Exploring a Spatial Dynamic Approach and Landmark Detection for Leakage/Burst Event Characterisation in Water Distribution Networks

Sotudeh Hoseini-Ghafari, Jorge Francés-Chust, Olivier Piller, David Ayala-Cabrera

iEMSs International Congress on Environmental Modelling and Software

11th International Congress on Environmental Modelling and Software
Brussels, Belgium, Ann van Griensven, Jiri Nossent, Stefan Reis (Eds.)
https://scholarsarchive.byu.edu/iemssconference/2022/

EXPLORING A SPATIAL DYNAMIC APPROACH AND LANDMARK DETECTION FOR LEAKAGE/BURST EVENT CHARACTERISATION IN WATER DISTRIBUTION NETWORKS

Sotudeh Hoseini-Ghafari[a], Jorge Francés-Chust[b], Olivier Piller[c], David Ayala-Cabrera[d]
[a]PhD student, CWRR-School of Civil Engineering, University College Dublin, sotudeh.hoseinighafari@ucdconnect.ie
[b]Business Development, Aguas de Bixquert, S.L., Xàtiva, jorge@abxat.com
[c]Senior Research Scientist, INRAE, UR ETTIS, olivier.piller@inrae.fr
[d]Ad Astra Fellow - Assistant Professor, CWRR-School of Civil Engineering, University College Dublin, david.ayala-cabrera@ucd.ie

Abstract: Extracting useful information from sensors that record water distribution network (WDN) data is essential to improve network performance, increase network preparedness and resilience, and advance network digitalisation. Due to the large volume of data generated, analysis of the pressure head requires advanced techniques to reduce dimensionality. While previous works were typically based on comparing hydraulic simulations and observed data, there is a lack of study on pattern recognition, a helpful method for event detection, localisation, and prevention. Since the number of metering devices and their operativity has a crucial role in the recognition of key patterns, a spatial evaluation of network behaviour (with a focus on resilience) is conducted in this study. Comparing the heatmaps leads to extracting key patterns (i.e., landmarks), which will be helpful for decision-makers to increase the preparedness by making arrangements against critical events and allow classification and prediction of the network behaviour. This paper focuses on recognising the possible landmarks in the network representing a key feature (particularly pressure) in the presence and absence of leakage through spatial analysis with the objective of dimensionality reduction. A dataset of incidents, leakage/burst events, and ordinary network operations were captured through sensors and expert knowledge in a WDN in Spain to obtain relevant information (in the form of landmarks) from them. Results were promising, recognising the patterns that characterise the network behaviour when influenced by leakage/burst events.

***Keywords**:* Data mining; Pressure sensor; Spatial-temporal analysis; Resilience and preparedness; Water distribution networks.

1 INTRODUCTION

Water distribution networks (WDNs) are known as complex critical infrastructures that provide water to consumers with a certain level of pressure. Leakage/burst in WDN, as a pressure-dependent event that affects the available pressure for consumption nodes, is a challenging issue for the network resilience assessment (Ayala-Cabrera et al. 2019). The potential negative impacts of such events can be minimised by extracting helpful information from sensors, which record WDN data. Extracting key information related to the network behaviour is a crucial way to improve network performance, increase network preparedness and resilience, and advance toward network digitalisation (Ayala-Cabrera et al. 2018, Taghlabi et al. 2020).
Several leakage detection, localisation, and prevention studies have been conducted to identify the anomalous behaviour of the network by comparing the data obtained from sensors with simulated pressure data (Taghlabi et al. 2020, Marzola et al. 2022). In this sense, advancing in the systematic

characterisation of this behaviour, through e.g., pattern recognition techniques, can increase the efficiency of the classification and prediction process. All of this information is essential for decision-makers to increase the network's preparedness by providing decision criteria to make the necessary arrangements against potential critical events.

