

Article

An Automated Geographical Information System-Based Spatial Machine Learning Method for Leak Detection in Water Distribution Networks (WDNs) Using Monitoring Sensors

Doha Elshazly ¹, Rahul Gawai ¹ , Tarig Ali ^{1,*} , Md Maruf Mortula ¹ , Serter Atabay ¹  and Lujain Khalil ² 

¹ Department of Civil Engineering, American University of Sharjah, Sharjah P.O. Box 26666, United Arab Emirates; g00058601@alumni.aus.edu (D.E.); rgawai@aus.edu (R.G.); mmortula@aus.edu (M.M.M.); satabay@aus.edu (S.A.)

² Department of Computer Science and Engineering, American University of Sharjah, Sharjah P.O. Box 26666, United Arab Emirates; g00082632@alumni.aus.edu

* Correspondence: atarig@aus.edu

Abstract: Pipe leakage in water distribution networks (WDNs) has been an emerging concern for water utilities worldwide due to its public health and economic significance. Not only does it cause significant water losses, but it also deteriorates the quality of the treated water in WDNs. Hence, a prompt response is required to avoid or minimize the eventual consequences. This raises the necessity of exploring the possible approaches for detecting and locating leaks in WDNs promptly. Currently, various leak detection methods exist, but they are not accurate and reliable in detecting leaks. This paper presents a novel GIS-based spatial machine learning technique that utilizes currently installed pressure, flow, and water quality monitoring sensors in WDNs, specifically employing the Geographically Weighted Regression (GWR) and Local Outlier Factor (LOF) models, based on a WDN dataset provided by our partner utility authority. In addition to its ability as a regression model for predicting a dependent variable based on input variables, GWR was selected to help identify locations on the WDN where coefficients deviate the most from the overall coefficients. To corroborate the GWR results, the Local Outlier Factor (LOF) is used as an unsupervised machine learning model to predict leak locations based on spatial local density, where locality is given by k-nearest neighbours. The sample WDN dataset provided by our utility partner was split into 70:30 for training and testing of the GWR model. The GWR model was able to predict leaks (detection and location) with a coefficient of determination (R^2) of 0.909. The LOF model was able to predict the leaks with a matching of 80% with the GWR results. Then, a customized GIS interface was developed to automate the detection process in real-time as the sensor's readings were recorded and spatial machine learning was used to process the readings. The results obtained demonstrate the ability of the proposed method to robustly detect and locate leaks in WDNs.

Keywords: geographical information system; spatial machine learning; leak detection; automation model; water distribution networks; pressor; flow; water quality; chlorine residual



Citation: Elshazly, D.; Gawai, R.; Ali, T.; Mortula, M.M.; Atabay, S.; Khalil, L. An Automated Geographical Information System-Based Spatial Machine Learning Method for Leak Detection in Water Distribution Networks (WDNs) Using Monitoring Sensors. *Appl. Sci.* **2024**, *14*, 5853. <https://doi.org/10.3390/app14135853>

Academic Editor: Norhisham

H. Bakhary

Received: 18 April 2024

Revised: 7 June 2024

Accepted: 17 June 2024

Published: 4 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Water is supplied in urban areas through a water distribution network (WDN). WDNs aim to deliver adequate clean water to consumers. Around 25% of the water produced by utility companies is lost during the distribution process every year as a result of theft, leakage, or metering errors, among which the main contributor is pipe leaks [1]. Water loss from faulty pipes can reach 50% of the supply, resulting in water and financial losses. Financial losses include the cost of raw water, its treatment, and delivery. Additionally, leakage deteriorates pipes in the form of erosion and pipe breakage, which can lead to damaged roads and building foundations.

There have been several studies on water leakage and leak detection techniques. The most commonly used leak detection methods include acoustic, infrared thermography, pressure and flow sensor signals, and ground penetrating radars [2,3]. These methods are point-based and rely on previous assumptions of leak locations. For this reason, they are limited in their ability to efficiently handle an entire large WDN since most experiments have been conducted on pilot-scale laboratory models. Some concerns still relate to the lack of time efficiency and the necessity of skilled labourers on-site. Moreover, the high dependency on the use of data collected by pressure sensors was proven to be inaccurate because there are several factors affecting pressure readings apart from leakage.

WDNs are often built with flow, water quality, and pressure sensors for continuous monitoring. Leaks within the WDN are often accompanied with an instant pressure and flow drop, with a deterioration in water quality as a result [4]. Extensive research has been carried out on the use of pressure drops in WDNs for leak detection due to the wide availability of low-cost sensors and ease of installation [5]. Sala and Kolakowski proposed an integrated leak detection method complemented by an electronic system that aids in the remote transmission of pressure data based on global systems for mobile communication (GSM) protocols [6]. However, the application of this system was limited to small-scale networks and its feasibility on large WDNs was not tested. Another study was conducted to develop an in-pipe leak detection system to spot leaks based on radial pressure gradients. The method was tested on a lab-scale prototype simulating real leaks and was found to accurately identify leaks using pressure readings [7].

Another study used the Head Loss Ratio (HLR) to detect leaks [8]. Although successfully proven theoretically, this study was applied to simulated small WDNs in EPANET Version 3 software. Hence, the method's suitability is still questionable in large, real WDNs [8]. Another study explored the angle method, which was compared with the least-squares optimization and the correlation methods [9]. The method was purely theoretical and mathematical in its formulation and eventually applied to a model Hanoi network. It was found to be 75% efficient in locating leaks using pressure data. However, its applicability in real, large WDNs is questionable as the extent of uncertainty in a large WDN cannot always be mathematically formulated with the approach of this study. A statistical classifier model was derived by another study while considering pressure irregularities in the network to improve the accuracy of the leak localization [10]. The proposed method was further tested by the Hanoi network benchmark. This method was compared with the angle method and was found to be 80% accurate compared to the 50% accuracy of the angle method [10]. Similar to other studies, no attempt was made to investigate the applicability in a large WDN. A state wireless pressure detection sensor was developed to identify leaks based on the measurement of relative indirect pressure changes in plastic pipes [11]. This method was capable of performing continuous monitoring without operator intervention. However, the applicability in large WDNs is still unexplored. Similarly, Asgari and Maghrebi examined the feasibility of using nodal pressure data to locate leaks by extracting several pipes and determining the nodal pressure patterns across the lines [12]. The technique was not sensitive to the size of simulated leak discharge, even though it was found efficient. EPANET was used to investigate leak detection formulations to locate leak positions and their sizes with the help of pressure readings at each pipe node [13]. The average detection accuracy was 82.45% for k-nearest neighbours (KNNs), 66.31% for the naive Bayes network, and 48.90% for decision tree networks.

A submersible optics-based pressure sensor was proposed to identify, spot, and quantify leaks based on hydrostatic and pressure transient responses [14]. However, the sensitivity of this sensor was found to be lower than those of traditional sensors. This study mostly focused on the sensor development, not necessarily for a large WDN. Using anomaly detection algorithms, pipe failures were successfully detected based on data obtained from external pressure and temperature sensors [15]. However, it was a point-based leak detection system. Another study on the point-based leak detection system was investigated and successfully used the pressure residual vector method to locate the leaks, taking into

account the pipe material and external factors like soil load [16]. The technology was tested in a hypothetical model in EPANET.

A study investigated the applicability of sum of squared errors (SSE) optimization functions and successfully detected leaks with various scenarios of leak sizes using simulated network pressure [17]. Another study used sensitivity analysis integrated with mass and energy conservation equations to detect the variations in pressure and flow data in the network in order to aid leak localization [18]. This methodology was capable of detecting water leaks of up to 30 L/min [18]. A different methodology that utilizes pressure data was successfully accompanied by an iForest algorithm to obtain accurate results [5]. The approach was investigated theoretically in a small network in EPANET. The Value of Information (VOI) and Transinformation Entropy (TE) methods, in concert with an optimization model, were able to identify leaks using pressure readings, as well as localize pressure sensors for more accurate results [19]. Another study utilized pattern recognition and the clustering of transient pressure signals to train Artificial Neural Networks (ANNs) for leak localization and estimation. It was found to be successful for real WDNs [20]. Similarly, a real-time detection method could detect pressure irregularities within the network by the use of an unsupervised learning anomaly detection method with an accuracy of 92% and 94% for 2013 and 2017 datasets, respectively [21]. Shao et al. employed a temporal-based leak detection method with the help of multiple pressure sensor placements. They investigated the difference between real-time and predicted model data in order to enhance the detection capability [22]. Sun et al. used machine learning methods including linear discriminant analysis (LDA) and neural networks (NNETs) in integration with Bayesian temporal reasoning with results reaching up to 80% accuracy [23]. A similar study employed the random forest algorithm and performed model training and testing with the help of available pressure data with 96% accuracy [24]. Gungor et al. successfully applied the use of pressure reduction valves (PRVs) to monitor and detect leaks with the help of an existing supervisory control and data acquisition (SCADA) system, GIS, and customer information system (CIS) databases [25].

Few studies focused on the flow sensor data to determine the leak locations within a network. Mulholland et al. employed a linear programming algorithm based on network hydraulics to identify the leak spots using sparse flow data [26]. However, this study identified the ability to handle large datasets at once, being an advantage. Another real-time leak detection technique using the traditional flow balance integrated with the minimum night flow (MNF) approach was used to provide flow thresholds for accurate leak identification. This method allowed the reduction in future leaks by 36% [27]. A similar study used the computed leakage rate, components, and reduction potential based on MNF to reduce leak frequencies [28]. Another detection methodology utilized nonlinear partial differential equations with the help of flow data for modelling and irregularities detection within the network for leak localization [29]. The method was tested on a pilot-scale model and provided reasonable leak location estimations. Furthermore, another study performed extensive simulations using optimization models that use available flow sensor data to locate possible leaks [30]. It was observed that the majority of the studies using flow data presented low accuracy.

