Prelocalization and Leak detection in water drinking distribution network using modeling-based algorithms: Case study: The city of Casablanca (Morocco)

Faycal Taghlabi [1], Laila Sour [1], Ali Agoumi [2]

5   [1] Laboratory of Processes and Environment, Faculty of Science and Technology of Mohammedia, Hassan II university of Casablanca, 28806, Morocco
[2]Laboratory of Civil, hydraulic Engineering, Environment and Climate, Hassania School of Public Works, Casablanca,20200, Morocco

Correspondence to: Faycal Taghlabi (ftaghlabi@yahoo.fr)

10   **Abstract.** The role of a water drinking distribution network (WDDN) is to supply high-quality water at the necessary pressure at various times of the day for several consumption scenarios. Locating and identifying priorities of water leakage areas becomes major preoccupation for manager of the water supply, to optimize and improve constancy of supply. In this paper, we present the results obtained on the field from a research conducted in order to identify and to locate leaks in (WDDN) focused on the resolution of the Fixed And Variable Area Discharge (FAVAD) equation by use of the prediction algorithms

15   in conjunction with hydraulic modeling and the Geographical Information System (GIS).The leak localization method is applied in the oldest part of Casablanca. We have used, in this research, two methodologies in different leak episodes: (i)The first episode is based on a simulation of artificial leaks on the MATLAB platform using the EPANET code to establish a database of pressures that describe the network's behaviour in the presence of leaks. The data thus established has fed into a machine learning algorithm called Random Forest, which will forecast the leakage rate and its location in the network; (ii)The

20   second was field-testing a real simulation of artificial leaks by opening and closing of hydrants, on different locations with a leak size of 6l/s and 17 l/s. The two methods converged to comparable results, the leaks position is spotted within a 100 m radius of the actual ones.

Keywords: FAVAD; WDN; Leak localization; prediction; Epanet.

1 Introduction

25   Climate Change (CC) is a major global issue, more and more important on the international scene. It affects all components of the hydrological cycle. The situation of water resources in Morocco is already critical with a state of water scarcity forecasted for 2020. This problem is accentuated by the effects of CC and may hinder any further sustainable development. The expected CC for Morocco would have direct and indirect harmful consequences on the water resources potential, in terms of both quantity and quality, on the water demand and on the efficiency of use of this resource by the different users. An anticipation

30   of the adaptation to the effects of this CC must pass by the valorization of the use of the resources and especially the

minimization of the water losses. In this regard, in Moroccan urban areas, drinking water distribution networks have particularly low yields. The location and prioritization of leaking areas is a major concern for the public authorities to optimize the use of water resources, reduce losses and improve continuity of service.

To guarantee the high-level service of pressure, the detection and repair time of leaks is certainly the most common factor used
35 in the analysis of decreases in contract pressures.

For most of the time, before starting a leak detection campaign in a Discrete Hydraulic Sector (DHS) we start with the analysis of the flows into and out of the sector, in particular the minimum night flow (MNF) between 2:00 AM and 4:00 AM, as well as the volumes of major consumers (Alkasseh et al., 2013).

In the literature it is possible to detect leaks in the DHS, Usually the leakage rate is permanent over time, if the DHS records
40 an increase in night flow, this increase should also appear during normal consumption time (Oasen, 2015a). According to research by Farley et al., (2008) An increase in minimum night flow can be used for targeting all DHS where leakage is more likely. It is therefore possible to detect leaks in a DHS by making a hydraulic balance between the volume of billed consumption and the volume distributed, by comparing the expected demand and the actual water consumption (Bakker, 2014).

Once new leaks by DHS are identified, various techniques are used to locate the leaks. Acoustic leak-detection is a technique
45 which has evolved a lot in recent years and is developing rapidly (Farley, 2003). Some of these techniques require partitioning a WDDN into smaller DHS, by closing certain valves on the network, which can sometimes shutdown the system (Colombo, 2009).

In addition, various research projects noted that it is difficult to apply the leak-detection to certain areas due to the complexity of isolating and partitioning (Andrea et al., 2011). Through the applied works in modeling leakage, in particular, those of Babel
50 et al., (2009) and Sebbagh et al., (2018), a reduction in pressure at the inlet of DHS, induces a reduction in leakage rate. For Al-Ghamdi et al (2011), a 25% reduction in pressure contributes to a leakage flow reduction of about 25% for a 50% rigid 50% plastic network.

Our approach, as we will see through the following paragraphs, is to do a virtual leak search without partitioning a WDDN into smaller DHS.

55 The deficiency of leak management is one of the key problems, given its impacts on production cost and resource exhaustion. The scope of this paper will be mainly focused on the application of the two approaches in two different leak events. The case study is a pilot sector in the city of Casablanca (Morocco), which covers around 24 000 inhabitants as displayed in Fig. 1