Landmarks are the significant reference points representing properties of a shape, environment, and road, among others. For example, Alghani (2021) proposed machine learning techniques to detect and track keypoints from a human face. Landmarks are the key points to measuring behavioural characteristics of something. As given by Kim et al. (2021), detecting landmarks makes it possible to obtain a key point of the face. Landmarks and the connectivity between them can be used to capture the properties of a shape (Ibragimov et al. (2014). Yesiltepe et al. (2021) considered landmarks as not only the objects themselves but also their relationship to their surroundings. Thus, they should contrast with their background or have a precise shape or another specific characteristic that makes them prominent. Elements of the network are spatially correlated (e.g., nodal demands and pressures., etc.) (Jahanpour, 2018). So, spatial knowledge helps recognise these correlation relationships. In this paper, we explore the identification of the landmarks, in WDNs, as a potential tool that would incorporate spatial knowledge into the process of network characterisation. This landmark exploration can facilitate the identification of specific characteristics of the network that represent the behaviour of the network at specific operating conditions and at a particular time. Focusing on landmarks instead of analysing all the elements would reduce the dimensionality since landmarks are expected to provide sufficient information to analyse the network.

This paper aims to extract key patterns by a temporal-spatial evaluation of network behaviour (with a focus on resilience) using pressure head data from sensors. It is proposed that extraction of these patterns will make it possible to recognise landmarks, which allow leakage detection, localization, and prevention, among others of water leakage events. With the proposed approach, events can be traced even when there is a lack of either historical data from sensors or records of utility experts. In addition, the information given by sensors can help validate the analysis. If the history of an event is missed, temporal-spatial analysis of pressures (and other parameters) can be practical. With a dynamic spatial approach, landmark detection was conducted considering ordinary network operation and operation under leakage.

2 METHODS

As mentioned above, this paper seeks to capture the behavioural patterns of WDNs in operating conditions considered as normal/ordinary (without any abnormal event) and during the occurrence of an abnormal event (in this case, water leakage event). This work is based first on pressure heads (which will be called pressure in this paper) obtained by sensors and incorporates both a temporal and spatial analysis for this data. In this sense, this paper proposes a methodology that allows for recognising potential landmarks from the pressure data available in the network in a specific period. These landmarks are the basis for characterising the network behaviour and will allow, among others, to make advances in detection, localisation, and prevention of leaks. These are essential components to increase the network preparedness to face future leak events and prepare the water utility for other types of events with more severe consequences (e.g., droughts). The proposed methodology consists of three steps: (1) generation of a spatial distribution matrix, (2) pressure-based anomaly indicator generation, and (3) migration process. Details of these steps are described in this section.

2-1 Generation of a spatial distribution matrix

Let us consider the elements associated with pressure for a WDN as a set the nodes n_n (which can be tanks, reservoirs, or consumption nodes). The nodes that are pressure sensors (n_s) can be an indexed subset from the set of all nodes ($n_s \subset n_n$). The nodes have different attributes as the coordinates x and y; and pressure. For pressure sensors, it can be defined as $sx = \left[sx_1, \dots, sx_{n_s}\right]$ for x coordinate; $sy = \left[sy_1, \dots, sy_{n_s}\right]$ for y coordinate; and $sp = \left[sp_1, \dots, sp_{n_s}\right]$ for pressure.

A homogeneous mesh was created for both the x and y coordinates. Each of these meshes contained all the network nodes and had a specific stride Δx and Δy. The vector for coordinates x and y indicates all the elements desired in the new mesh (X_q and Y_q) and these are defined as $dx =$

$\{x_{\min} - a_{\mathrm{x}}: \Delta x: x_{\max} + a_{\mathrm{x}}\}$ and $dy = \{y_{\min} - a_{\mathrm{y}}: \Delta \mathrm{y}: y_{\max} + a_{\mathrm{y}}\}$. Where $x_{\min}$ and $y_{\min}$; and $x_{\max}$ and $y_{\max}$ represent the minimum and maximum coordinates for all nodes, respectively. a_{x} and a_{y} correspond to arbitrary numbers placed to extend the final window beyond the minimum and maximum coordinates. dx and dy have a total of elements n and m; respectively. MATLAB's griddata function (f), for scattered data was used to construct the pressure matrix, P (of size m × n), for each evaluated time, see (1). m and n represent the resolution for the x-axis and y-axis of the spatial coordinates, respectively. Matrix P was built from the nodes with sensors that were operational at the specific time evaluated.

