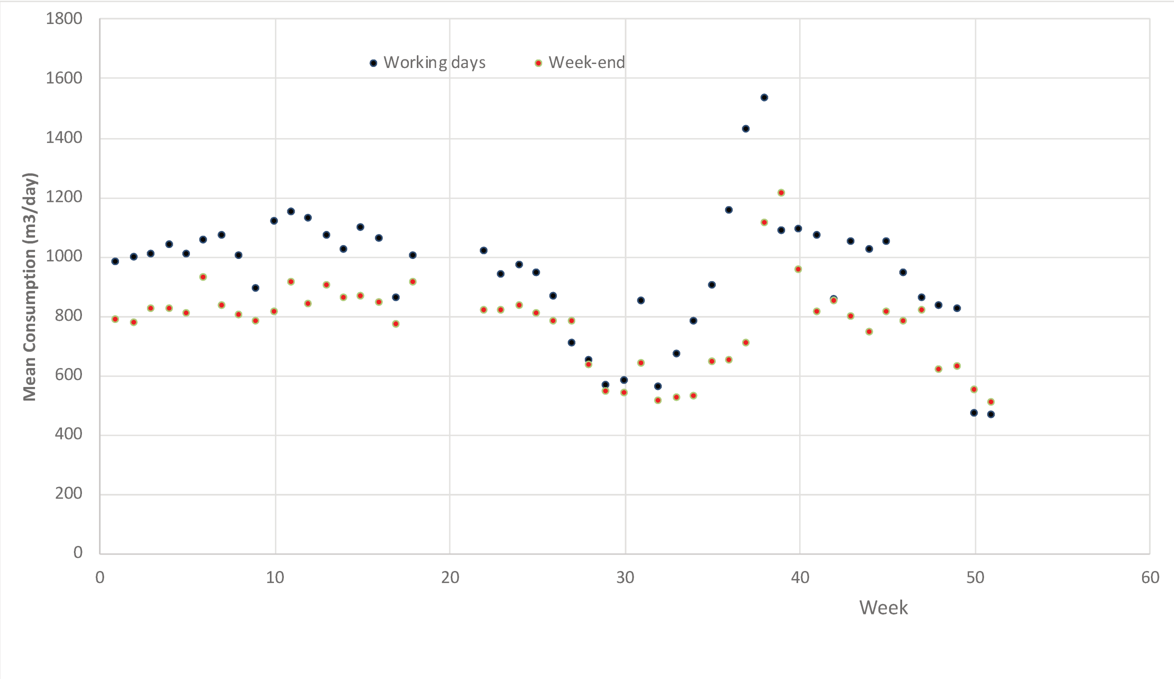


Figure 3.25Comparison of the in the working days and the weekend

Table 3.11: Statistic analysis of the ratio water supply in the working days to that in the weekend

	Total	F1	F2	F3
Average	1.22	1.33	1.07	1.23
Minimum	0.85	0.71	0.71	0.69
Maximum	2.02	2.17	1.70	2.49

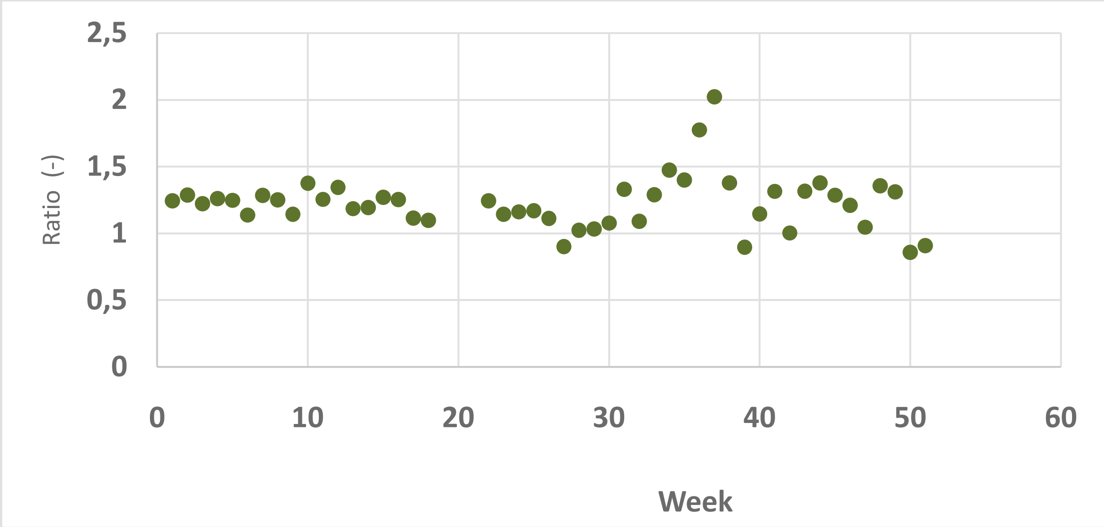


Figure 3.26 Ratio between water consumption in the working days to that in the weekend

3.6.4 Analysis of the night water consumption - determination of permanent leak

Figure 3.27 shows the variation of the night hourly consumption (Qn) in the campus (around 4:00 am). It indicates a significant consumption with an average value of 26 m^3/hour. Since the activity of the campus at 4:00 am is almost stopped, in particular for the water usage, this consumption could be attributed to permanent leak. This leak could be related to small leaks in the water network and in the buildings. If we assume that this leak is continuous over the day, we can evaluate the daily permanent water leak (Qpl = 24*Qn) and determine its percentage of the global daily consumption (Qd). Figure 3.28 shows this percentage varies between zero and 69% with an average of 28%. This result highlights the importance of the permanent leak in the campus and the necessity to control the water network and the water system in buildings to reduce this leak.