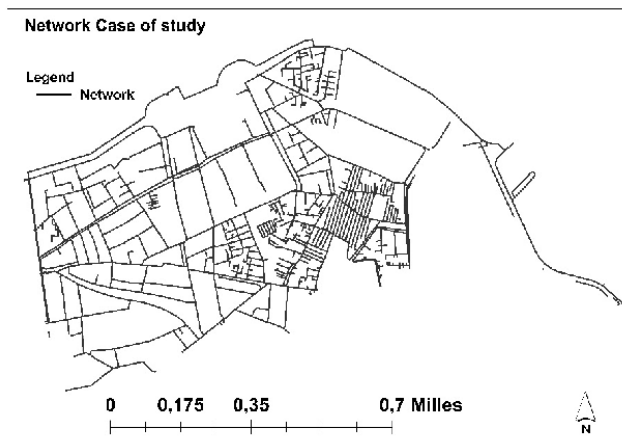
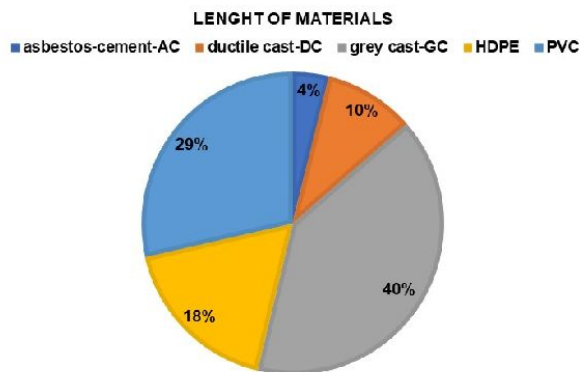



Figure 1 Delimitation of the study area.

60   The study area has three inlets, 493 nodes and 42 km of pipes. Concerning the instrumentation, the network flow and pressure are monitored through flowmeter in diameter 300 mm and pressure sensors at each inlet.

Figure 2 reveals that the network consists mainly of greycast iron (40%). The average age is 40 years, increasing vulnerability and promoting leaks.

65   **Figure 2 Types of materials constituting the network**

2 Materials and methods

2.1 Software

EPANET is a free software developed by the US Environmental Protection Agency (U.S. EPA). From the representation of the distribution network (nodes, pipes, tank, valves, pump, etc.), it allows the hydraulic balancing of the network by the

70   calculations of pressure losses, flow velocity, flow in the pipes and pressure at the nods. (Rossman, 2000).

 In practice EPANET is used by water utilities (EPA 2005) and in literature (Farina et al., 2014).

The basic demand for the hydraulic modeling software EPANET 2.0 is defined as a water output at each node, We consider that there are two main methods to simulate a water leak in EPANET, as an additional demand; or even water flow rate through a Valve, the formula to calculate the head loss (Darcy-Weisbach) was used with the default values for the roughness (Brown,

75   2002).

2.2 Method

(i) Relationships between pressure and leakage rates in distribution networks.

Pressure management not only involves reducing pressure, but also other pressure control and optimization methods without compromising customer service. A definition of pressure management in its broadest sense is given by Thornton et al., (2005),

80   "pressure management is about controlling the pressure of the system to achieve a level of optimal service, to ensure an efficient supply to consumers while avoiding the unnecessary excesses of this pressure which would unduly increase leaks.

Water utilities often take to design their distribution networks the minimum pressure that occurs at the critical point at maximum demand. Understanding this concept is of great importance as pressure regulation can significantly reduce leakage without compromising the level of customer service.

85   The term "FAVAD" comes from the English "Fixed and Variable Area Discharge Paths". This concept, through the definition of an exponent N1, defines the relationship between the leak rate and the pressure (Rozental, 2010):

The relationship between pressure and flow leaks is given by Eq. (1):

$$L1 = L0 \times (P1/P0)^{N1} \tag{1}$$

With $L0$ and $L1$: the leakage flow before and after pressure reduction.

90   P0 and P1: the pressure before and after reduction.

According to Al Ghamdi, (2011) et Cobacho et al., (2014), the main method for representing leaks in a hydraulic network model is through adding a leakage valve for each node, the emitter parameter is used to model flow rate through a valve. These emitters devices permit the modelling of flow evacuated to the atmosphere through a nozzle.

the equation below represents the concept of FAVAD, through a flow rate, pressure and emitter coefficient Eq. (2):

95   $$Qleak \, j = C \times Pj^{\,N} \tag{2}$$

where Qleak is the flow rate at node j, C is the emitter coefficient, P pressure at node j

The exponent N1 of the above equation varies according to the material of the pipeline (mainly its elasticity), for a circular opening on a rigid pipe (cast iron, steel), N1 is of the order of 0,5 whereas it reaches 1,5 or more for longitudinal slots on plastic materials (PVC, PEHD). However, international feedback shows a variation of N1 between 0.36 and 2.95 depending on the networks experienced, as shown in Fig. 4 (Rozental, 2010). Figure 3 illustrates the influence of N1 the exponent emitter on the impact of pressure reduction on leakage rate.

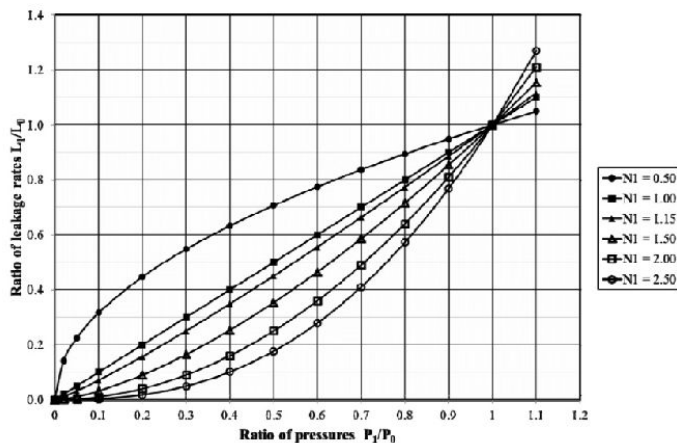



Figure 3 Relationships between pressure and leakage rate using the N1 Approach (Rozental, 2010)

The data collection relative to the various components of the drinking water networks (pipes, reservoirs, well, drilling, pumps, valves) is made by means of shapefile exported from the database cart@jour. It's an interface GIS (Geographical Information System) available for consultation in intranet which includes three networks managed by the drinking water operators as well as all of the hydraulic structure which constitutes them. The interface allows extracting all the desired layers while geometrically targeting the study area (zone of study). the following Table illustrates the roughness values used during modeling (Chadwick et al., 2013).