$$P = f(sx, sy, sp, X_q, Y_q) \tag{1}$$

2-2 Pressure-based anomaly indicator

To better understand the pressure response of the network to anomalies (in this case, water leaks), a pressure-based anomaly indicator was proposed by Hoseini-Ghafari et al. (2022). This indicator was obtained using a temporal-spatial analysis of the system pressure (pressure obtained through sensors). Based on the pressure matrix P, it is possible to create a matrix of maximum pressures, $P_{\max}$, for each leakage case and a given period, $t = \{1,2, \dots, T\}$ with T as the total time evaluated, see (2). It can serve as the basis for the generation of a pressure-based anomaly indicator R, see (3). In this indicator, it is assumed that the highest pressure value (of each cell of P) can be a reference point to compare with the current evaluated pressure.

$$P_{max\,i,j} = \max\left(p_{t_1\,i,j}, p_{t_2\,i,j}, \dots, p_{t_T\,i,j}\right) \tag{2}$$

$$R_{t\,i,j} = \frac{P_{t\,i,j}}{P_{max\,i,j}} \tag{3}$$

where, $p_{t\,i,j}$ is the pressure in the i-th row ($i = 1{:}\,m$) and j-th column ($j = 1{:}\,n$) in the matrix at time step t, $P_{max\,i,j}$ is the maximum pressure during the total period (T) in each cell (i, j) of the matrix. $R_{t\,i,j}$ refers to the matrix of anomaly indicator constructed for each cell at each time step.

2-3 Migration process

According to the observations, the pressure-based anomaly indicator, (3), allows for recognising changes in the network behaviour from the ordinary state to the state affected by leakage. This indicator could better represent the behaviour of the network if the dependencies among the different elements of the evaluation system could be considered in the analysis. In other words, this consideration could mitigate the effect of pressure behaviour at each point and requires an implementation of pressure from one point to another. (4) referred to as the matrix of behavioural indicator, M, which is proposed to take into account the anomaly indicator affected by the surroundings for each cell at each time step. Obtaining M, is an iterative process that allows the migration of information towards a central point (in this case, each cell of R). In the first iteration, the central cell captures the behaviour of its eight closest neighbours. With each subsequent iteration, the central cell can obtain information from a wider radius of neighbours. Figure 1 shows a schematic representation of this idea to condense the information about the behaviour of the entire network in particular preferred points.

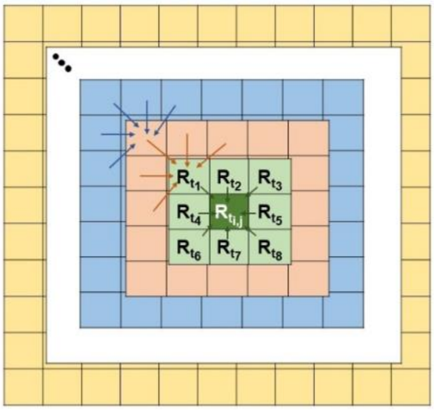

Figure 1. Dependencies between different points in the network.

As mentioned, capturing the coherence of the entire network's behaviour over a relatively large spatial extent can be conducted through successive iterations of the information migration (100 iterative/migration steps were explored in this paper). This is to obtain a stable relationship, which means that each cell absorbs the effect of as many cell perimeter layers as iterations are performed. In (4), value of each cell in each iteration is calculated by the average value of the eight cells around plus the cell itself. The closer a layer is to each specific point, the stronger the effect of corresponding cells to that point.

$$M_{t,i,j} = \frac{\left(\sum_{k=1}^{k=8} R_{t,cell_k}\right) + R_{t,i,j}}{9} \tag{4}$$

where $M_{t_{i,j}}$ is the matrix of behavioural indicator as an improved version of $R_{t_{i,j}}$ that mitigates the effect of pressure from surrounding cells. $cell = \{(i+1, j-1), (i+1, j), (i+1, j+1), (i, j-1), (i, j+1), (i-1, j-1), (i-1, j), (i-1, j+1)\}$; represents the pairwise of each neighbouring cell. k is the indices of the elements in $cell$; $k = \{1, \dots, 8\}$.