Several studies have explored the feasibility of utilizing the pressure and flow meter data to detect leaks. Nevertheless, accurate identification of the leak locations can be challenging when relying solely on separate parameters (flow/pressure) due to the presence of fluctuations along the network length caused by other factors. Additionally, research on combining those parameters in addition to water quality parameters is still in its infancy. The use of machine learning techniques for identifying leak locations using these parameters is nonexistent. There is an apparent necessity to integrate this innovative pressure, flow, and chlorine residual with machine learning-based techniques to improve the accuracy of leak detection and avoid false positive scenarios.

Machine learning models that are most suitable for anomaly detection tasks have been explored for this paper. Tornyeviadzi et al. [31] performed a comparative study on different

anomaly detection models on the LeakDB dataset, a benchmark dataset for leak detection. After comparing model performances, the authors concluded that proximity-based models performed best on the LeakDB dataset, with the Local Outlier Factor (LOF) model achieving a high entropy weight TOPSIS value of 0.9943. This approach, however, is semi-supervised and does not consider the fact that data collected from WDNs are unlabelled. Another study by Alghushairy et al. [32] reviewed the use of different LOF algorithms for big data streams, showcasing the model's efficiency in various applications including intrusion detection, fraud detection, and medical applications. Desmet and Delore [33] used two anomaly detection models for leak detection in compressed air systems, the LOF and Autoencoders. Despite the complexity of the Autoencoder, the LOF was able to perform better with an Area Under Curve (AUC) value of 0.94. This study, however, used synthetic data to test their algorithms due to the lack of available real data in their domain.

Along with the LOF, Geographically Weighted Regression (GWR) is a machine learning technique that we found suitable for leak detection in WDNs [34]. GWR has been used in this study with the dataset provided by the Sharjah Electricity and Water Authority (SEWA) to (a) integrate the three leakage predictors (flow, pressure, and chlorine readings in the WDN) given their spatial variability throughout the WDN, and (b) to automate the leak detection workflow. This approach is particularly relevant for large WDNs. To correctly use a GWR model, it is recommended to first test the ordinary least squares (OLS) linear regression to check if the assumptions are met. GWR first requires testing if the residual distribution is not spatially autocorrelated [35,36], and then to utilize the OLS predictors to build a GWR model [37,38]. This model was used in this study to help with the workflow automation process and improve the accuracy of predicting leak areas using the sensor recordings. This study aimed to develop an automated GIS-based leak detection method for WDNs using GWR with flow, pressure, and chlorine readings as conditioning factors. Spatial variability maps were generated with the help of the GIS. Using these maps, pipelines with a significant drop in flow and pressure are identified and compared with the areas where water quality drop is located, yielding increased leak location identification accuracy. A customized GIS interface was developed to automate the process in real-time.

2. Methodology

2.1. Study Area

This study was carried out on part of the WDN of the city of Sharjah, UAE. The Sharjah WDN is shown in Figure 1 below. The network details were obtained from the SEWA. This network was selected as it represents a large WDN and is suitable for the application of this model. The pipe network consists of approximately 70,000 nodes, and the pipeline extends to about 3152 km through the entire city. Water in Sharjah is obtained from ground water and desalinated seawater and distributed to the pipe network from three main treatment plants: Layyah, Hamriyah, and Sajja. Chlorine is further injected into the network as a secondary disinfectant. The WDN is subdivided into districts, each of which contains one or multiple pressure, flow, and water quality sensors. Sensor readings are recorded from the 13 monitoring zones daily to ensure the best quality of supply and to allow for anomaly detection. In this study, the proposed method was tested on data collected by SEWA over the span of a year.

2.2. Water Pressure Analysis

SEWA's WDN is monitored by 13 pressure sensors placed in 13 different zones, each of which records the pressure in 15 min intervals. Average daily pressure values were calculated to match the daily flow and water quality data provided by SEWA. Pressure drops in the system are influenced by many factors; those fluctuations remain acceptable up to a certain limit [39,40]. If the pressure drops below this value of 1.7 bar, it is indicative of possible anomalies. In this study, this was the threshold used for analysis.

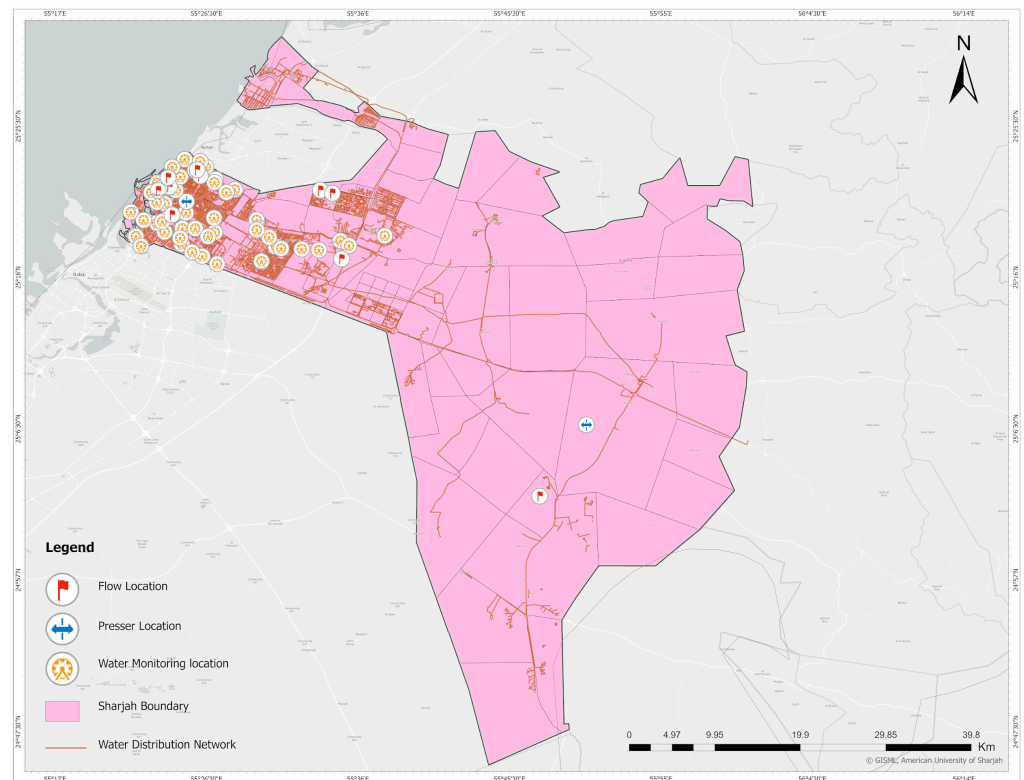


Figure 1. Water distribution network of Sharjah.

2.3. Water Flow Analysis

The WDN considered in this study is monitored by the 13 flow sensors. The net flow can determine the water flowing through the system, identified as the difference between the sum of water passing through the inlet and outlet points.

$$\text{Net flow} = \text{Inlet flow} - \text{Outlet flow} \quad (1)$$

The daily inflow capacity provided to the system is driven by consumer demand. As the leak event occurs, the water flow through the pipe is reduced, resulting in a decrease in the outlet flow sensor readings. Hence, in addition to the effluent, the net flow in this case will account for the water lost to leaks as well. This can be rewritten as follows.

$$\text{Net flow} = \text{Inlet flow} - \text{Outlet flow} - \text{Leak flow} \quad (2)$$

The data provided by SEWA do not include any information regarding the leak flow. The proposed alternative is to pinpoint the areas with exceedingly low net flows as areas of suspected leaks. Based on the experiences of operators of SEWA, a threshold of low net flow was used as $1000 \text{ m}^3/\text{day}$.

2.4. Chlorine Residual Analysis

Chlorine is injected into the WDN at the desalination plants and is used as a secondary disinfection. Adequate chlorine concentrations along with a maintained pressure are essential to keep the water free of contaminants. As pipe walls break, the disturbances caused by pressure result in contaminant intrusion and reduced chlorine residual and water quality. The SEWA network contains 46 water quality monitoring points distributed throughout the 13 zones. Various water quality parameters (temperature, pH, chlorine) are often monitored to determine the status of the WDN, with residual chlorine being the most critical water quality indicator [41]. Any drops in chlorine levels below 0.2 mg/L would

be possible indicators of leaks in the network [42]. In this study, the chlorine residual of 0.2 mg/L was used as the threshold for leak detection.

2.5. Spatial Variability Maps

Spatial variability occurs when a quantity at different locations throughout an area exhibits different values. Spatial variability maps are raster images generated to display those variations across the surface. In this study, pressure and flow spatial variability maps were generated in ArcGIS using the inverse distance weighting (IDW) interpolation to spot the locations with relatively low readings. Pressure and flow dips within the distribution system are often caused by leaks, which deteriorate water quality as a result. To improve the accuracy of the results, chlorine residual values were examined at the suspected locations. The three parameters were combined using spatial variability maps to direct the user to the areas suspected of leaks. Using these maps, a customized GIS interface was developed to automate the proposed method which will enable the identification of possible leak locations during its execution.

2.6. Data Preprocessing

After interpolating, the resulting dataset consisted of 782 data points, each with its corresponding Longitude, Latitude, pressure, chlorine, and flow values. Initial analysis involved generating histograms to understand the distributions of the features, as shown in Figure 2. To avoid mistaking normal locations as leaks, any data point with pressure, chlorine, or flow values that are 3 standard deviations above their respective means were dropped. This was carried out following the assumption that values at the higher extreme are not representative of leaks, but rather are exceptional conditions to water pipes that are irrelevant to the analysis.