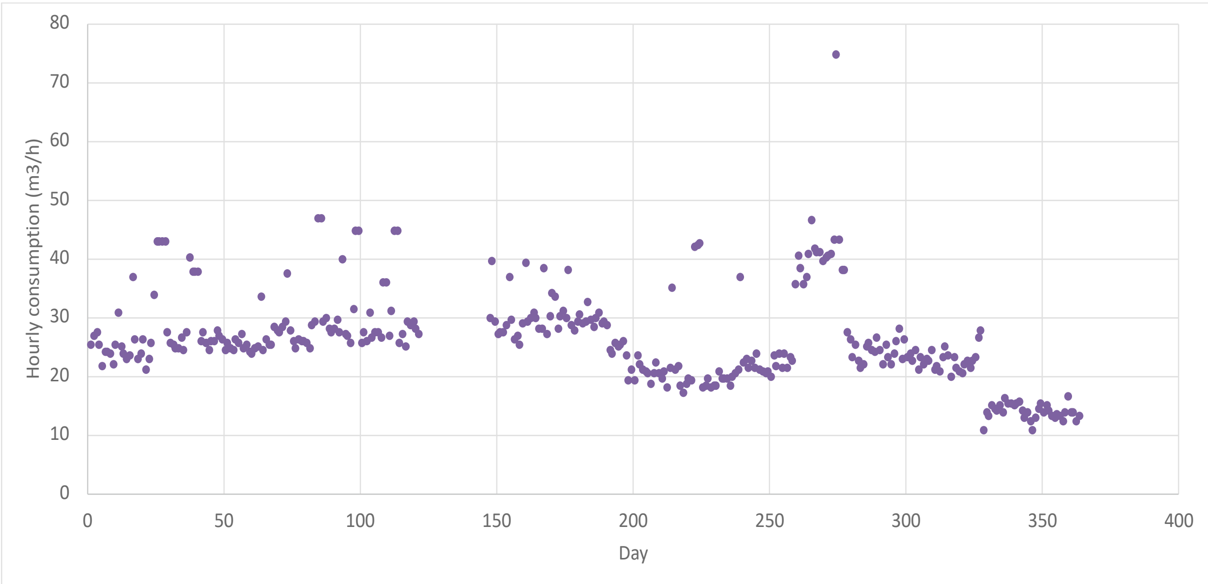


Figure 3.27: Night consumption (Qn) of the scientific campus in 2015.

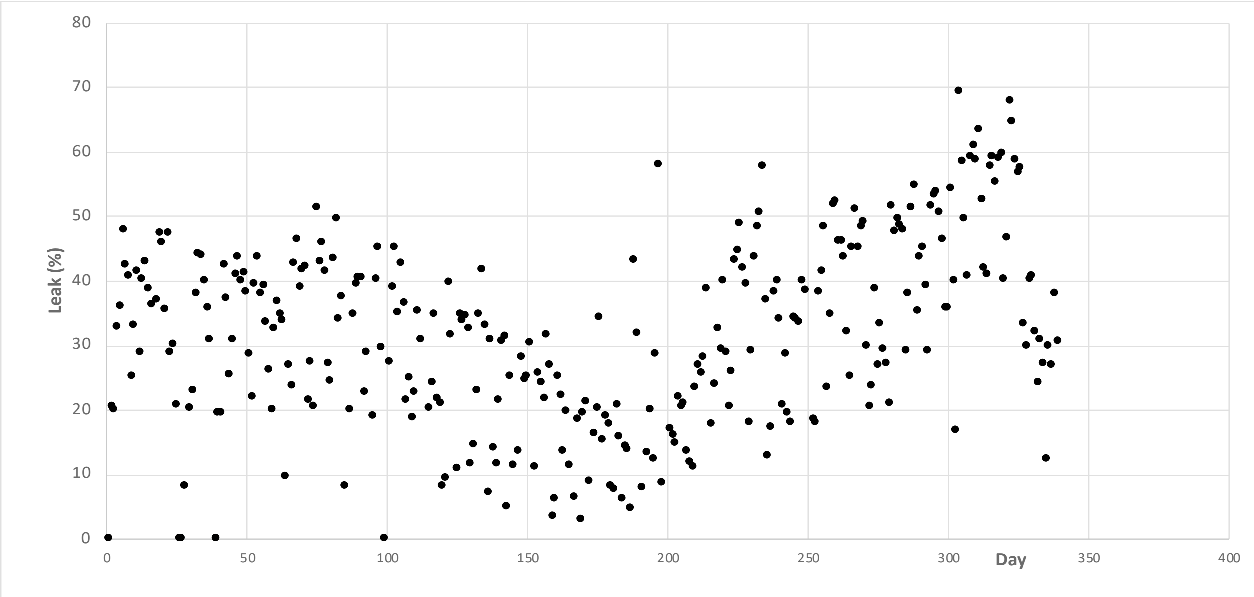


Figure 3.28 Percentage of the permanent water leak Qpl/Qd (%)

3.6.5 Leak analysis

The estimation of leak events is based on the identification of abnormal water consumptions. Figure 3.29 and 3.30 show the events with a water consumption exceeding 1200 m^3/day (Average water consumption + 1.5 standard deviation). We observe the presence of five groups of events, which are summarized in table 3.12.The first group (G1) corresponds to the day 76 with a consumption exceeding by about 464 m^3 the average water consumption (Qav), followed by the day 86 (G2),which exceeds Qav by 326 m^3. The third group corresponds to days 260 and 261, with water consumption exceeding Qav by 390 and 467 m^3, respectively. The fourth group is related to days 264 - 275, with water consumption exceeding Qav by values included in the interval 311 - 790 $m^{3.}$The last leak (G5) occurred day 327 with a consumption exceeding Qav by 318 m^3.

Figure 3.31 and table 3.13 show the repartition of the water supply ratios related to leak events. It shows that the ratio related to F1 is higher than those related toF2 and F3. F1 water supply accounts for 56% of the total water supply for groups G1, G2 and G5, while F2 and F3 account for around 22 % each. For groups G3 and G4, F1 accounts for around 46% of the campus water supply, while F2 and F3 account for around 27% each.

For the localization of leaks events G1 to G5, the water supply ratios corresponding to the leak events are reported in figure 3.32. It could be observed that leaks G1, G2 and G5 well match with leaks in zone 1 of the campus, while leaks G3 and G4 well match with leaks in zone 2.