3 CASE STUDY, DISCUSSION

In this section, a real WDN, working under normal and abnormal/degraded operating condition has been selected to apply the proposed methodology.

3-1 Case study

The network selected in this study corresponds to a small-size utility network (Figure 2) located in Spain. The reason for choosing this network was the availability of data from both sensors and the availability of the operators to generate a structured database of the incidents that occurred in the network, including information on the leak detection and repair process. The model consists of 146 demand nodes, each node responsible for delivering water to many consumers (mainly houses), 212 links (40 km), two pumping stations, two reservoirs, and four tanks. There were 23 pressure sensors (recording pressure values every 15 minutes) in the network with different working conditions at different times (i.e., it is worth mentioning that the sensors were not always operative or, in some cases, relocated). Eight random cases of short-term leakage events (see Table 1) with various characteristics such as duration, location, cause, intensity, and available sensor data were selected to implement how effective is the recognition of patterns of pressure values in the whole network (see Figure 2). Maximum pressure was considered a basis for constructing an anomaly indicator, which was used to determine the network behavior to leakage.

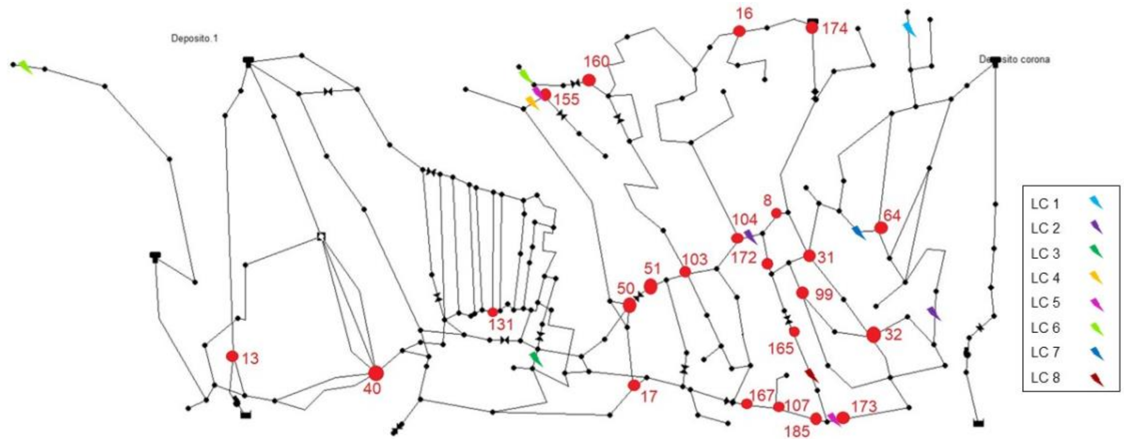

Figure 2. Case study network with the location of pressure sensors and leaks.

Table 1. Status of pressure sensors for the leakages.

Leakage ID	LC 1	LC 2	LC 3	LC 4	LC 5	LC 6	LC 7	LC 8
Date	2021-08-06	2021-08-09	2021-08-23	2021-09-20	2021-10-25 (two overlapping leaks)	2021-11-03 and 2021-11-04 (two overlapping leaks)	2021-11-15	2021-11-23
Number of sensors in operative conditions	11	11	11	11	13	11	15	16

3-2 Dataset

The dataset used in this paper includes historical data from pressure sensors (dataset 1) and utility expert knowledge (dataset 2; e.g., type of leakage, cause, location, detection, and repair information). Leakage analysis was initially conducted through a data-driven method via sensors' time-series data. Then leakages were temporally labelled (as no event/ordinary operation and event/abnormal operation) by an event-driven approach from the records by utility expert knowledge. Removing outliers in the data was conducted through manual inspection and collaboration with the system operator.