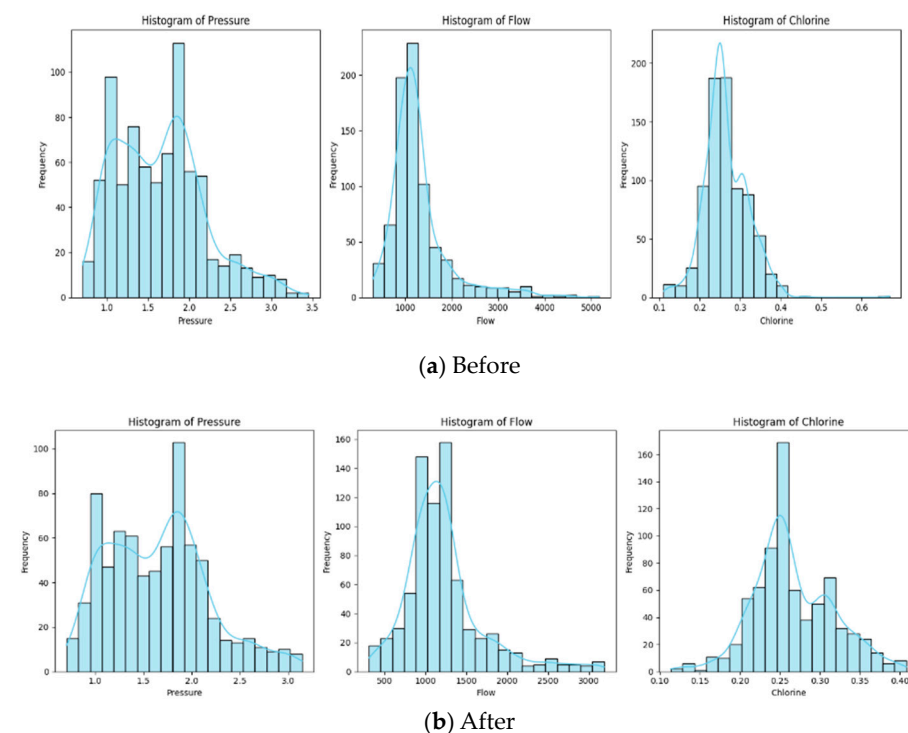


Figure 2. Pressure, flow, and chlorine before and after dropping extremes.

A feature, called 'ratio_pf', was created by dividing the pressure by flow for a better representation of the data points. This relationship could be more indicative of outliers. Further, it was important to scale or normalize the parameter values to avoid one feature overpowering the other. To determine the most suitable method of scaling, statistical tests were run to check the hypothesis that the features follow a Gaussian Distribution.

Shapiro–Wilk, Kolmogorov–Smirnov, and Anderson–Darling tests were run, and the results can be seen in Tables 1 and 2 below.

Table 1. *p*-Values of Shapiro–Wilk and Kolmogorov–Smirnov tests from all 4 features.

	Pressure	Flow	Chlorine	Ratio-pf
Shapiro–Wilk	8.12×10^{-12}	1.34×10^{-23}	1.93×10^{-8}	5.44×10^{-12}
Kolmogorov–Smirnov	0.0	0.0	5.11×10^{-216}	3.69×10^{-175}

Table 2. Statistic value of Anderson–Darling from all 4 features.

	Pressure	Flow	Chlorine	Ratio-pf
Anderson–Darling	5.461	26.021	6.831	6.086

Since the *p*-values of all features for both the Shapiro–Wilk and Kolmogorov–Smirnov tests were below 0.05, we could reject the null hypothesis. This was further proven by the Anderson–Darling statistic for all features being below the critical value of 0.783. This conclusion allowed us to normalize the parameters using MinMaxScaler, rather than StandardScaler. After normalization, the data were ready to be fed through the machine learning model.

2.7. Geographically Weighted Regression (GWR)

Geographically Weighted Regression (GWR) provides a local model for fitting a regression equation to every feature in the dataset. GWR incorporates the independent variables falling within the bandwidth identified based on the chosen Kernel type, which was the *k*-neighbours in this case (identical to the one used for the LOF model as explained later). GWR is a supervised machine learning method that allows the fitted regression model to vary spatially considering adjacency or local variations. A major parameter for the GWR is the Kernel function, which is used to set the distance between a point and its closest neighbour. In this study, the Gaussian Kernel function was used for its known robustness where the bandwidth corresponds to the standard deviation of the function. To run the GWR, a 70:30 data splitting was performed between training and testing. This was carried out using the `sklearn.model_selection.train_test_split` function in Python Version 3.11.

2.8. Local Outlier Factor Model

The LOF model is an unsupervised anomaly detection model that identifies anomalies in the data based on local neighbourhood densities. It is especially useful when the density of the data varies globally across the dataset. This makes the LOF especially useful in this case since the presence of a leak will affect the pressure, flow, and chlorine levels that are closer in distance to the suspected leak.

The LOF score of a data point is calculated to determine whether it is considered an outlier. The closer the LOF score is to -1 , the more “likely” it is to be an outlier. To calculate the LOF score of a datapoint, we first need to understand *k*-distance and *k*-neighbours. The *k*-distance of any point *A* is defined as the distance between point *A* and its *k*th nearest neighbour. The number of *k*-neighbours from point *A* is denoted as $N_k(A)$, where $N_k(A)$ is always less than or equal to *k*. The measure of distance used in this analysis is the Minkowski distance with $p = 4$, where *p* is the parameter that determines the order of the distance calculation. Since each data point in this dataset contains 4 features (pressure, flow, chlorine, and ratio_pf), the parameter *p* is set to 4. The Minkowski distance between two points *A* and *B* can be computed as follows:

$$Minkowski(A, B) = \left(\sum_{i=1}^4 |A_i - B_i|^4 \right)^{\frac{1}{4}} \quad (3)$$

The Reachability Distance (RD) between two points A and B can then be defined as the maximum of the k -distance of A and the distance between A and B :

$$RD(A, B) = \max\{k - \text{distance}(A), \text{distance}(A, B)\} \quad (4)$$

The Local Reachability Distance (LRD) of a point A is defined as the inverse of the average RD of A from its k -neighbour. The lower the LRD value, the further the closest cluster is to point A .

$$LRD_k(A) = \frac{1}{\sum_{X_j \in N_k(A)} \frac{RD(A, X_j)}{\|N_k(A)\|}} \quad (5)$$

Finally, the LOF score of a datapoint A is calculated as the ratio of the average LRD of its k -neighbours to the LRD of A :

$$LOF_k(A) = \frac{\sum_{X_j \in N_k(A)} LRD_k(X_j)}{\|N_k(A)\|} \times \frac{1}{LRD(A)} \quad (6)$$

Evaluating how well the model was able to cluster anomalies was performed using the silhouette score of the data. The silhouette score is a metric used to evaluate the quality of clusters generated by a model. It is used in this analysis as a standard to decide how well the model performs. The silhouette score $S(A)$ for a datapoint A is calculated as follows:

$$S(A) = \frac{f(A) - g(A)}{\max\{f(A), g(A)\}} \quad (7)$$

where $f(A)$ is the smallest average distance from A to points in a different cluster, and $g(A)$ is the average distance from A to other points in the same cluster. The silhouette score for the dataset can then be calculated as the average of all individual silhouette scores:

$$\text{Silhouette Score} = \frac{1}{N} \sum_{i=1}^N S(i) \quad (8)$$

where N is the total number of data points. The closer the silhouette score is to 0, the more overlapping there is between the clusters. An A value closer to -1 indicates possible misclustering of the data, whilst a value closer to 1 indicates more distinct clustering of the data.

2.9. GIS-Based Interface Workflow

A GIS-based interface was developed using model builder, combining the data measured to automate the leak detection process. An average daily value was generated, and the flow and pressure data were compared to their average monthly and minimum acceptable ranges, respectively. An IDW variability raster was developed for each of the two parameters to examine the parametric variation throughout the network. Subsequently, the areas showing pressure and flow readings below the threshold values were clipped out of the network and the two models were intersected and presented as potential leak locations. Finally, the residual chlorine levels at the locations nearest to the areas of intersection obtained in the previous step were inspected, and locations with values less than the acceptable range were confirmed as suspected leak locations. In the case of a leak event, the GIS interface would provide a visualization of the leak locations overlaid with the WDN and road layers. The following flowchart shows the workflow of the proposed GIS-based, water leak detection system (Figure 3).

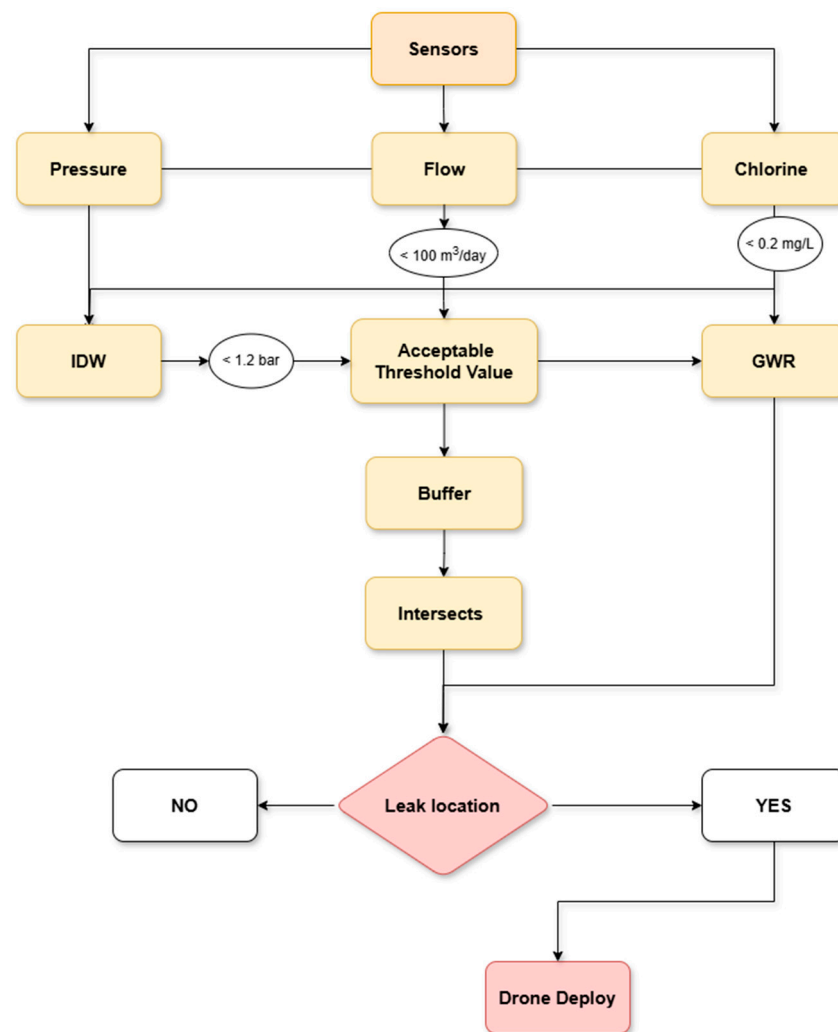


Figure 3. Proposed methodology flowchart.