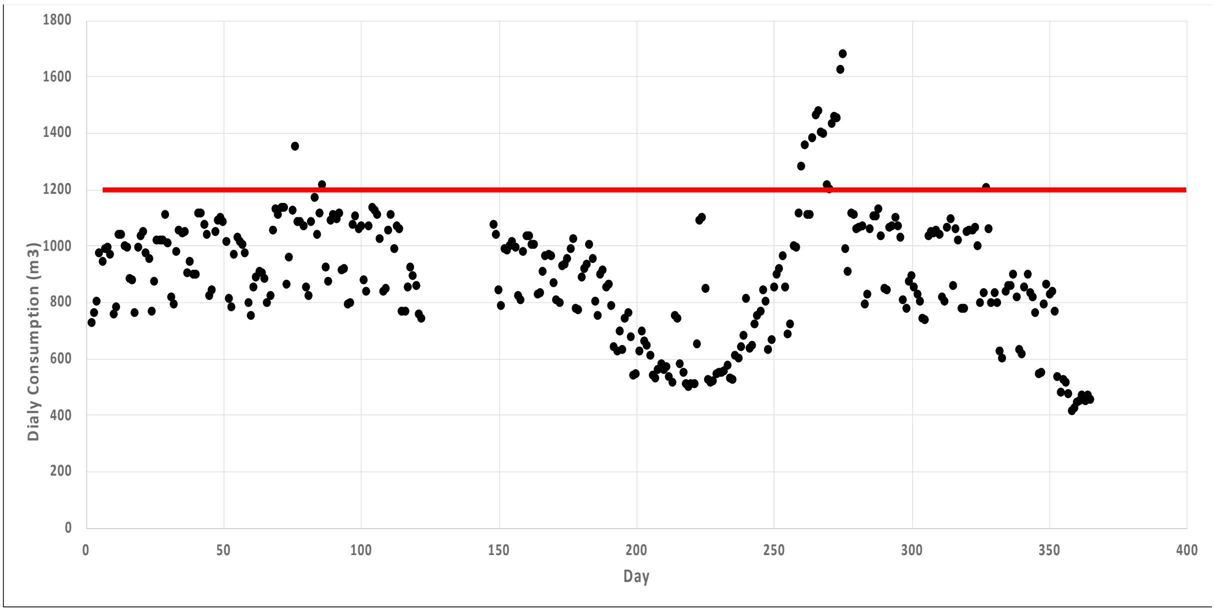


Figure 3.29: Identification of leak events in the water distribution system of the scientific campus

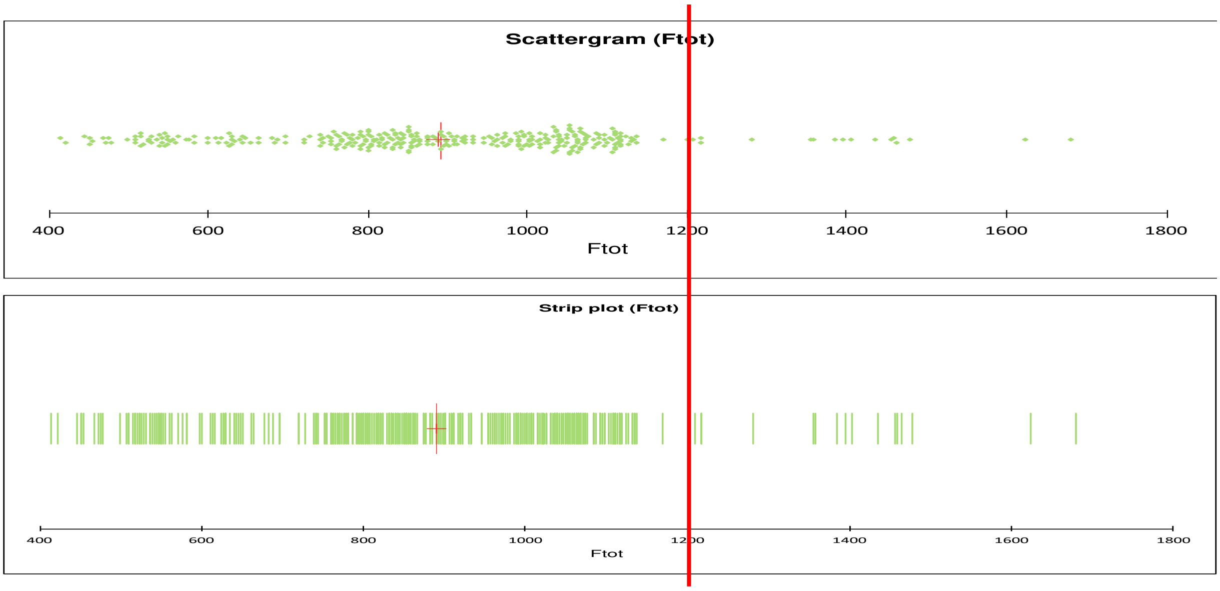
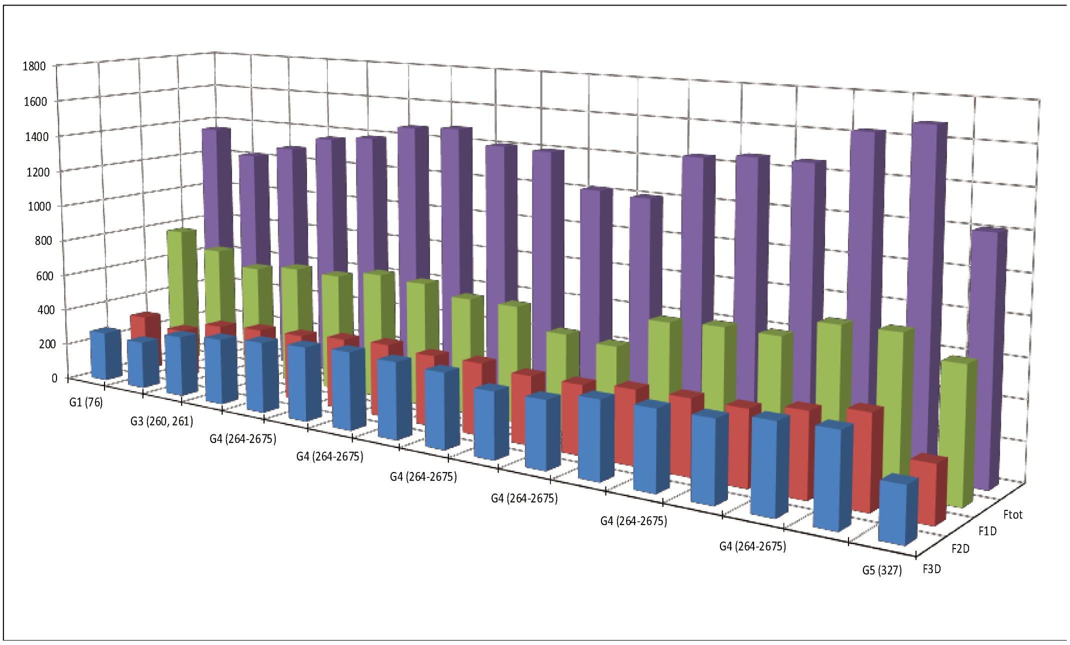