Table 1 Roughness coefficient of materials (values from Chadwick et al., 2013)

Material	Age [year]	Darcy-Weisbach roughness coefficient [mm]
AC	0-10	0.10
	10-40	0.15
	>40	1.2
All plastics: HPE, PVC, ZPE	0-10	0.05
	10-40	0.10
	>40	0.15
ST	0-10	0.3
	10-40	1.0
	>40	2.0

The elevations are extracted from the Digital Elevation Model (DEM) layer and automatically assigned to network nodes. The annual average consumption for 2017 in addition, are distributed into each node in the model according to the geographical distribution of subscribers within the tour. Once the network template is prepared using ArcGIS, it is transferred to the

115   EPANET software in .inp file.

The EPANET hydraulic simulation model calculates node pressure and pipe flow for a fixed reservoir level and variable water demands over time and space. The calculation of the head and the flow at a particular point in time involves the simultaneous resolution of the flow conservation equation for each of the nodes and the equation of the pressure drop in each pipe of the network. Dynamic simulation is used to describe the operation of the network during a given period, while taking into account

120   the variation in customers' consumption over time.

Several simulations using the Epanet software were used to determine the roughness coefficients of the pipes to obtain calculated pressures which indicate the actual pressures in different nodes of the hydraulic system.

To solve this localization of leakages we will use the random forest algorithm (R Learning and Prediction Algorithm).

Algorithm 1 (Section 2,3,2) illustrates briefly the suggested leakage detection procedure.

125   The leakage localization methodology displayed in Fig. 4 is based on data mining algorithms, the starting point of the algorithm is the learning of the data obtained by simulation using EPANET simulator. Then four training data elements are used to predict the location of the leak:

- The distance between simulation node and the sensor
- The leakage flow

130   • Emitter Coefficient
- Pressure at the sensors

In order to estimate the network flow, the leakage outflow at each node, and the distance of leak from sensors, a leakage model is implemented within a classical hydraulic simulation model (EPANET).

(ii) Leak simulation in Epanet:

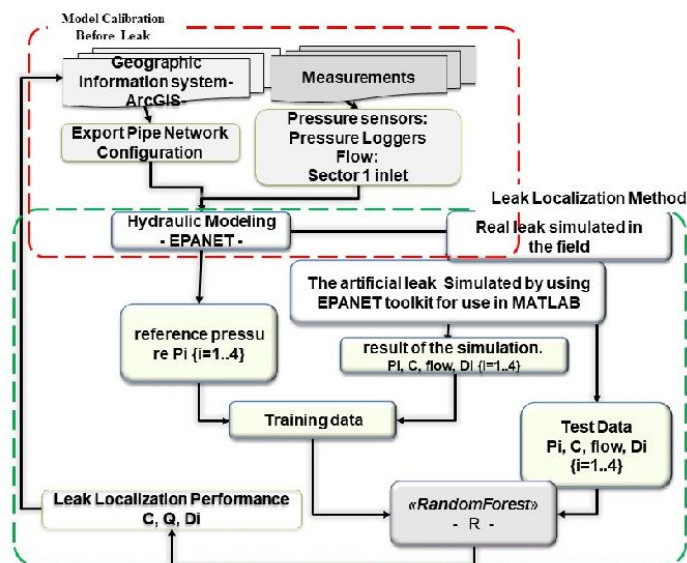
135   This training data is obtained by using EPANET toolkit available in [https://github.com/OpenWaterAnalytics/EPANET-Matlab-Toolkit] for use in MATLAB. It is an open-source software that provides programming interface for EPANET within MATLAB framework. It is easy to modify, simulate and plot the result produced by EPANET libraries. Figure 5 reveals that 493 nodes were defined to simulate leaks with different emitter coefficients: 0.8, 1.6, 2.4 and 3.2, which corresponds approximately to leak rates of 5, 10, 15 and 20 l/s as per Eq. (2)

140

Figure 4 Leak pre-localization procedure

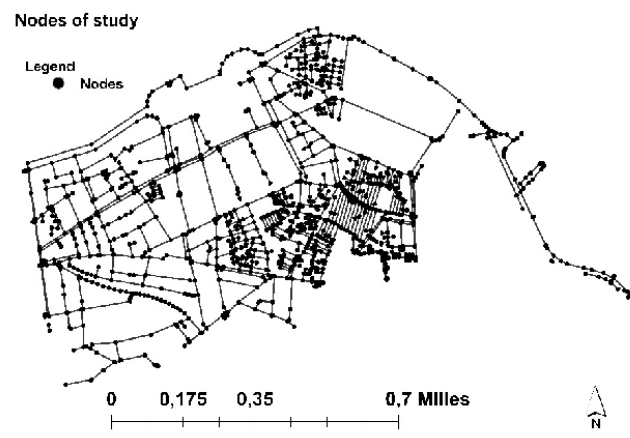



Figure 5 Leak simulation nodes in sector 1

Thus, at each node, we simulated 4 leaks with emitter coefficients equal to 0.8, 1.6, 2.4 and 3.2, for a total of 1972 simulations.