2.10. ArcGIS Model Builder

The model builder module in ArcGIS Pro 3.0 was utilized to automate the leak detection workflow process. Model builder is a tool used to create and modify geoprocessing models. It displays the selected functions in a diagram format, chaining separate tools together in a manner that enables the output of one function to be the input of the subsequent one [43]. The flowchart in Figure 3 illustrates the methodology used in this study. The steps of the proposed method are shown below:

- Read the WDN sensor recordings of the three leakage predictors: pressure, flow, and chlorine.
- Apply the threshold values of the three readings.
- Use the buffer around the sensors' locations and intersect operators to assess the variability in the three predictors throughout the WDN.
- Use IDW to create raster layers for pressure, flow, and chlorine in ArcGIS Pro for the WDN.
- Use GWR to seamlessly integrate the three predictors and, at the same time, automate the leak detection workflow.
- Identify potential leakage locations.

A drone can then be deployed to the areas of the potential leakage locations determined by our proposed method. The IR images taken for the potential leakage locations would be processed to accurately identify the leak locations. The IR drone images were processed using our method in [44], which is based on multitemporal infrared thermography.

The pressure and flow data were inserted into the GIS model and compared with the thresholds. Subsequently, an IDW was created for each parameter to provide values to unmonitored locations. Following this procedure, a fishnet mimic of the area was created to select certain portions of the network using the query table tool (values below zero for pressure and above zero for flow). The vector fishnet was narrowed down, using the spatial join tool, to the locations suspicious of a leak, identified using the query table in the previous step. Each of the mentioned steps was applied to the pressure and flow individually. Next, the suspicious locations were merged using the intersect tool to identify the areas with drops in both pressure and flow. The same tool was later employed to select the parts of the network experiencing those drops. Finally, the chlorine values in the vicinity were identified by the intersect tool and were further compared with the minimum acceptable value using the query table to further narrow down the leak locations. The roles of each of the mentioned functions are discussed below.

3. Results

3.1. Model Performances

Fine tuning the LOF model to achieve the best possible silhouette score was performed by running different combinations of the *n_neighbours* and contamination parameters. The *n_neighbours* parameter specifies the number of neighbours considered when calculating the LOF score, while the contamination parameter defines the expected proportion of outliers in the dataset. The results of the applied grid search can be seen in Table 3.

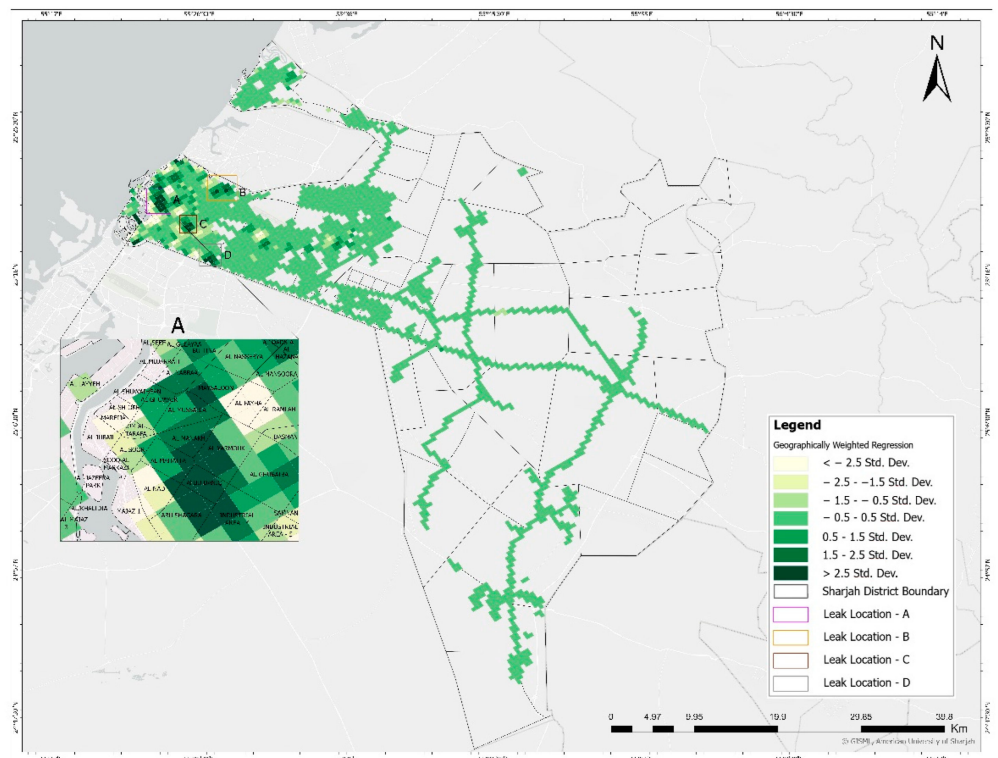
Table 3. Grid search for optimal contamination and *n-neighbours* parameters.

Contamination	n_Neighbours			
	2	5	10	15
0.005	0.0247	0.0688	0.2525	0.3478
0.010	0.0080	0.1050	0.2144	0.3404
0.050	−0.0476	0.0038	0.1480	0.1378
0.100	−0.0298	0.0091	0.0659	0.0774

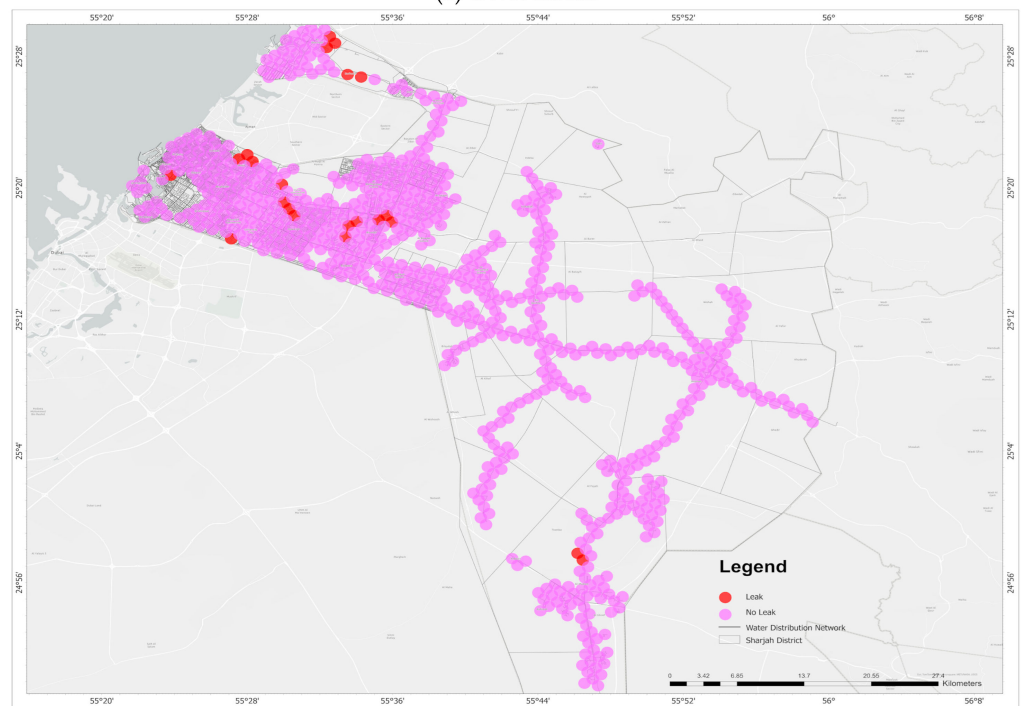
Despite the best silhouette score being observed at a 0.005 contamination rate and 15 *n_neighbours*, the model with a contamination rate of 0.01 was chosen, with its results shown in Figure 4b. This was decided due to the small number of data points existing in the dataset, with 0.005 limiting the number of anomalies to only eight data points. The more anomalies that are detected, the less likely it is for the model to misclassify an actual leak as normal.

The machine learning-model-predicted leak locations are shown in Figure 4a for the GWR model and in Figure 4b for the LOF model. Note that since leakage causes drops in the values of the three independent variables, positive standard deviations of the dependent variables were used to determine the possible leak locations. GWRs of pressure, flow, and chlorine show large numbers of sites with positive and negative residual values, indicating a lack of stationarity across Sharjah (Figure 4a). Moreover, these variables show similar sub-regions with high positive residuals in industrial areas of the region, leak locations A, B, C, and D, suggesting some parallel localized dependency on the high density of residuals. A geographic pattern was observed where residuals are stationary throughout most of the study area, with some clusters of high residual variation in the western part (Figure 4a), and an isolated patch of low residual variation in the western part of the study areas because there are fewer settlements. The obtained coefficient of determination (R^2) for the GWR model was 0.909. The main advantage of using GWR in this study is to explore the varying spatial relationships between the three parameters (the explanatory variable) and the leakage as the dependent variable. The results show that the relationship between the leakage (dependent variable) and the explanatory variables is spatially consistent. The high

degree of correlation indicates the suitability of the model. Figure 5 shows a comparison between the results of the GWR and LOF models, which showed an overall match of 80% in the dataset. Note that in the figure, a circle shows a match and a rectangle shows no match between the models, only for the visually clear leak locations. A match of 80% is also indicative of the effectiveness of this particular model.