Figure 3.30: Identification of leak events

Table 3.12 : Leak events in the water distribution of the scientific campus

Day	**Group**	**Qd (m^3/day)**	**Qd - Qaverage (m^3/day)**
76	G1 (76)	1354	464
86	G2 (86)	1216	326
260	G3 (260. 261)	1280	390
261	G3 (260. 261)	1357	467
264	G4 (264-2675)	1383	493
265	G4 (264-2675)	1463	573
266	G4 (264-2675)	1477	587
267	G4 (264-2675)	1404	514
268	G4 (264-2675)	1396	506
269	G4 (264-2675)	1217	327
270	G4 (264-2675)	1201	311
271	G4 (264-2675)	1435	545
272	G4 (264-2675)	1459	569
273	G4 (264-2675)	1455	565
274	G4 (264-2675)	1624	734
275	G4 (264-2675)	1680	790
327	G5 (327)	1208	318



3.31:Repartition of the water supply ratios related to leak events

Table 3.13 Repartition of the water supply ratios related to leak events

Day	Groupe	FL3 (%)	FL2 (%)	FL1 (%)
76	G1 (76)	20	23	57
86	G2 (86)	21	22	57
260	G3 (260, 261)	26	26	48
261	G3 (260, 261)	26	26	48
264	G4 (264-2675)	28	26	46
265	G4 (264-2675)	27	26	47
266	G4 (264-2675)	28	27	45
267	G4 (264-2675)	29	27	44
268	G4 (264-2675)	28	27	44
269	G4 (264-2675)	28	29	42
270	G4 (264-2675)	29	30	41
271	G4 (264-2675)	28	27	46
272	G4 (264-2675)	28	26	46
273	G4 (264-2675)	28	26	46
274	G4 (264-2675)	27	26	47
275	G4 (264-2675)	27	27	45
327	G5 (327)	22	23	55

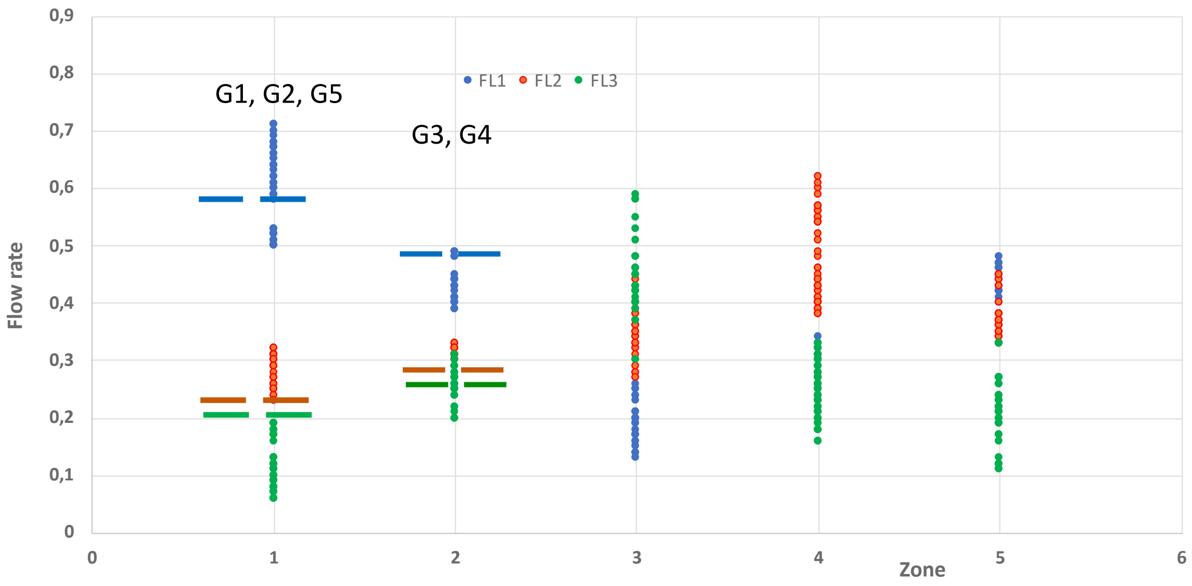


Figure 3.32: Localization of the week events (G1 to G5) in the campus

3.7 Conclusion

This chapter presented the use of the Machine Learning techniques for the localization of leaks in the water distribution systems. This issue is very important because the localization of water leaks in urban area is very complex, due to the high density of constructions and the complexity of urban water networks. The proposed methodology is based on the creation of zones in each city. For each zone, sensors are used to measure the water supply of the zone as well as the water pressure variation in different locations of the zone. Data of water supply and pressure variation are then used for the construction of the Machine Learning Models.

The scientific campus of Lille University was used as a support for this research. This campus includes 140 buildings with around 25 000 users. The water company ensures the water supply of the campus through three supply sections (North, West and South). The water network was subdivided in five hydraulic zones. The EPANET software was then used for the determination of the water supply and pressure data resulting from 215 leak events in the campus. These data were then used for training and testing six Machine Learning techniques:

- Three supervised methods: Logistic Regression, Decision Tree and Random Forest
- Two unsupervised methods: The Hierarchical Classification method and a combination of the PCA and K-means classification method.
- The Artificial Neural Network

The application of these methods on the 215 leak events generated by EPNAT showed (i) excellent performance of the supervised methods, in particular the Logistic Regression and Random Forest (ii) excellent performances of the Artificial Neural Network (iii) difficulties in the exploitation of the clustering capacity of the unsupervised methods in leak localization because of clusters' overlapping.

Real water supply data were then used for the analysis of the water leak in the campus. This analysis showed an important permanent leak in the campus as well as some leak events. The results of the leak localization were then used for the determination of the zones corresponding to leak events. This work was based only on water supply data. The use of pressure data could improve the precision of leak localization in the campus.