3-3 Spatial distribution matrix of the pressure

The pressure matrix, P (of size 74×33), was created for each evaluated time following the method mentioned before (see Section 2). The matrix P was built from the nodes with sensors that were operational at the specific time evaluated. A homogeneous mesh was created for both the x and y coordinates. Each of these meshes contained all the nodes of the network. As results of multiple iterations, the selected mesh stride for both Δx and Δy corresponds to 100; a_{x} and a_{y} are equal to 100 based on the coordinates of this case study.

3-4 Pressure-based anomaly indicator and migration process

Figure 3 shows the patterns of M during the studied period for each case, where x-axis shows time and y-axis represents the cells' location in the matrix.

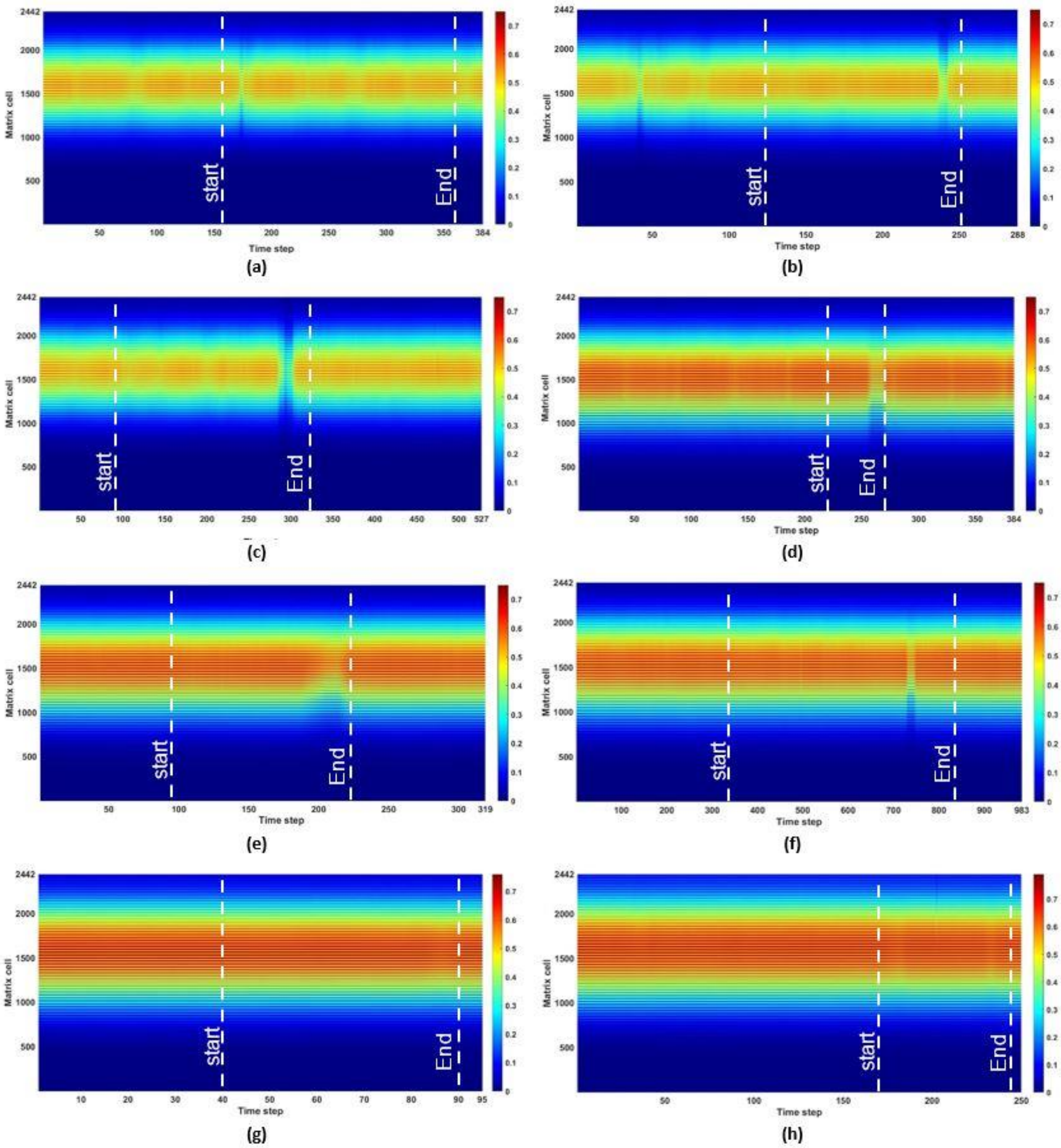

Figure 3. Behavioural indicator (M) for (a) LC 1, (b) LC 2, (c) LC3, (d) LC 4, (e) LC 5, (f) LC 6, (g) LC 7, and (h) LC 8.

Figure 3, indicates the changes in network behaviour during the period in which leakage impacted. Considering sudden or frequent changes of behaviour outside the abnormal period, there can be discovered some issues such as undetected leakage (for example, in Figures 3(b) and 3(d)). It is suggested that flow analysis and other potential parameters can help identify the meaning of these abnormalities. Another advantage of this type of analysis is the independency of the indicator to the number of available sensors. Whereas available sensors were fewer in LC 1, LC 2, LC3, LC 4 and LC 6 and higher in LC 7 and LC 8, the possibility to recognise patterns was not affected by this factor.