(a) GWR model



(b) LOF Model

Figure 4. The predicted leak locations using the GWR and LOF models.

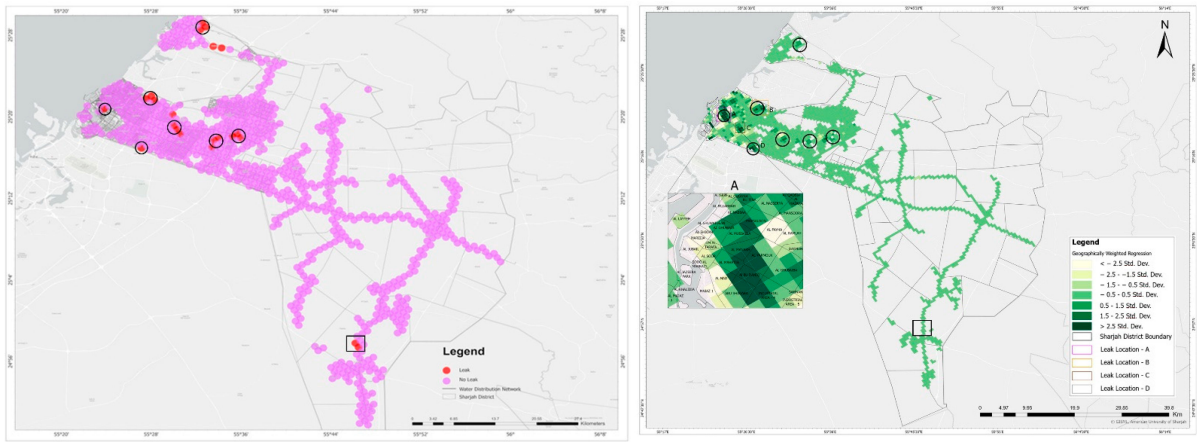


Figure 5. Comparison between GWR and LOF model results (circle: match; rectangle: no match).

3.2. GIS Interface

The automation process was carried out on ArcGIS Pro 3.0 utilizing the model builder tool. This process returns a series of visual outputs at individual functions. The first visual output was two IDW raster surfaces representing individual pressure and flow data points (sensor readings). Following this step, the locations selected to be indicative of leaks, based on numerical criteria, in each IDW were overlaid using the intersect tool and were subsequently intersected with the network layer as well. The third visual output was a capture of the chlorine monitoring points located in the vicinity selected in the previous steps. The chlorine values were further compared with the minimum acceptable limit and the values that appeared to be within the limit were discarded from the map. The final visual output of the automation process then returned a network area showing drops in all three parameters that can indicate the presence of a leak. Figure 6 below illustrates the series of outputs returned during the automation process. These different subfigures indicate the pressure and flow patterns throughout the WDN.

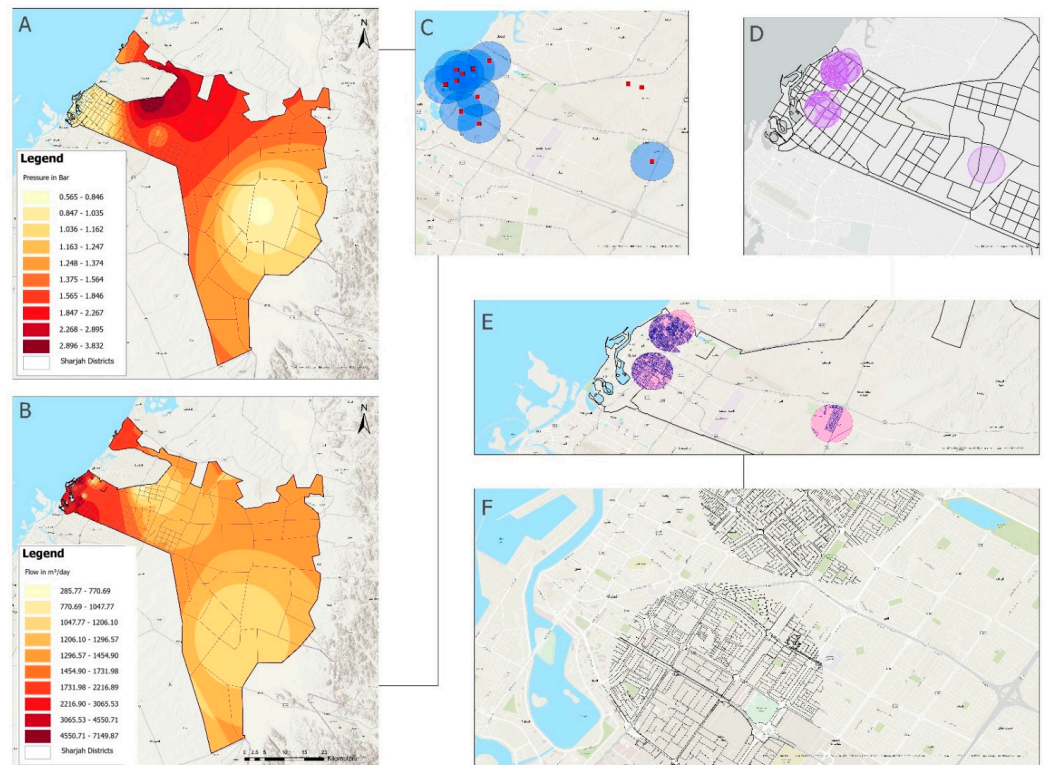


Figure 6. Series of GIS outputs—(A) pressure, (B) flow, (C) buffer, (D) intersect, (E,F) intersected WDN.

The IDW method was employed to generate density grids presenting the spatial distribution of pressure, flow, and water quality data throughout the study area. Three scenarios were carried out to test the efficiency of the automation model in leak detection. The first days of October, January, and July were selected to test the leak progression in water pipes throughout the seasons autumn, winter, and summer, respectively.

3.2.1. Autumn

For consistency, the first day of October was chosen in this scenario to test the automation model. Figure 7 shows the spatial distribution of pressure and flow over the network area obtained from the IDW. The cream colour in Figure 7 indicates the lowest observed values and the burgundy colour indicates the highest values. The pressure IDW indicated low pressure values (0.6 to 1.2 bar) observed at a few areas including Butina, Al Ghuwair, Industrial Area 4, Industrial Area 6, Al Ghaphia, Maysaloon, Al Fayah, and Al Nasserya. However, the remaining vicinities covering the Al Rahmaniya 1, Al Rahmaniya 3, Al Qadsia, Al Sabkha, and Barashi areas indicated normal pressure behaviour across its network (ranging from 2 to 3.75 bar). The IDW returned for the net flow data for the analysed day indicated a major increase covering the eastern sector of the network which included Al Rahmaniya 1, Al Rahmaniya 3, Butina, Maysaloon, Al Nasserya, Barashi, and Al Fayha, while the remaining locations indicated normal flow patterns (110 to 7380 m³/day). The individual pressure and flow drops were selected by the model (shown in brick red) and both the selected vicinities in the pressure and flow layers were intersected. Since pressure and flow drops are indicative of multiple defects other than leaks, the individual layers were intersected to determine the regions with a high risk of a leak based on both parameters. Finally, one chlorine monitoring point showing a low chlorine value of 0.15 mg/L narrowed down the suspected area to a vicinity extending from Butina to Industrial Area 4, including a pipe length of about 320 km.