3.8 References

Caputo, A. C., &Pelagagge, P. M. (2003). Using Neural Networks to Monitor Piping Systems. Process Safety Progress, 22(2), 119–127. https://doi.org/https://doi.org/10.1002/prs.680220208

Chan, T. K., Chin, C. S., & Zhong, X. (2018). Review of Current Technologies and Proposed Intelligent Methodologies for Water Distributed Network Leakage Detection. IEEE Access, 6, 78846–78867. https://doi.org/10.1109/ACCESS.2018.2885444

Ciupke, K. (2018). Leak Detection Using Regression Trees. https://doi.org/10.1007/978-3-319-62042-8

Da Cruz R.P., da Silva F.F., Fileti A.M.F. Machine learning and acoustic method applied to leak detection and location in low-pressure gas pipelines. Clean Technologies and Environmental Policy (2020), doi:10.1007/s10098-019-01805-x.

Farah, E., & Shahrour, I. (2017). Leakage Detection Using Smart Water System: Combination of Water Balance and Automated Minimum Night Flow. Water Resources Management, 31(15), 4821–4833. https://doi.org/10.1007/s11269-017-1780-9

Farah, E., Abdallah, A., & Shahrour, I. (2017). SunRise: Large scale demonstrator of the smart water system. International Journal of Sustainable Development and Planning. https://doi.org/10.2495/SDP-V12-N1-112-121

Han, Q., Nguyen, P., Eguchi, R. T., Hsu, K. L., &Venkatasubramanian, N. (2017). Toward An Integrated Approach to Localizing Failures in Community Water Networks. Proceedings - International Conference on Distributed Computing Systems, 1250–1260. https://doi.org/10.1109/ICDCS.2017.81

Haxton, T., Janke, R., Murray, R., Burkhardt, J., &Grayman, W. (2019). EPANET: Future of Water Distribution Modeling and Data Analytics Tools. https://cfpub.epa.gov/si/si_public_file_download.cfm?p_download_id=539762&Lab=NHSRC

Mounce, S. R., Pedraza, C., Jackson, T., Linford, P., & Boxall, J. B. (2015). Cloud based machine learning approaches for leakage assessment and management in smart water networks. Procedia Engineering, 119(1), 43–52. https://doi.org/10.1016/j.proeng.2015.08.851

Mounce, Stephen R., Mounce, R. B., & Boxall, J. B. (2011). Novelty detection for time series data analysis in water distribution systems using support vector machines. Journal of Hydroinformatics, 13(4), 672–686. https://doi.org/10.2166/hydro.2010.144

Rojek, I., &Studzinski, J. (2019). Detection and localization of water leaks in water nets supported by an ICT system with artificial intelligence methods as a way forward for smart cities. Sustainability (Switzerland), 11(2). https://doi.org/10.3390/su11020518

Romano M., Kapelan Z., Savic D.A. (2014) Evolutionary algorithm and expectation maximization strategies for improved detection of pipe bursts and other events in water distribution systems J. Water Resour. Plan. Manag., 140 (5) (2014), pp. 572-584

Salam, A. E. U., Tola, M., Selintung, M., & Maricar, F. (2014). On-line monitoring system of water leakage detection in pipe networks with artificial intelligence. ARPN Journal of Engineering and Applied Sciences, 9(10), 1817–1822.

Shahrour, I., Abbas, O., Abdallah, A., AbouRjeily, Y., Afaneh, A., Aljer, A., Ayari, B., Farrah, E., Sakr, D., & Al Masri, F. (2017). Lessons from a Large-Scale Demonstrator of the Smart and Sustainable City BT - Happy City - How to Plan and Create the Best Livable Area for the People (A. Brdulak& H. Brdulak (eds.); pp. 193–206). Springer International Publishing. https://doi.org/10.1007/978-3-319-49899-7_11

Shravani, D., Prajwal, Y. R., Prapulla, S. B., Salanke, N. S. G. R., Shobha, G., & Ahmad, S. F. (2019). A Machine Learning Approach to Water Leak Localization. 2019 4th International

Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), 4, 1–6. https://doi.org/10.1109/CSITSS47250.2019.9031010

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

Soldevila, A., Fernandez-Canti, R. M., Blesa, J., Tornil-Sin, S., & Puig, V. (2017). Leak localization in water distribution networks using Bayesian classifiers. Journal of Process Control, 55, 1–9. https://doi.org/10.1016/j.jprocont.2017.03.015

van der Walt, J. C., Heyns, P. S., & Wilke, D. N. (2018). Pipe network leak detection: comparison between statistical and machine learning techniques. Urban Water Journal, 15(10), 953–960. https://doi.org/10.1080/1573062X.2019.1597375

Wu, Y., & Liu, S. (2017). A review of data-driven approaches for burst detection in water distribution systems. Urban Water Journal, 14(9), 972–983. https://doi.org/10.1080/1573062X.2017.1279191

Wu, Z. Y., El-Maghraby, M., & Pathak, S. (2015). Applications of deep learning for smart water networks. Procedia Engineering, 119(1), 479–485. https://doi.org/10.1016/j.proeng.2015.08.870

Yalçin, B. C., Demir, C., Gökçe, M., &Koyun, A. (2018). Water leakage detection for complex pipe systems using hybrid learning algorithm based on ANFIS Method. Journal of Computing and Information Science in Engineering, 18(4), 1–10. https://doi.org/10.1115/1.4040130

Zhang, Q., Wu, Z. Y., Zhao, M., Qi, J., Huang, Y., & Zhao, H. (2016). Leakage Zone Identification in Large-Scale Water Distribution Systems Using Multiclass Support Vector Machines. Journal of Water Resources Planning and Management, 142(11), 4016042. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000661

Zhou, X., Tang, Z., Xu, W., Meng, F., Chu, X., Xin, K., & Fu, G. (2019). Deep learning identifies accurate burst locations in water distribution networks. Water Research, 166, 115058. https://doi.org/10.1016/j.watres.2019.115058

4 General Conclusion

This research concerned the detection and localization of leaks in urban water distribution networks. This issue is of major concern in the management of the water distribution systems because leaks in the water distribution system cause significant economic, social, and environmental impacts as well as severe damages to the surrounding soils and infrastructures. Despite the important researches on the development and use of hardware and software-based methods for the detection and localization and localization of water leaks, professionals still need efficient and cost-effective methods for the detection of water leaks in complex water distribution systems.

The recent progress in smart monitoring and Artificial Intelligence provides significant opportunities for the development of data-based methods for leak detection and localization. The literature review showed an important concern in the use of these methods. However, on the one hand, the majority of the applications of the Intelligent Artificial methods remain at the research stage. On the other hand, the literature review revealed a lack of a comprehensive use of these methods. This research work aimed to fill the gap in this area through a complete investigation of the Machine Learning methods to detect and localize leaks in the water distribution system.