For each case, matrix M was flattened for ordinary (Figure 4(a)) and abnormal (Figure 4(b)) periods, individually. The overall behaviour of the network was achieved by the mean values of M for the ordinary state of each leakage case. The resulted M_{avg} (landmark) was sorted from lower to higher and shown in figure 4. It should be mentioned that the location number was not necessarily the same for different cases. It means that the x-axis in Figure 4 presents the cell locations not ordered by cell ID but by its corresponding M_{avg} . As shown in this figure, the pattern of ordinary and abnormal are not similar. The possibility of distinguishing a network with or without leak is promising by comparing the curve with the landmark, which can be helpful for leakage detection and prediction. Another interesting point is that the ordinary curves, shown in Figure 4(a), show the overall improvement in the network performance (based on the defined indicator) by date. One possible reason can be, for example, replacing the affected pipe, which might have also led to solving other problems; or it can be relevant to other factors and uncertainties that change during the time (such as leakage severity, location, demand pattern flow., etc) since lower (i.e., LC 1-Ord) and upper (i.e., LC 8-Ord) curves belong to earlier and later dates of occurrence in the year, respectively.

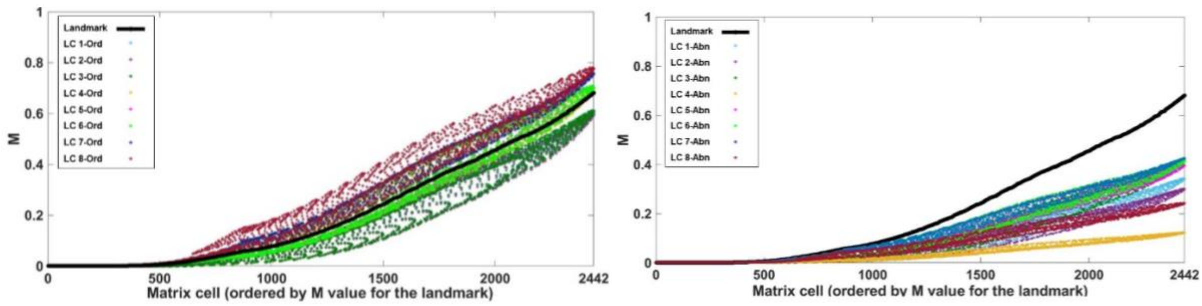

Figure 4. Landmark (black curve) compared with flattened matrices M for: (a) ordinary operation and (b) abnormal state.