145   Each time, we note the leakage rate, the emitter coefficient, the location of the leakage through the hydraulic distances of the

simulation node with respect to the four sensors P14, P57, P58 and PC, as well as the maximum, minimum pressures. and average at these measurement points, the results are reported in Table 2.

Table 2 Result of Model Leak Simulations

Data	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8	……
Pmax1	38,91	38,85	38,78	38,70	38,91	38,85	38,78	38,70	……
Pmin1	37,33	37,16	36,98	36,80	37,33	37,16	36,98	36,80	
Pmoy1	38,39	38,29	38,17	38,04	38,39	38,29	38,17	38,04	
Pmax2	37,48	37,43	37,36	37,29	37,48	37,43	37,36	37,29	
Pmin2	35,66	35,50	35,34	35,16	35,66	35,50	35,34	35,16	
Pmoy2	36,87	36,77	36,65	36,53	36,87	36,77	36,65	36,53	
Pmax3	36,89	36,84	36,78	36,71	36,89	36,84	36,78	36,71	
Pmin3	34,91	34,75	34,59	34,42	34,91	34,75	34,59	34,41	
Pmoy3	36,20	36,11	35,99	35,87	36,20	36,11	35,99	35,87	
Pmax4	34,51	34,46	34,40	34,33	34,51	34,46	34,40	34,33	
Pmin4	32,17	32,02	31,86	31,69	32,17	32,02	31,86	31,69	
Pmoy4	33,66	33,56	33,45	33,33	33,66	33,56	33,45	33,33	
C	0,8	1,6	2,4	3,2	0,8	1,6	2,4	3,2	
Q	4,96	9,90	14,83	19,73	4,96	9,90	14,83	19,73	
D1	478,21	478,21	478,21	478,21	475,51	475,51	475,51	475,51	
D2	399,55	399,55	399,55	399,55	483,06	483,06	483,06	483,06	
D3	486,43	486,43	486,43	486,43	483,06	483,06	483,06	483,06	
D4	729,73	729,73	729,73	729,73	726,19	726,19	726,19	726,19	

150

Or :

- D1: the distance from the simulation node to the P14
- D2: the distance from the simulation node to the P57
- D3: the distance from the simulation node to the P58
155 - D4: the distance from the simulation node to the PC
- Q: the leakage
- C: emitter coefficients
- Pmax 1: The maximum pressure during low consumption period at P14
- Pmin 1: the minimum pressure in rush hour period at P14

160 •   Pmoy 1: average daily pressure at P14

•   Similarly, for the indices 2,3 and 4 which correspond respectively to the P57, P58 and PC,

The table thus constructed constitutes the input data of our algorithm

2.3 Location and simulation of leaks on the ground

2.3.1 leak pre-locating

165   The technique used to search preventive leakage on distribution networks is organized around three distinct but complementary operations: sectoring, pre-location, followed by localization. These methods must be adapted according to the dimensions and the degree of knowledge of the targeted.

The objective of sectorization is to define priorities between different sectors and to estimate or even quantify the level of leakage. It defines fugitive areas larger than the linear kilometer;

170   The objective of the pre-location is to check the presence of leaks in a given sector and to determine their position with a precision of the order of magnitude of the hundred meters;

The objective of the location is to define the position of a leak with a precision of the order of one meter.

2.3.2 Pre-localization of leaks by learning data

-   Random Forest

175   The type of learning we apply to anomaly detection in this article is a supervised learning. In Zhang et al., (2008), decision tree forests are used to detect intrusions from the network.

To execute the anomaly detection method by supervised learning, we used the statistical software -R (Zhang et al. 2008).

Random Forest is a type of tree based supervised learning algorithm (Ho, 1995). It uses many decision trees to aggregate the answer. In this paper, the supervised Random Forest algorithm was used as technique to detect the leaks (Breiman, 2001). In

180   addition to its efficiency, this algorithm is famous for its ability to treat big-data.

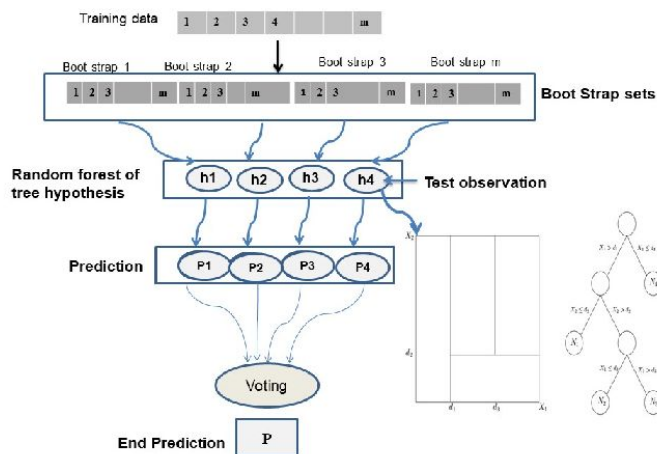
The random forest optimization principle is based on the combination of multiple decision trees, to extract different classes from the original raw dataset. Then, the average classes are determined based on the classes outputted by the decision trees used. Thus, the performance of the resulting model is enhanced, compared to one decision tree model, and the ability to apply the resulting model in other datasets is acquired. Figure 6 illustrates the principle of running a random forest algorithm.

185

Figure 6 Random Forest Tree

- Leakage Detection Algorithm:

The input data are: Training data td, (number of sensor node, the distance between simulation node and the sensor, the leakage flow Emitter Coefficient and Pressure at the sensors).