The water network of the scientific campus of Lille University was used as a support for this research. This use is motivated by the campus representatively of a small town, the complexity of the water network, and the availability of data about the water network asset and water consumption. The water network is monitored by about 93 Automated Meter Reading (AMR) that record the water supply and consumption in the main buildings at an hourly-time interval.

The physical water network was completed by the construction of a Lab pilot of this network to investigate, under well-controlled conditions, the impact of the position of a leak on the water flow rates. Results of experiments showed an evident influence of the leak position leak on the water supply flow rates, when the leak is in the proximity of the water supply. For other locations, the impact is not clear, which means that the leak position could not be systematically determined from only the supply flow rates. In the future, it could be interested

in monitoring the pilot with pressure cells to investigate the possibility of improving the leak localization using the water supply flow rates and the pressure variation in the water network.
A large data set was built about the impact of leaks in the water network on the scientific campus on the variation of the water supply flow rates and the pressure in five zones of the campus. This data set was constructed using the hydraulic software EPANET. The dataset included the responses of the water network to 215 individual and double leaks.
The dataset was used for training and testing the following six Machine Learning methods:

- Three supervised methods: Logistic Regression, Decision Tree, and Random Forest.
- Two unsupervised methods: The Hierarchical Classification method and a combination of the PCA and K-means classification method.
- The Artificial Neural Network

The results of tests conducted on these methods showed:

- -Excellent performance of the supervised methods in the localization of leaks in the water network. Both the Logistic Regression and the Random Forest predicted the position of the leak with an Accuracy = 1.0, while the Decision Tree predicted leaks with an Accuracy = 0.98 with pressure and flow data.
- Excellent performances of the Artificial Neural Network for the localization of the water leaks in the water network (Accuracy = 1.0).
- Some difficulties in the exploitation of the clustering capacity of the unsupervised methods in the leak localization because of clusters' overlapping.

The results of this research were used for the investigation of the position of water leaks in the campus using water flow data rates recorded in 2015. Difficulties were encountered in the determination of the position of leaks because of a lack of pressure data. In the future, we recommend extending the monitoring of the campus water network by adding cell pressure on the campus as well as flow rates in critical sections of the water network.

5 References

Adedeji, K., Y. Hamam, B. Abe, and A. M. Abu-Mahfouz. 2017. "Leakage Detection Algorithm Integrating Water Distribution Networks Hydraulic Model." in SimHydro: Choosing the right model in applied hydraulics.

Afan, Haitham Abdulmohsin, Ahmed El-Shafie, Zaher Mundher Yaseen, Mohammed Majeed Hameed, Wan Hanna Melini Wan Mohtar, and Aini Hussain. 2014. "ANN Based Sediment Prediction Model Utilizing Different Input Scenarios." Water Resources Management. DOI: 10.1007/s11269-014-0870-1

Ahadi, Majid and Mehrdad Sharif Bakhtiar. 2010. "Leak Detection in Water-Filled Plastic Pipes through the Application of Tuned Wavelet Transforms to Acoustic Emission Signals." Applied Acoustics. DOI: 10.1016/j.apacoust.2010.02.006

Caputo, A. C., &Pelagagge, P. M. (2003). Using Neural Networks to Monitor Piping Systems. Process Safety Progress, 22(2), 119–127. https://doi.org/https://doi.org/10.1002/prs.680220208

Chan, T. K., Chin, C. S., & Zhong, X. (2018). Review of Current Technologies and Proposed Intelligent Methodologies for Water Distributed Network Leakage Detection. IEEE Access, 6, 78846–78867. https://doi.org/10.1109/ACCESS.2018.2885444

Charalambous, B., D. Foufeas, and N. Petroulias. 2014. "Leak Detection and Water Loss Management." Water Utility Journal.

Ciupke, K. (2018). Leak Detection Using Regression Trees. https://doi.org/10.1007/978-3-319-62042-8

Cobaner, Murat. 2011. "Evapotranspiration Estimation by Two Different Neuro-Fuzzy Inference Systems." Journal of Hydrology. DOI: 10.1016/j.jhydrol.2010.12.030

Da Cruz R.P., da Silva F.F., Fileti A.M.F. Machine learning and acoustic method applied to leak detection and location in low-pressure gas pipelines. Clean Technologies and Environmental Policy (2020), doi:10.1007/s10098-019-01805-x.

Di Nardo, A., M. Di Natale, C. Giudicianni, D. Musmarra, J. M. Rodrigue. Varela, G. F. Santonastaso, A. Simone, and V. Tzatchkov (2017). "Redundancy Features of Water Distribution Systems." in Procedia Engineering. DOI: 10.1016/j.proeng.2017.03.244

El-Shafie, A., O. Jaafer, and A. Seyed. 2011. "Adaptive Neuro-Fuzzy Inference System Based Model for Rainfall Forecasting in Klang River, Malaysia." International Journal of Physical Sciences.

Farah E. (2016) « Detection of water leakage using innovative smart water system – Application to SunRise Smart City Demonstrator» Thèsesoutenue le 2 novembre2016 ;Université de Lille.

Farah E., A. Abdallah & I. Shahrour (2017). SunRise: LARGE SCALE DEMONSTRATOR OF THE SMART WATER SYSTEM. Int. J. Sus. Dev. Plann. Vol. 12, No. 1 (2017) 112–121

Farah F., I. Shahrour (2017) Smart Water for Leakage Detection: Feedback about the Use of Automated Meter Reading Technology. Sensors Networks Smart and Emerging Technologies (SENSET), Beirut, Lebanon, 2017, IEEE Xplore, DOI: 10.1109/SENSET.2017.8125061.

Farah, E. & I. Shahrour (2017) Leakage Detection Using Smart Water System: Combination of Water Balance and Automated Minimum Night Flow. Water Resources Management, 31, 4821-4833.