4 CONCLUSIONS AND RECOMMENDATIONS

In this paper, we explored network behaviour to pressure for a dataset including eight cases of leakage events and the ordinary operation of a real WDN in Spain. Sensor records and expert knowledge provided the data. A pattern recognition technique was proposed to analyse the spatial behaviour of the network pressure in the condition of a limited number of available sensors. The results illustrated the effectiveness of intervention of pressure of the points that have strong dependencies to each point, which is necessary to capture the behaviour of the pressure head as a determining hydraulic parameter with/without leakage. The results were promising, recognising the patterns of pressure head values throughout the network. It was observed that the network behaviour would allow recognising landmarks when a leakage/burst event influences the network.

The output of this preliminary study would be advantageous to develop research studies in many aspects, such as:

- If any sensor fails or is relocated, it is still possible to identify an abnormal incident in the network by spatial analysis. It means that the reflection of an event would be independent of only one specific sensor and will be obtained through the extracted patterns.
- The ability to extract relevant patterns (i.e., feature maps) from the preliminary results of the pressure head heatmaps allows for appropriate detection of landmarks.
- It is possible to recognise the critical areas in the network to a specific parameter with/without leakage. Many factors can be considered to make the best decisions to improve the preparedness of WDN. For example, a delay in leakage detection (as an absorptive phase of resilience) might negatively impact the pressure of the entire network depending on the affected part. Developing this approach could help identify the potential landmarks.
- This analysis with spatial dimension can be improved by including parameters such as flow.
- The behavioural indicator could support the decision-making process regarding the implementation/deployment of actions likely to mitigate the effects of the event. In future research, we will investigate how to anticipate future events by increasing network preparedness, being proactive in preventing the occurrence of an event, and/or responding more quickly to events.
- Intelligent data analysis tools are recommended for a comprehensive study of influencing parameters for this approach.

ACKNOWLEDGMENTS

This work has been developed under the support of the Scholarship granted by the University College Dublin for the first author and the Ad Astra Start Up grant for the last author.

REFERENCES

Alghani Khalid, I., 2021. Face Landmark Detection using Python (https://towardsdatascience.com/face-landmark-detection-using-python-1964cb620837)

Ayala-Cabrera, D., Piller, O., Deuerlein, J., Herrera, M., 2018. Key Performance Indicators to Enhance Water Distribution Network Resilience in Three-Stages. Water Utility Journal.19, 79-90.

Ayala-Cabrera, D., Piller, O., Herrera, M., Gilbert, D., Deuerlein, J., 2019. Absorptive Resilience Phase Assessment Based on Criticality Performance Indicators for Water Distribution Networks. Journal of Water Resources Planning and Management. 145 (9), 04019037.

Hoseini-Ghafari, S., Francés-Chust, J., Piller, O., Ayala-Cabrera, D., 2022. Analysis of online pressure for resilience phase characterisation of leakage/burst events. In: 2nd International Joint Conference on Water Distribution System Analysis & Computing and Control in the Water Industry (WDSA/CCWI). (*Under review*).

Ibragimov B, Likar B, Pernuš F, Vrtovec T. 2014. Shape representation for efficient landmark-based segmentation in 3-d. IEEE Trans Med Imaging. 33 (4), 861-74.

Jahanpour, M., 2018. Pressure Sensor Placement for Leak Diagnosis under Demand Uncertainty in Water Distribution Systems, Ph.D. thesis, University of Waterloo, Ontario, Canada.

Kim, T., Mok, J., Lee, E., 2021. Detecting Facial Region and Landmarks at Once via Deep Network. Sensors. 21, 5360.

Marzola, I., Mazzoni, F., Alvisi, S., Franchini, M., 2022. Leakage Detection and Localization in a Water Distribution Network through Comparison of Observed and Simulated Pressure Data. Journal of Water Resources Planning and Management. 148 (1), 04021096.

Taghlabi, F., Sour, L., Agoumi, A., 2020. Prelocalization and leak detection in drinking water distribution networks using modeling-based algorithms: a case study for the city of Casablanca (Morocco), Drink. Water Eng. Sci. 13, 29-41.

Yesiltepe, D., Conroy Dalton, R., Ozbil Torun, A., 2021. Landmarks in wayfinding: a review of the existing literature. Cogn Process. 22, 369-410.