190 The process engaged in the proposed leak detection is concisely discussed in the following Algorithm:

Algorithm:

1: Start {

2: Load Training data network parameters

3: Read network parameters and initialise

195 4: for node i= 1 to nt, (nt: The number of nodes in the network)

Run hydraulic analysis and compute leakage Qleak

5: if Qleak < tolerance (relatively low)

6: Print "L= 0" No leaking node"

7: else

200 8: i: Print "Leaking node ID"

9: if Qleak > tolerance

10: Print "L= 1" leaking node"

11: else

12: i: Print "Leaking node ID"

205 13: Display "Distance from sensor, pressure node, Qleak ..."

14: end if

15: end if

16: end for i

17: Stop}

210 **2.4 Preparation of input data for the algorithm**

Several pressure profiles for reference (without leaks) are required to attain satisfactory level of prediction from the data analysis algorithms. This pressure profiles are obtained by using EPANET. Around 7 references cases are added to the table 1. A first case is the pressure reference that has been simulated, the others constitute a translation of the reference curve of +0.1, +0.2, +0.3, -0.1, -0.2, and 0.2 meters, forming envelope with 0.6m amplitude as shown in Fig. 7 below case of P14

215 measuring point.

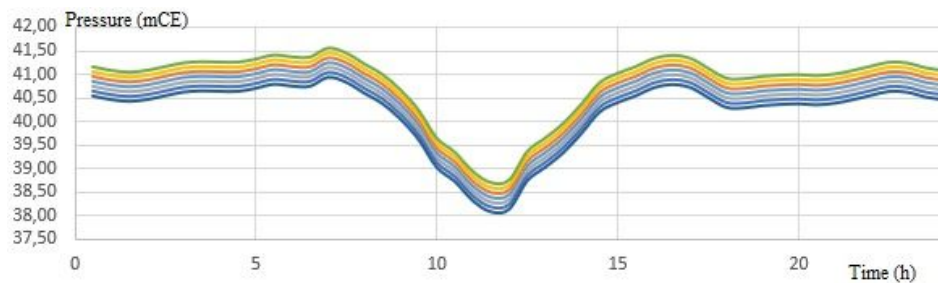



Figure 7 The reference pressures profile at P14

For each measurement point we considered seven reference curves. This envelope is the area of variation of the daily pressure, it was calculated from several measurements of pressure spread over several days.

220 The artificial leaks are measured with fire hydrant controller (FHC).

Using 3 artificial leaks and one real leak, the leak localization method is validated.

A total of 4 leaks were simulated by opening fire hydrants rated A, B and C, the artificial leaks locations are shown in Fig. 8.

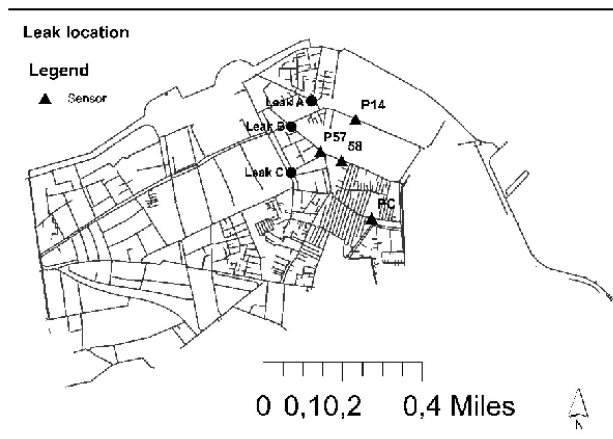



Figure 8 Sensors and leaks locations

225 The leak localization method depends on the head loss, itself depends, among other factors, on the size of the leak. Thus, if we have a limited leak, the sensors could not detect the small head loss, a leak of 6 l/s is chosen as a lower simulation limit. As upper limit, we chosen the 17 l/s leak, and we presume that those bypassing this limit finish by surfacing, and therefore do not require any localization process. (Pérez et al., 2014a).

At "PI A" we simulated two leaks of 6 l/s and 17 l/s; for the "PI B" a small leak of 6 l/s and for "PI C" a leak of 17 l/s (Table 230 3).

Table 3 The artificial leaks that were created for each location

ID Leaks	fire hydrants	Flow
LA-1	PI A	6 L/s
LA-2	PI A	17 L/s
LB	PI B	6 L/s
LC	PI C	17 L/s

3. Results and Discussion

3.1 Data-reading pressure at sensor

The performance of the pressure for the artificial leak simulation during May the 3rd is shown in Fig. 9. The blue line in Fig.9 235 shows the daily pattern of pressure at P14.

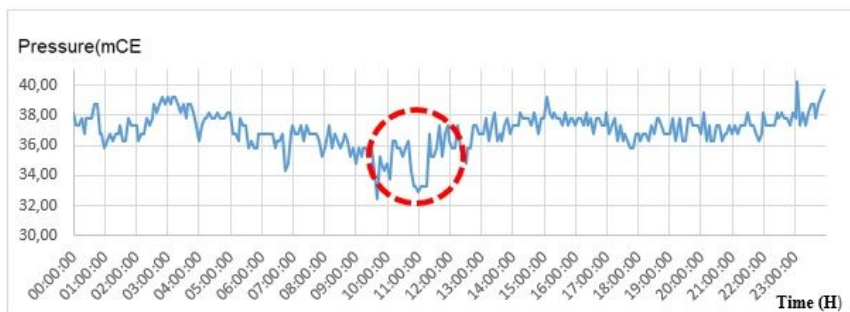



Figure 9 Pressure profile of 3/5/17 at P14

The red circle shows the decrease of the pressure during the real simulation

These leaks were simulated on the EPANET model by choosing the same time slot of the real simulations

240   The model was calibrated without errors. It was tested and validated by comparing the measured and simulated pressures

The results of the simulation are very close to what is measured as shown in Fig. 10 and 11.