Farah, E., & Shahrour, I. (2017). Leakage Detection Using Smart Water System: Combination of Water Balance and Automated Minimum Night Flow. Water Resources Management, 31(15), 4821–4833. https://doi.org/10.1007/s11269-017-1780-9

Farah, E., Abdallah, A., & Shahrour, I. (2017). SunRise: Large scale demonstrator of the smart water system. International Journal of Sustainable Development and Planning. https://doi.org/10.2495/SDP-V12-N1-112-121

Gao, Y., M. J. Brennan, P. F. Joseph, J. M. Muggleton, and O. Hunaidi. 2004. “A Model of the Correlation Function of Leak Noise in Buried Plastic Pipes.” Journal of Sound and Vibration .DOI: 10.1016/j.jsv.2003.08.045

Golmohamadi, Marcia. 2015. “Pipeline Leak Detection.” MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY. Electronic OCLC # 913485554

Hamilton, Stuart and Bambos Charalambous. 2017. Leak Detection: Technology and Implementation. DOI: 10.26530/oapen_578133

Han, Q., Nguyen, P., Eguchi, R. T., Hsu, K. L., &Venkatasubramanian, N. (2017). Toward An Integrated Approach to Localizing Failures in Community Water Networks. Proceedings - International Conference on Distributed Computing Systems, 1250–1260. https://doi.org/10.1109/ICDCS.2017.81

Haxton, T., Janke, R., Murray, R., Burkhardt, J., &Grayman, W. (2019). EPANET: Future of Water Distribution Modeling and Data Analytics Tools. https://cfpub.epa.gov/si/si_public_file_download.cfm?p_download_id=539762&Lab=NHSRC

Hou, Chun Xia and Er Hua Zhang. 2013. “Pipeline Leak Detection Based on Double Sensor Negative Pressure Wave.” in Applied Mechanics and Materials. DOI: 10.4028/www.scientific.net/AMM.313-314.1225

Kanakoudis, Vasilis and Habib Muhammetoglu. 2014. “Urban Water Pipe Networks Management towards Non-Revenue Water Reduction: Two Case Studies from Greece and Turkey.” Clean - Soil, Air, Water. DOI: 10.1002/clen.201300138

Kashid, Santosh Gautam and Sanjay A. Pardeshi. 2014. “A Survey of Water Distribution System and New Approach to Intelligent Water Distribution System.” in 1st International Conference on Networks and Soft Computing, ICNSC 2014 - Proceedings.

Kumar, Dileep, Dezhan Tu, Naifu Zhu, Reehan Ali Shah, Dibo Hou, and Hongjian Zhang. 2017. “The Free-Swimming Device Leakage Detection in Plastic Water-Filled Pipes through Tuning the Wavelet Transform to the Underwater Acoustic Signals.” Water (Switzerland). DOI: 10.3390/w9100731

Lidia Ogiela. 2017. Cognitive Information Systems in Management Sciences 1st Edition. Academic Press. ISBN:0128038039 9780128038031

Mirzavand, Mohammad, Benyamin Khoshnevisan, ShahaboddinShamshirband, Ozgur Kisi, Rodina Ahmad, and ShatirahAkib. 2015. “Evaluating Groundwater Level Fluctuation by Support Vector Regression and Neuro-Fuzzy Methods: A Comparative Study.” Natural Hazards. DOI: 10.1007/s11069-015-1602-4

Mohanty, S., Madan K. Jha, S. K. Raul, R. K. Panda, and K. P. Sudheer. 2015. "Using Artificial Neural Network Approach for Simultaneous Forecasting of Weekly Groundwater Levels at Multiple Sites." Water Resources Management. DOI: 10.1007/s11269-015-1132-6

Mounce, S. R., Pedraza, C., Jackson, T., Linford, P., & Boxall, J. B. (2015). Cloud based machine learning approaches for leakage assessment and management in smart water networks. Procedia Engineering, 119(1), 43–52. https://doi.org/10.1016/j.proeng.2015.08.851

Mounce, Stephen R., Mounce, R. B., & Boxall, J. B. (2011). Novelty detection for time series data analysis in water distribution systems using support vector machines. Journal of Hydroinformatics, 13(4), 672–686. https://doi.org/10.2166/hydro.2010.144

Muggleton, J. M. and M. J. Brennan. 2004. "Leak Noise Propagation and Attenuation in Submerged Plastic Water Pipes." Journal of Sound and Vibration. DOI: 10.1016/j.jsv.2003.10.052.

Mukherjee, Amritendu and Parthasarathy Ramachandran. 2018. "Prediction of GWL with the Help of GRACE TWS for Unevenly Spaced Time Series Data in India : Analysis of Comparative Performances of SVR, ANN and LRM." Journal of Hydrology. DOI: 10.1016/j.jhydrol.2018.02.005

Murvay, Pal Stefan and IoanSilea. 2012. "A Survey on Gas Leak Detection and Localization Techniques." Journal of Loss Prevention in the Process Industries. DOI: 10.1016/j.jlp.2012.05.010

Mutikanga, harrison 2012. "WATER LOSS MANAGEMENT TOOLS AND METHODS FOR DEVELOPING COUNTRIES." Delft- Netherland.ISBN: 9780415636483

Rajeev, Pathmanathan, Jayantha Kodikara, Wing Kong Chiu, and Thomas Kuen. 2013. "Distributed Optical Fibre Sensors and Their Applications in Pipeline Monitoring." Key Engineering Materials. DOI: 10.4028/www.scientific.net/KEM.558.424

Renzetti, Steven and Diane Dupont. 2013. "Buried Treasure: The Economics of Leak Detection and Water Loss Prevention in Ontario." Environmental Sustainability Research Centre (ESRC) Working Paper Series. http://hdl.handle.net/10464/4279

Rojek, I., &Studzinski, J. (2019). Detection and localization of water leaks in water nets supported by an ICT system with artificial intelligence methods as away forward for smart cities. Sustainability (Switzerland), 11(2). https://doi.org/10.3390/su11020518

Romano M., Kapelan Z., Savic D.A. (2014) Evolutionary algorithm and expectation maximization strategies for improved detection of pipe bursts and other events in water distribution systems J. Water Resour. Plan. Manag., 140 (5) (2014), pp. 572-584

Saab C., I. Shahrour, F. Hage-Chehade (2017) Smart Technology for Water Quality Control: Feedback about use of water quality sensors. Sensors Networks Smart and Emerging Technologies (SENSET), Beirut, Lebanon, 2017, IEEE Xplore, DOI: 10.1109/SENSET.2017.8125060.

Saab Ch. (2018) "Smart system for water quality control: feedback from large-scale experimentation", Thèsesoutenue le 2 juillet 2018 ; Université de Lille, UniversitéLibanaise.