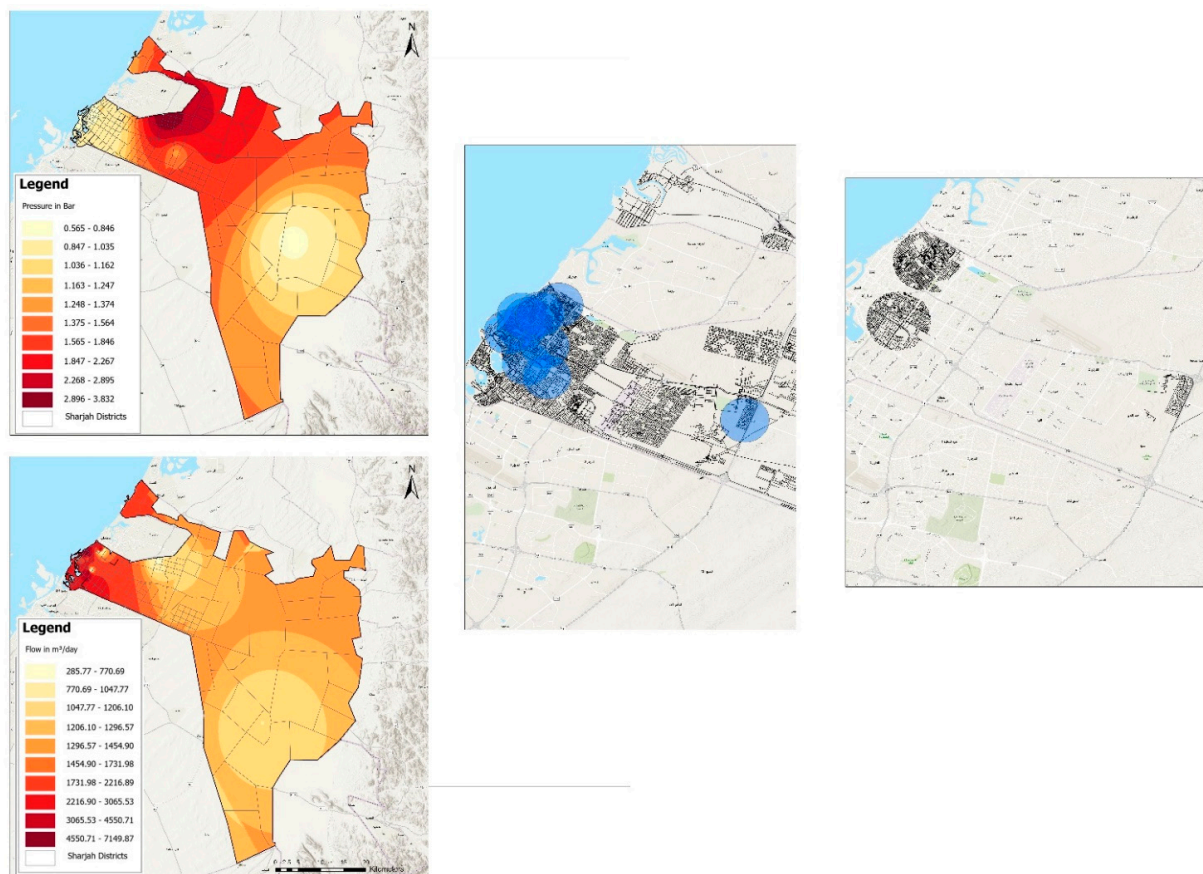


Figure 7. Analysis maps for October.

3.2.2. Winter

Day 1 of January was selected to carry out this scenario. Figure 8 shows the spatial distribution of pressure throughout the study area. The cream colour indicated the lowest observed values and burgundy indicated the highest values, which varied between 0.67 to 4.07 bar, respectively. Areas including Al Rahmánya 1, Al Rahmánya 3, and Al Barashi pressure values were indicative of normal pressure patterns across that specific sector of the network, while Al Sabka, Al Ghuwair, Industrial Area 4, Industrial Area 6, Butina, Al Qadsia, Al Nassreya, Al Gaphia, Al Maysaloon, and Al Fayah were displayed in burgundy, which reflected the highest pressure drops (below 1.3 bar) in the system for the analysed day. Hence, the dark burgundy areas at the western and eastern ends were selected (depicted in brick red) by the automation model to carry out further analysis.

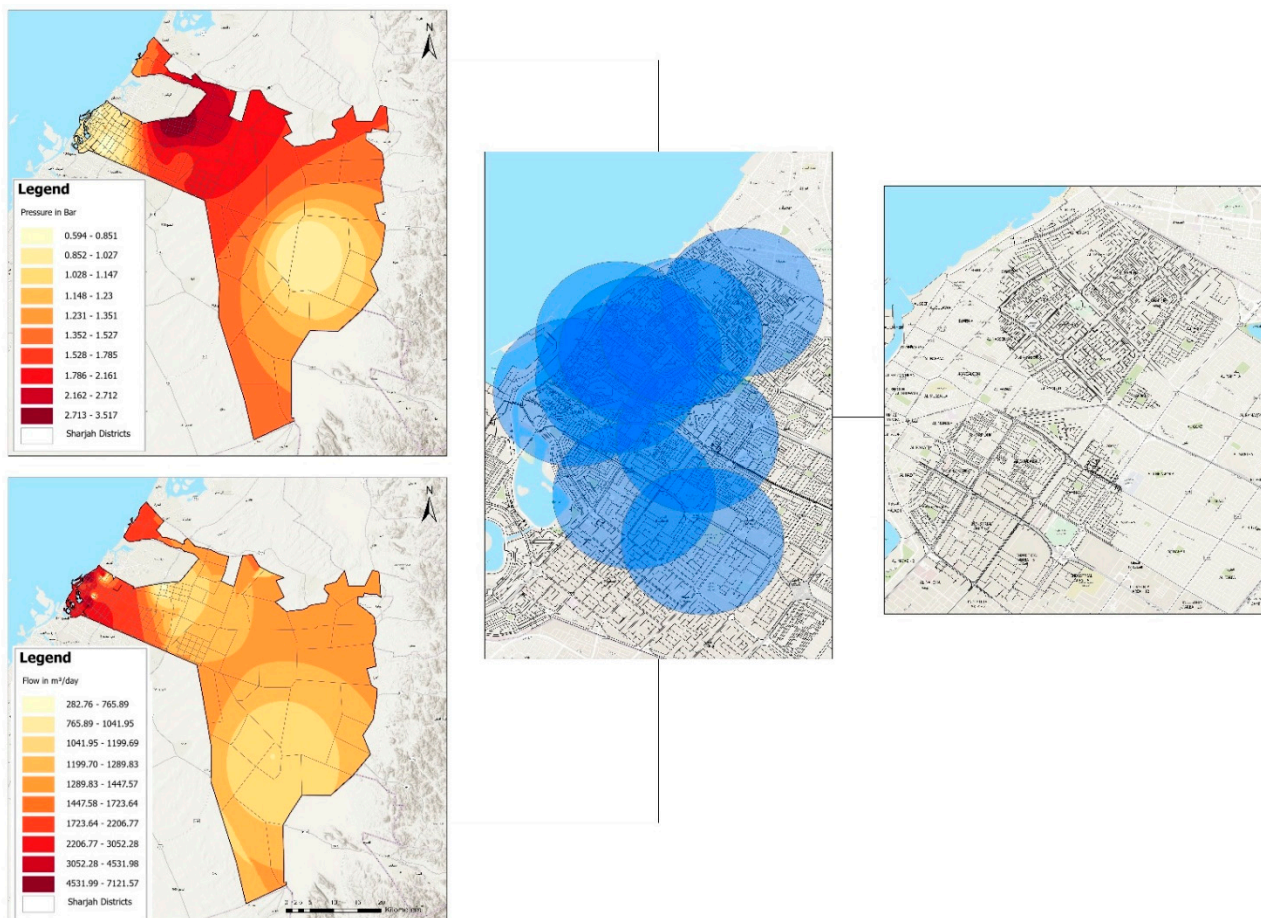


Figure 8. Analysis maps for January.

The IDW created for the flow data utilized the same colour patterns, with burgundy indicating the areas with the highest net flow values (target) and cream specifying areas with the lowest net flow. The net flow values within the study area fluctuated from 30 to 6500 m³/day. Based on the daily averages calculated by the model for each area, the values above the average were selected to be indicative of the leak. The suspected leak area (shown in brick red) was found to extend from Al Rahmánya 1 to Al Ghuwair, including Al Sabka and Al Qadsia. The remaining areas showed a relatively regular distribution of net flows within the network. The chosen areas based on pressure and flow data were overlaid to indicate the most vulnerable areas of the network; however, they were further intersected to determine the common areas where both flow and pressure drops were observed. Finally, two chlorine monitoring points, 0.2 mg/L, and 0.14 mg/L, respectively, were identified to be lower than the minimum acceptable value. As shown in the last step of the flowchart, the automation model selected the part of the network where a leak was

identified based on the three parameters (Figure 8). This network section extended from Al Sabka through Al Ghafia to Al Ghuwair including a pipe length of around 480.0 km.

3.2.3. Summer

The month of July was selected to carry out this run based on the data recorded on the first day of the month. Figure 9 illustrates the subsequent model outputs which return the final leak location specified in the last step (month of June). The spatial variability pressure map showed a varied pattern including pressure drops across most areas excluding Al Rahmaniya 1, Al Rahmaniya 3, and Barashi with pressure values varying from 0.65 to 2.3 bar. The net flow IDW represented relatively mild variations throughout the network based on the calculated averages per area, ranging from 950 to 8950 m³/day. The areas selected by the automation model in step 2 from the individual IDWs of pressure and flow were somewhat similar, indicating a high potential of leak in the observed locations. The layers were overlaid to determine the vulnerable sections of the network and further intersected to represent the vicinities that are at a higher risk. The total suspected leak area appeared to have drastically increased as compared to the previous runs. Finally, five chlorine monitoring points showing low chlorine readings were found to lie within the selected network section (varying from 0.14 to 0.2 mg/L).