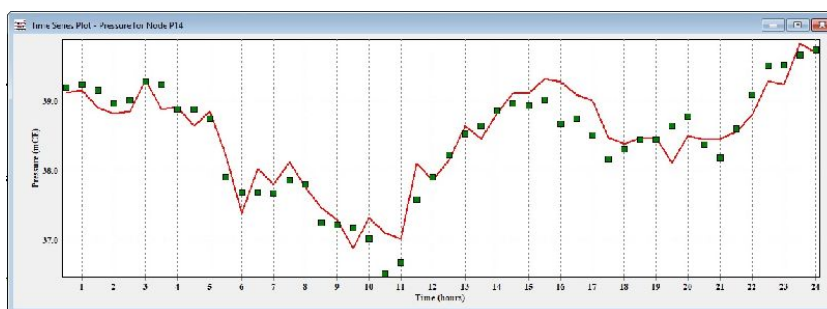
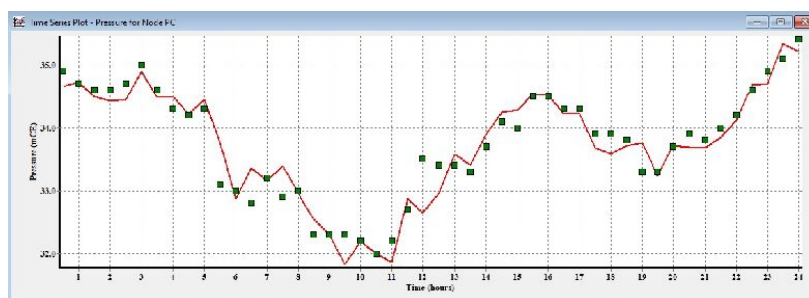



Figure 10 Pressure profile at P14

245 **Figure 11 Pressure profile at PC**

The red line corresponds to the simulation within EPANET and the green nodes to the pressures measured by the sensors.

To calculate the magnitude of the leak for each time and node, Eq.2 was used. The following Table shows the emitter coefficients simulated on each network node.

Table 4 Emitter coefficients corresponding to simulated leaks

ID Leak	Flow	Emitter coefficient
LA-1	6 l/s	1
LA-2	17 l/s	2,2
LB	6 l/s	1
LC	17 l/s	2,4

250 Table 5 resumes where the leaks are spatially spotted and the value of leakage flow at each node of the network. using the Random Forest algorithm.

Table 5 Quantification of the leaks.

Date	L	Emitter coefficient	Flow rate	D1	D2	D3	D4
22/04	0	0,41	2,47	91	45	39	84
23/04	1	2,21	13,77	529	362	336	358
24/04	0	0,33	2,62	122	47	41	86
25/04	0	0,34	2,31	48	53	44	148
26/04	1	2,99	17,89	525	344	320	332
27/04	0	0,32	2,10	65	37	46	45
28/04	0	0,01	0,04	4	0	0	2
09/05	0	0,01	0,03	3	0	0	2
10/05	0	0,01	0,03	3	0	0	2
11/05	0	0,03	0,27	37	16	14	18
Leak LA-1	1	2,35	14,12	229	239	291	591
Leak LA-2	1	2,60	16,04	201	214	259	517
Leak LB	1	2,62	16,25	253	207	271	464
Leak LC	1	3,00	18,34	288	110	164	345

Note that for the 4 last leaks we have L = 1 which means that the algorithm classified them as leaks. However, the leak is over-estimated. The algorithm shows the days of 23/04 and 26/04 as being cases of leaks.

255 The results of these two days are not used because the change in pressure profile during these days was not because of a leak but because there was closure of some valves in the area to do some work,

For other days we have indeed L = 0

The data analysis confirms that, for an emission coefficient of at least 2, the leaks pre-localization via the adopted method is possible in particularly for flows passing 10 l/s. Indeed, these values of flow provoke important "head loss" easily detected by

260 the pressure sensors implemented within the acting zone, which confirms the hypothesis made at the beginning of this study.

Two phenomena can explain the limits of the current method in terms of its capacity to detect leaks with low flow values. First of all, according to Jarrige et al., (2011) several factors may influence the leak noise propagation til the sensors, such as: the material type, the pipe diameter, and more importantly the pipe roughness. In fact, the misevaluation of this last factor influences the reference pressure calculation. According to Paquin et al., (2000) the results prove that it is necessary to measure

265 the real roughness in order to interpret correctly a simulated and a measured pressure.

The second reason beyond the limited performance of the proposed method to spot leak characterized by low flow was highlighted by a study of Mirats-Tur et al., (2014). In their paper the authors demonstrated that a mis-calibrated hydraulic model (in terms of the topographic structure and its parameters), the precision regarding the estimation of the spatial water need distribution within the sector of application and the precision of the sensors embedded in the network are all eventual

270 causes of the mentioned detection issues.