Salam, A. E. U., Tola, M., Selintung, M., & Maricar, F. (2014). On-line monitoring system of water leakage detection in pipe networks with artificial intelligence. ARPN Journal of Engineering and Applied Sciences, 9(10), 1817–1822.

Sanz, Gerard, Ramon Perez, and Antoni Escobet. 2012. "Leakage Localization in Water Networks Using Fuzzy Logic." in 2012 20th Mediterranean Conference on Control and Automation, MED 2012 - Conference Proceedings. DOI: 10.1109/MED.2012.6265711

Savic, Dragan A., Zoran S. Kapelan, and Philip M. R. Jonkergouw. 2009. "Quo Vadis Water Distribution Model Calibration?" Urban Water Journal. DOI: 10.1080/15730620802613380

Seckin, Neslihan, Murat Cobaner, Recep Yurtal, and TefarukHaktanir. 2013. "Comparison of Artificial Neural Network Methods with L-Moments for Estimating Flood Flow at Ungauged Sites: The Case of East Mediterranean River Basin, Turkey." Water Resources Management. DOI: 10.1007/s11269-013-0278-3

Shahrour I., O. Abbas, A. Abdallah, Y. AbouRjeily, A. Afaneh, A. Aljer, B. Ayari, E. Farrah, D. Sakr, F. Al Masri (2017). Lessons from a Large Scale Demonstrator of the Smart and Sustainable City. In Happy City Book, A. Brdulak and H. Brdulak (eds.), Springer International Publishing AG 2017, DOI 10.1007/978-3-319-49899-7_11

Shahrour, I., Abbas, O., Abdallah, A., AbouRjeily, Y., Afaneh, A., Aljer, A., Ayari, B., Farrah, E., Sakr, D., & Al Masri, F. (2017). Lessons from a Large Scale Demonstrator of the Smart and Sustainable City BT - Happy City - How to Plan and Create the Best Livable Area for the People (A. Brdulak& H. Brdulak (eds.); pp. 193–206). Springer International Publishing. https://doi.org/10.1007/978-3-319-49899-7_11

Shiri, Jalal, Mohammad ZounematKermani, Ozgur Kisi, and Sahar Mohsenzadeh Karimi. 2019. "Comprehensive Assessment of 12 Soft Computing Approaches for Modeling Reference Evapotranspiration in Humid Locations." Meteorological Applications. DOI: doi.org/10.1002/met.1841

Shravani, D., Prajwal, Y. R., Prapulla, S. B., Salanke, N. S. G. R., Shobha, G., & Ahmad, S. F. (2019). A Machine Learning Approach to Water Leak Localization. 2019 4th International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), 4, 1–6. https://doi.org/10.1109/CSITSS47250.2019.9031010

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

Soldevila, A., Fernandez-Canti, R. M., Blesa, J., Tornil-Sin, S., & Puig, V. (2017). Leak localization in water distribution networks using Bayesian classifiers. Journal of Process Control, 55, 1–9. https://doi.org/10.1016/j.jprocont.2017.03.015

Soltani, Foad, Reza Kerachian, and Ehsan Shirangi. 2010. "Developing Operating Rules for Reservoirs Considering the Water Quality Issues: Application of ANFIS-Based Surrogate Models." Expert Systems with Applications. DOI: 10.1016/j.eswa.2010.03.057

Sousa, Joaquim, Luísa Ribeiro, João Muranho, and Alfeu Sá Marques. 2015. "Locating Leaks in Water Distribution Networks with Simulated Annealing and Graph Theory." in Procedia Engineering. DOI: 10.1016/j.proeng.2015.08.854

Thornton, J., R. Sturm, and George Kunkel. 2008. Water Loss Control. Second. McGraw Hill. ISBN: 9780071499187

Tinelli, Silvia and IlanJuran. 2019. "Artificial Intelligence-Based Monitoring System of Water Quality Parameters for Early Detection of Non-Specific Bio-Contamination in Water Distribution Systems." Water Science and Technology: Water Supply. DOI: 10.2166/ws.2019.057

Valizadeh, Nariman and Ahmed El-Shafie. 2013. "Forecasting the Level of Reservoirs Using Multiple Input Fuzzification in ANFIS." Water Resources Management. DOI: 10.1007/s11269-013-0349-5

van der Walt, J. C., Heyns, P. S., & Wilke, D. N. (2018). Pipe network leak detection: comparison between statistical and machine learning techniques. Urban Water Journal, 15(10), 953–960. https://doi.org/10.1080/1573062X.2019.1597375

Wan, Jiangwen, Yang Yu, Yinfeng Wu, Renjian Feng, and Ning Yu. 2012. "Hierarchical Leak Detection and Localization Method in Natural Gas Pipeline Monitoring Sensor Networks." Sensors. DOI: 10.3390/s120100189

Wu, Y., & Liu, S. (2017). A review of data-driven approaches for burst detection in water distribution systems. Urban Water Journal, 14(9), 972–983. https://doi.org/10.1080/1573062X.2017.1279191

Wu, Z. Y., El-Maghraby, M., & Pathak, S. (2015). Applications of deep learning for smart water networks. Procedia Engineering, 119(1), 479–485. https://doi.org/10.1016/j.proeng.2015.08.870

Yalçin, B. C., Demir, C., Gökçe, M., &Koyun, A. (2018). Water leakage detection for complex pipe systems using hybrid learning algorithm based on ANFIS Method. Journal of Computing and Information Science in Engineering, 18(4), 1–10. https://doi.org/10.1115/1.4040130

Yalçin, Baris Can, Cihan Demir, Murat Gökçe, and Ahmet Koyun. 2018. "Water Leakage Detection for Complex Pipe Systems Using Hybrid Learning Algorithm Based on ANFIS Method." Journal of Computing and Information Science in Engineering. DOI: 10.1115/1.4040130

Zhang;, Q., Wu;, Z. Y., Zhao;, M., Qi;, J., Huang;, Y., & Zhao, H. (2016). Leakage Zone Identification in Large-Scale Water Distribution Systems Using Multiclass Support Vector Machines. Journal of Water Resources Planning and Management, 142(11), 4016042. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000661

Zhou, X., Tang, Z., Xu, W., Meng, F., Chu, X., Xin, K., & Fu, G. (2019). Deep learning identifies accurate burst locations in water distribution networks. Water Research, 166, 115058. https://doi.org/10.1016/j.watres.2019.115058