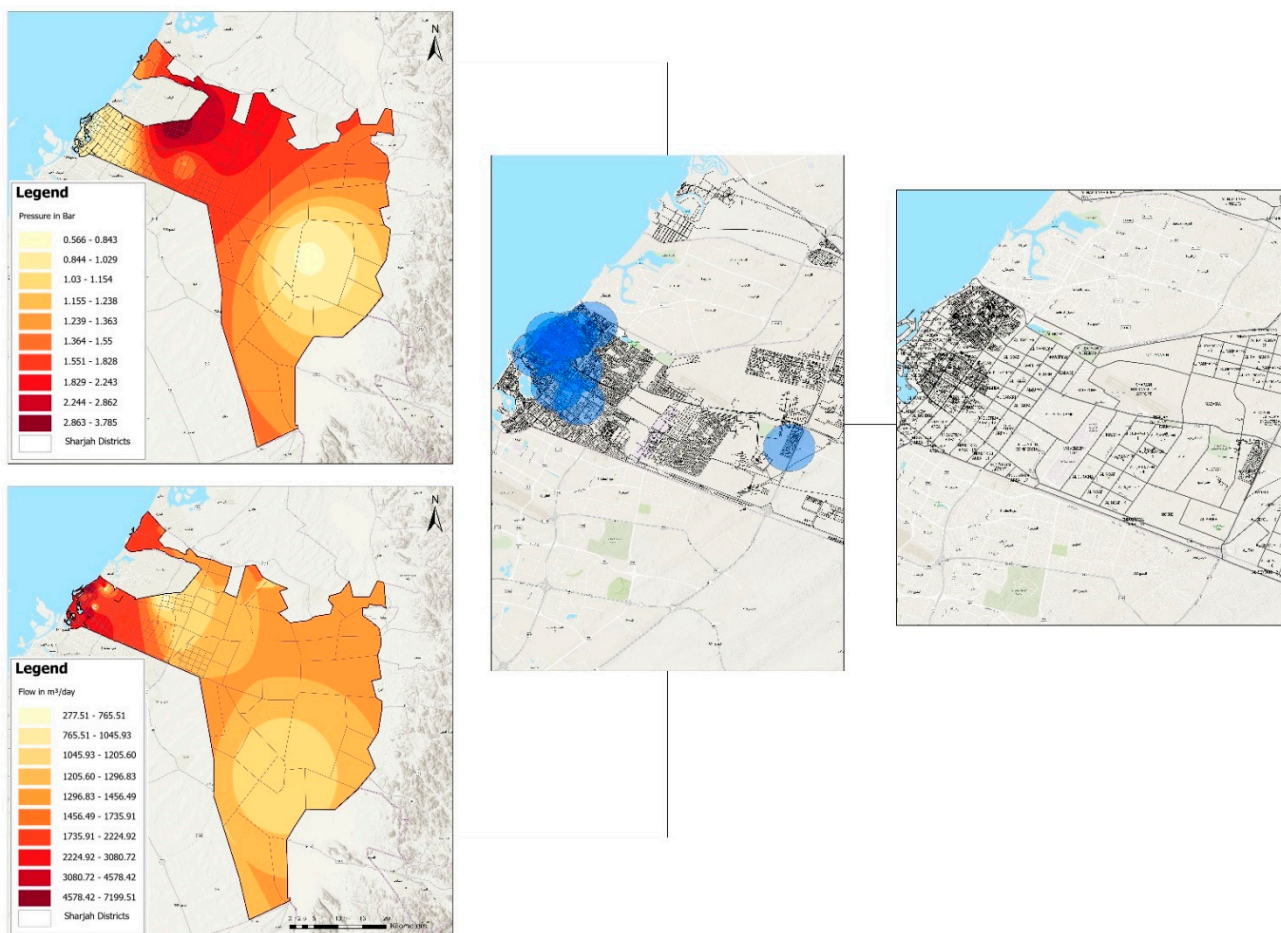


Figure 9. Analysis maps for June.

4. Discussion

The model on ArcGIS 3.0 Pro was successfully tested on a section of the Sharjah WDN. The results obtained from the analysed data showed a gradual increase in the suspected leakage area from October to July. This can be attributed to the extreme temperature fluctuations between the tested seasons as well as the time factor. This indicates that there is

more pipe breakage/leakage in summer and winter compared to the rest of the seasons [45]. SEWA also confirmed the vulnerability of pipe leakage in that area [46].

Typically, there is less rainfall during the summer, causing the ground to be very loose. This allows the pipes to shift and/or expand, which can lead to pipe bursts and leakage as a result. Additionally, hot and dry weather calls for more showers, the frequent recreational use of water in general, the frequent watering of lawns and plants, and more drinking water for our consumption. All of these create a higher-than-normal water demand, which poses a high demand for the water pipes as well. If the pipes are not in good condition, are aging, or are rusting, an increase in water demand can lead to a leak somewhere in the line, or worse, they may burst. Moreover, if a pipe is clogged, the pipeline pressure can increase to an extent that pushes the pipe to crack and leak. Temperature changes in winter can cause the pipes to contract upon cooling, causing cracks and leaks as a result. Furthermore, there is typically a water consumption reduction during this season, causing a rise in the water pipe pressure which might lead to leaks. In transition months like October, the water consumption and weather conditions are typically less extreme and hence do not pose a threat to the distribution pipes. Additionally, any change in soil characteristics due to weather fluctuations would cause changes in pipeline infrastructure, which would cause an imbalance between pipeline sections and fittings, and cause pipe wall rupture, water leakage, and pipe bursts [47].

Another perspective to consider is the time factor between the trials. Pipes that had been damaged during October (due to aging, temperature change, and construction anomalies) could continue to deteriorate over the year if no measures are taken to fix/replace those [48].

Considering that the model results agreed with the typically expected weather, soil, and temporal impacts on the water network and consistency with the literature in potential leakage location, this proves the model's accuracy and efficiency in water leak detection [49]. Additionally, the real-time analysis creates an added benefit of increasing the time to respond or take action, minimizing future risk, understanding the network's behaviour as leaks occur, and identifying the common areas of leaks/bursts, while simultaneously saving unnecessary costs [49]. Finally, the model results can be displayed in the GIS in combination with other layers, including the topographic layer of the city, the road layer, and the water treatment plants in the vicinity, which can provide an understanding of the network behaviour and critical zone identification for optimal management and operation of the pipe system. Furthermore, the method utilizes existing pipe sensors which can eliminate the costs associated with other means of leakage detection. The method also proved its applicability in a large-scale network.

5. Conclusions

Leakage in WDNs is a major concern for water utilities. Hence, a timely inspection of the network is required for the early detection of fluctuations in the system that are indicative of leaks. This raises the necessity of exploring possible approaches to remedy the issue. In this paper, we have described an automatic GIS-based method for leak detection using machine learning, specifically GWR (supervised ML) and LOF (unsupervised ML) models. The method combines sensor readings of flow, pressure, and chlorine using GWR and LOF to detect irregularities that occur in the WDN in real-time. In collaboration with SEWA, the method was tested on a part of Sharjah's WDN successfully. The seasonal variations in water parameters have been predicted to be consistent with the temperature variations. This study can be very useful for water utilities for prompt leak localization. The results indicated the effectiveness of our method in detecting the presence of leaks within the network. For large WDNs, carrying out leakage studies is often time-consuming and resource-intensive. This model can be used with the existing sensor network to identify potential leakage locations.

The proposed approach has limitations even though the results were successful in this particular WDN. The ability of the model is dependent on the availability, size, and

accuracy of the dataset. Even though the use of different hydraulic and water quality parameters can reduce inaccuracies, a large number of monitoring points are needed for a large WDN for the model to be accurate. Future studies should investigate the applicability of this model with more datasets and in other large WDNs.

Author Contributions: Conceptualization, T.A. and M.M.M.; methodology, T.A. and D.E.; software, R.G., D.E., and L.K.; validation, T.A., M.M.M., S.A., and L.K.; formal analysis, D.E. and T.A.; investigation, T.A. and M.M.M.; resources, M.M.M. and T.A.; data curation, T.A. and M.M.M.; writing—original draft preparation, D.E. and T.A.; writing—review and editing, T.A., M.M.M., and S.A.; visualization, D.E. and R.G.; supervision, T.A. and M.M.M.; project administration, M.M.M. and T.A.; funding acquisition, M.M.M. and T.A. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to acknowledge the support of the American University of Sharjah through grants FRG-22-C-15 and OAP24-CEN-088. This paper represents the opinions of the author(s) and does not mean to represent the position or opinions of the American University of Sharjah.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Some of the data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request. Unprocessed raw images originated from experiments using different types of devices are available for this purpose.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

WDN	Water distribution network
GWR	Geographically Weighted Regression
GIS	Geographical information system
IDW	inverse distance weighting
KNN	k-nearest neighbours
SCADA	supervisory control and data acquisition
CIS	customer information system
LDA	linear discriminant analysis
PRV	pressure reduction valves
SEWA	Sharjah Electricity and Water Authority
OLS	ordinary least squares
VOI	Value of Information
ML	Machine learning

References

- Hunaidi, O.; Chu, W.T. Acoustical Characteristics of Leak Signals in Plastic Water Distribution Pipes. *Appl. Acoust.* **1999**, *58*, 235–254. [[CrossRef](#)]
- Awad, A.; Yahyia, M.; Albasha, L.; Mortula, M.; Ali, T. Communication Network for Ultrasonic Acoustic Water Leakage Detectors. *IEEE Access* **2020**, *8*, 29954–29964. [[CrossRef](#)]
- Aslam, H.; Mortula, M.M.; Yehia, S.; Ali, T.; Kaur, M. Evaluation of the factors impacting the water pipe leak detection ability of GPR, infrared cameras, and spectrometers under controlled conditions. *Appl. Sci.* **2022**, *12*, 1683. [[CrossRef](#)]
- Hassani, R.A.; Ali, T.; Mortula, M.; Gawai, R. An Integrated Approach to Leak Detection in Water Distribution Networks (WDNs) Using GIS and Remote Sensing. *Appl. Sci.* **2023**, *13*, 10416. [[CrossRef](#)]
- Xu, W.; Zhou, X.; Xin, K.; Boxall, J.; Yan, H.; Tao, T. Disturbance Extraction for Burst Detection in Water Distribution Networks Using Pressure Measurements. *Water Resour. Res.* **2020**, *56*, e2019WR025526. [[CrossRef](#)]
- Sala, D.A.; Kołakowski, P. Detection of Leaks in a Small-Scale Water Distribution Network Based on Pressure Data—Experimental Verification. *Procedia Eng.* **2014**, *70*, 1460–1469. [[CrossRef](#)]
- Chatzigeorgiou, D.; Youcef-Toumi, K.; Ben-Mansour, R. Design of a Novel In-Pipe Reliable Leak Detector. *IEEE-ASME Trans. Mechatron.* **2015**, *20*, 824–833. [[CrossRef](#)]
- Ishido, Y.; Takahashi, S. A New Indicator for Real-Time Leak Detection in Water Distribution Networks: Design and Simulation Validation. *Procedia Eng.* **2014**, *89*, 411–417. [[CrossRef](#)]