In fact, like the Mirats-Tur et al. (2014) study, our proposed method is mainly relying on the sensitivity analysis of the pressure measures to the water demand fluctuation in the network nodes. The only difference is that we simulated leak flow series at the level of fire hydrant according to different buffer (radius of 100m, 200m, 300m..), and the only aim is to study the influence radius of each sensor to detect the real flow simulated. This analysis permitted an optimal distribution of sensors in the network.

275 For Mirats-Tur et al (2014), although the sensors were narrowly installed, the leak position was determined within a 150 m radius, compared to the 100 m radius of our method.

The presented results are outputted from a model established using measures from the network. Some of these measures are considered to have a good precision, and other have a certain level of uncertainty. For instance, the roughness and the nodes' elevations measures are highly impacted by uncertainties. Another factor that impact the precision of the proposed method is

280 the measuring devices in terms of their recording interval (the pressure is measured each 5 minutes). In order to optimize the detection, and to focus on the leaks with high head losses spotted by standard sensors, it is recommended to use sensors with high frequency, capable of recording a high number of samples. This will help detect the small pressure variation caused by low leak flow.

3.2 Displaying results

285 The proposed method outputs are not always reliable. In fact, instead of a deterministic mapping of the leaks, there is a probabilistic output that maps the probability of occurrence of leaks in space (Pérez et al., 2014b).

If the forecast indicates a leak, which is estimated, to locate it on the map, we have 4 distances from the 4 pressure sensors, around each sensor we draw a circle of radius corresponds to the given distance by the prediction Fig. 12, the ideal would then be that these four circles intersect at a single point which corresponds to the leak point,

290   The intersection of these 4 circles will then be at the maximum at 12 points if all the circles intersect with each other at two points, considering two circles, there are three cases:

•   The circles intersect at two points, one corresponds to the leak point and the other is his symmetrical with respect to the line passing through the centers of the two circles,

•   Circles tangent to each other intersect at a single point that corresponds to the leak point

295   •   The circles do not cross, but if the forecast is good, they can get closer in the leaking zone

By drawing all the points of intersection, around each of them, the leakage location is identified within a 100 m radius.

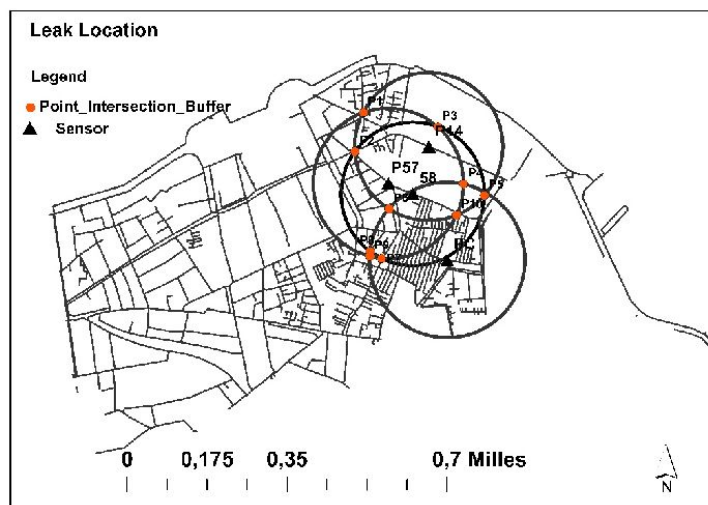
A good performance with highest probability of having the location of leak when there is a big agglomeration of circle that has a more intersect point. This agglomeration corresponds to the cumulative probability of the given nodes to experience a leakage (Fig. 13).

300

Figure 12 The spatial intersection of circles

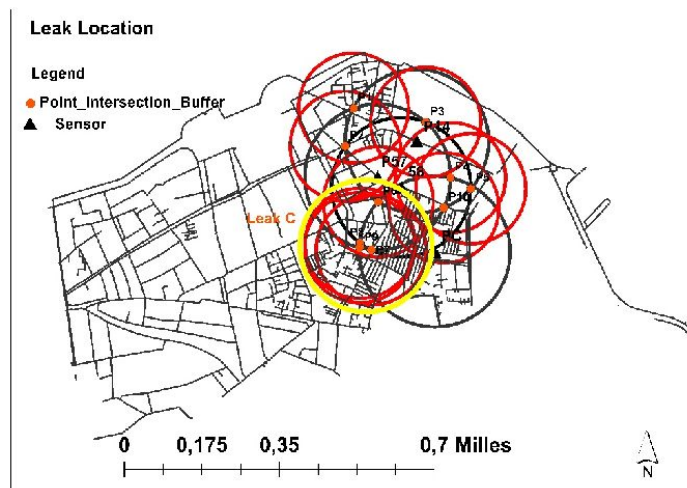



Figure 13 Spatial location of the leak LC

Our research objective was the purpose of discovery appropriate solution for detection and localization of leakages and

305   estimation of the size of leakages for a water distribution system.

The results obtained using this approach is satisfying. The leak is identified within a 100 m radius.

That said, the detection of a leak is extremely related to its location within the network. For instance, one located in the looped section of the network is less likely to be spotted in night time. In the mesh part of the network the pressure fallen at the sensor levels are too low, which could lead to disturb by uncertainties in the model, the measured pressures obviously involve

310   significant errors, which reduces, in the analysis, the possibility of detecting leaks of lesser importance.

4 Conclusions and perspectives

Our research objective was the purpose of _discovery appropriate solution for detection and localization of leakages and estimation of the size of leakages for a water distribution system. The FAVAD parameters were optimized via a prediction algorithm, to constitute the core of our adopted procedure. The adopted approach necessitates a coupled hydraulic-GIS

315   interface by mean of the random forest algorithm.

This work helped to spot critical leaking points, and therefore contribute in the effort of physical loss reduction. Although, the detection results were not always accurate in term of space localization, the radius of search is reduced substantially, which make the detection rates during field campaigns more successful and less time-consuming.

Conflicts of interests: the authors declare that they have no conflict of interest.

320 **References**

Al-Ghamdi, A.S.: Leakage–pressure Relationship and leakage detection in intermittent water distribution Systems, J. Water Supply Res. Technol. Aqua, 60, 178–183, doi:10.2166/aqua.2011.003, 2011.

Alkasseh, JMA., Adlan, MN., Abustan, I, Aziz HA., Hanif, ABM.: Applying minimum night flow to estimate water loss using statistical modeling: A case study in Kinta Valley, Malaysia. Water Resour Manag., 27(5), 1439–1455, 2013.

325 Andrea, C., Giuseppe, C., Egidio, B., Nicola, G.: A New Method for Detecting Leaks in Underground Water Pipelines, IEEE Sensors Journal., 12(6), 1660 – 1667, https://doi: 10.1109/JSEN.2011.2176484 ,2011.

Babel, M.S., Islam, M.S., Gupta, A.D.: Leakage Management in a low-pressure water distribution network of Bangkok. Water Sci. Technol. Water Supply 2009., 9, 141–147, doi:10.2166/ws.2009.088, 2009.

Bakker, M.: Optimised control and pipe burst detection by water demand forecasting, (Doctoral dissertation). TU Delft, Delft,
330 Netherlands,2014.

Breiman, L.: Random Forests, Mach. Learn., 45(1), 5-32, 2001.

Brown, G.O.: The history of the Darcy-Weisbach equation for pipe flow resistance, In A. Fredrich and J. Rogers, editors, Proceedings of the 150th Anniversary Conference of ASCE, Washington D.C., pages 34–43. American Society of Civil Engineers, November 3-6 2002. Available at http://bioen.okstate.edu/Darcy/DarcyWeisbach/Historyof
335 theDarcyWeisbachEq.pdf, accessed 2 August 2011.

Chadwick, A., Morfett, J. and Borthwick M.: Hydraulics in civil and environmental engineering, Crc Press 2013.

Cobacho, R., Arregui, F., Soriano, J., Cabrera, E., Jr.: Including leakage in network models: An application to calibrate leak valves in EPANET, J. Water Supply Res. Technol. Aqua., 64, 30, doi:10.2166/aqua.2014.197, 2014.

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

EPA: Water Distribution System Analysis: Field Studies, Modeling and Management, U.S. Environmental Protection Agency. EPA/600/R-06/028, United States, 2005.

Farina, G., Creaco, E., Franchini, M.: Using EPANET for modelling water distribution systems with users along the pipes, Civ. Eng. Environ. Systems, 31(1), 36–50, 2014.

345 Farley, M., Trow, S.: Losses in Water Distribution Networks. IWA Publishing UK, 2003.

Farley, M., Wyeth, G., Ghazali, ZM., Istandar, A., Singh, S., Dijk, N., Raksakulthai, V., and Kirkwood, E.: The manager's nonrevenue water handbook: a guide to understanding water losses, Ranhill Utilities Berhad and the United States Agency for International Developement, Bangkok, Thailand,2008.

Jarrige, P.A.,and Gancel, G.: Method for detecting anomalies in a distribution network, in particular a water distribution
350 network, United States Patent Application. US2017/0205267A1,2017.

Ho, T.K.: Random Decision Forest. Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, 278-282, 14-16 August 1995.

Mirats-Tura, J.M., Jarrige, P.A., Meseguera, J., and Cembrano, G.: Leak detection and localization using models: field results, Procedia Eng. ,70 ,1157 – 1165, 2014.

355   Oasen.: Eerste successen met DMA's in voorzieningsgebied Oasen behaald, Available at :https://www.oasen.nl/nieuws/eerste-successen-met-dma-s-voorzieningsgebied-oasen-behaald, (accessed 20/02/20), 2015a".

Paquin, P., Babineau, D., Brissette, F., Leconte, R. : Développement d'une méthodologie de repérage des conduites d'aqueduc présentant des fuites, Can. J. Civ. Eng., 27, 151-159, 2000.

Pérez, R., Cugueró, M.-A., Cugueró, J., Sanz, G.: Accuracy assessment of leak localisation method depending on available
360   measurements, Procedia Eng., 70, 1304–1313, 2014a.

Pérez R., Sanz G., Puig V., Quevedo J., Cuguero Escofet M.A., Nejjari F., Meseguer J., Cembrano G., Mirats Tur J.M. and Sarrate R.: Leak localization in water networks: A model-based methodology using pressure sensors applied to a real, 2014b.

Sebbagh, k., Abdelhamid, S., Zabot, M.: Pre-Localization Approach of Leaks on a Water Distribution Network by Optimization of the Hydraulic Model Using an Evolutionary Algorithm, Proceedings., 2, 588,
365   doi:10.3390/proceedings2110588, 2018.

Rossman, L.A.: EPANET 2 User's Manual, US Environmental Protection Agency. Water Supply and Water Resources Division, National Risk Management Research Laboratory, Cincinnati, OH. 45268, 2000.

Rozental M. : La modulation de Pression Intérêts Bénéfices et Mise en œuvre, SUEZ Environnement, 86 pages, 2010.