9. Casillas, M.V.; Garza-Castañón, L.E.; Puig, V. Extended-Horizon Analysis of Pressure Sensitivities for Leak Detection in Water Distribution Networks: Application to the Barcelona Network. In Proceedings of the 2013 European Control Conference (ECC), Zurich, Switzerland, 17–19 July 2013; 8, pp. 401–409. [[CrossRef](#)]
10. Ferrandez-Gamot, L.; Busson, P.; Blesa, J.; Tornil-Sin, S.; Puig, V.; Duviella, E.; Soldevila, A. Leak Localization in Water Distribution Networks Using Pressure Residuals and Classifiers. *IFAC-PapersOnLine* **2015**, *48*, 220–225. [[CrossRef](#)]
11. Sadeghioon, A.M.; Metje, N.; Chapman, D.; Anthony, C. SmartPipes: Smart Wireless Sensor Networks for Leak Detection in Water Pipelines. *J. Sens. Actuator Netw.* **2014**, *3*, 64–78. [[CrossRef](#)]
12. Asgari, H.; Maghrebi, M.F. Application of Nodal Pressure Measurements in Leak Detection. *Flow Meas. Instrum.* **2016**, *50*, 128–134. [[CrossRef](#)]
13. Ayadi, A.; Ghorbel, O.; Obeid, A.; BenSaleh, M.S.; Abid, M. Leak Detection in Water Pipeline by Means of Pressure Measurements for WSN. In Proceedings of the 3rd International Conference on Advanced Technologies for Signal and Image Processing, ATsip 2017, Fez, Morocco, 22–24 May 2017; pp. 1–6. [[CrossRef](#)]
14. Wong, L.; Deo, R.N.; Rathnayaka, S.; Shannon, B.; Zhang, C.; Chiu, W.K.; Kodikara, J.; Widyastuti, H. Leak Detection in Water Pipes Using Submersible Optical Optic-Based Pressure Sensor. *Sensors* **2018**, *18*, 4192. [[CrossRef](#)]
15. Sadeghioon, A.M.; Metje, N.; Chapman, D.; Anthony, C. Water Pipeline Failure Detection Using Distributed Relative Pressure and Temperature Measurements and Anomaly Detection Algorithms. *Urban Water J.* **2018**, *15*, 287–295. [[CrossRef](#)]
16. Abdulshaheed, A.; Mustapha, F.; Anuar, M.F.M. Pipe Material Effect on Water Network Leak Detection Using a Pressure Residual Vector Method. *J. Water Resour. Plan. Manag.* **2018**, *144*, 05018006. [[CrossRef](#)]
17. Amoatey, P.K.; Bårdossy, A.; Steinmetz, H. Inverse Optimization Based Detection of Leaks from Simulated Pressure in Water Networks, Part 2: Analysis for Two Leaks. *J. Water Manag. Model.* **2018**, *26*, 1–10. [[CrossRef](#)]
18. Salguero, F.J.; Cobacho, R.; Pardo, M.Á. Unreported Leaks Location Using Pressure and Flow Sensitivity in Water Distribution Networks. *Water Sci. Technol. Water Supply* **2018**, *19*, 11–18. [[CrossRef](#)]
19. Khorshidi, M.A.; Nikoo, M.R.; Taravatroy, N.; Sadegh, M.; Al-Wardy, M.; Al-Rawas, G. Pressure Sensor Placement in Water Distribution Networks for Leak Detection Using a Hybrid Information-Entropy Approach. *Inf. Sci.* **2020**, *516*, 56–71. [[CrossRef](#)]
20. Manzi, D.; Brentan, B.M.; Lima, G.M.; Izquierdo, J.; Luvizotto, E. Pattern Recognition and Clustering of Transient Pressure Signals for Burst Location. *Water* **2019**, *11*, 2279. [[CrossRef](#)]
21. Geelen, C.V.C.; Yntema, D.; Molenaar, J.; Keesman, K.J. Monitoring Support for Water Distribution Systems Based on Pressure Sensor Data. *Water Resour. Manag.* **2019**, *33*, 3339–3353. [[CrossRef](#)]
22. Shao, Y.; Li, X.; Zhang, T.; Chu, S.; Liu, X. Time-Series-Based Leakage Detection Using Multiple Pressure Sensors in Water Distribution Systems. *Sensors* **2019**, *19*, 3070. [[CrossRef](#)]
23. Soldevila, A.; Fernandez-Canti, R.M.; Blesa, J.; Tornil-Sin, S.; Puig, V. Leak Localization in Water Distribution Networks Using Bayesian Classifiers. *J. Process Control* **2017**, *55*, 1–9. [[CrossRef](#)]
24. Aymon, L.; Decaix, J.; Carrino, F.; Mudry, P.-A.; Mugellini, E.; Khaled, O.A.; Baltensperger, R. Leak Detection Using Random Forest and Pressure Simulation. In Proceedings of the 2019 6th Swiss Conference on Data Science (SDS), Bern, Switzerland, 14 June 2019. [[CrossRef](#)]
25. Güngör, M.; Yazar, U.; Cantürk, Ü.; Firat, M. Increasing Performance of Water Distribution Network by Using Pressure Management and Database Integration. *J. Pipeline Syst. Eng. Pract.* **2019**, *10*, 04019003. [[CrossRef](#)]
26. Mulholland, M.; Purdon, A.; Latifi, M.A.; Brouckaert, C.J.; Buckley, C.A. Leak Identification in a Water Distribution Network Using Sparse Flow Measurements. *Comput. Chem. Eng.* **2014**, *66*, 252–258. [[CrossRef](#)]
27. Farah, E.; Shahrour, I. Leakage Detection Using Smart Water System: Combination of Water Balance and Automated Minimum Night Flow. *Water Resour. Manag.* **2017**, *31*, 4821–4833. [[CrossRef](#)]
28. Al-Washali, T.M.; Sharma, S.; Al-Nozaily, F.; Haidera, M.; Kennedy, M.D. Modelling the Leakage Rate and Reduction Using Minimum Night Flow Analysis in an Intermittent Supply System. *Water* **2018**, *11*, 48. [[CrossRef](#)]
29. Jiménez-Cabas, J.; Romero-Fandiño, E.; Torres, L.; Sanjuán, M.; López-Estrada, F. Localization of Leaks in Water Distribution Networks Using Flow Readings. *IFAC-PapersOnLine* **2018**, *51*, 922–928. [[CrossRef](#)]
30. Pal, A.; Kant, K. Water Flow Driven Sensor Networks for Leakage and Contamination Monitoring in Distribution Pipelines. *ACM Trans. Sens. Netw.* **2019**, *15*, 1–43. [[CrossRef](#)]
31. Tornyeviadzi, H.M.; Mohammed, H.; Seidu, R. Semi-supervised anomaly detection methods for leakage identification in water distribution networks: A comparative study. *Mach. Learn. Appl.* **2023**, *14*, 100501. [[CrossRef](#)]
32. Alghushairy, O.; Alsini, R.; Soule, T.; Ma, X. A Review of Local Outlier Factor Algorithms for Outlier Detection in Big Data Streams. *Big Data Cogn. Comput.* **2021**, *5*, 1. [[CrossRef](#)]
33. Desmet, A.; Delore, M. Leak Detection in Compressed Air Systems Using Unsupervised Anomaly Detection Techniques. In Proceedings of the Annual Conference of the PHM Society, St. Petersburg, FL, USA, 2–5 October 2017; Volume 9.
34. Brunson, C.; Fotheringham, A.S.; Charlton, M. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geogr. Anal.* **2010**, *28*, 281–298. [[CrossRef](#)]
35. Sheehan, K.R.; Strager, M.P.; Welsh, S.A. Advantages of Geographically Weighted Regression for Modeling Benthic Substrate in Two Greater Yellowstone Ecosystem Streams. *Environ. Model. Assess.* **2012**, *18*, 209–219. [[CrossRef](#)]

36. Koh, E.; Lee, E.-H.; Lee, K. Application of Geographically Weighted Regression Models to Predict Spatial Characteristics of Nitrate Contamination: Implications for an Effective Groundwater Management Strategy. *J. Environ. Manag.* **2020**, *268*, 110646. [[CrossRef](#)]
37. Zhu, C.; Zhang, X.; Zhou, M.; He, S.; Gan, M.; Yang, L.; Wang, K. Impacts of Urbanization and Landscape Pattern on Habitat Quality Using OLS and GWR Models in Hangzhou, China. *Ecol. Indic.* **2020**, *117*, 106654. [[CrossRef](#)]
38. Nugroho, W.; Iriawan, N. Effect of the Leakage Location Pattern on the Speed of Recovery in Water Supply Networks. *J. Phys.* **2019**, *1402*, 022023. [[CrossRef](#)]
39. Ghorbanian, V.; Karney, B.W.; Guo, Y. Pressure Standards in Water Distribution Systems: Reflection on Current Practice with Consideration of Some Unresolved Issues. *J. Water Resour. Plan. Manag.* **2016**, *142*, 04016023. [[CrossRef](#)]
40. National Research Council, Division on Earth, Life Studies, Water Science, Technology Board, Committee on Public Water Supply Distribution Systems, Assessing, and Reducing Risks. *Drinking Water Distribution Systems: Assessing and Reducing Risks*; National Academies Press: Cambridge, MA, USA, 2007.
41. Langowski, R.; Brdyś, M.A. An Optimised Placement of the Hard Quality Sensors for a Robust Monitoring of the Chlorine Concentration in Drinking Water Distribution Systems. *J. Process Control* **2018**, *68*, 52–63. [[CrossRef](#)]
42. Al Alzarooni, E.; Ali, T.; Atabay, S.; Yilmaz, A.G.; Mortula, M.M.; Fattah, K.P.; Khan, Z. GIS-Based Identification of Locations in Water Distribution Networks Vulnerable to Leakage. *Appl. Sci.* **2023**, *13*, 4692. [[CrossRef](#)]
43. Toubal, A.K.; Achite, M.; Ouillon, S.; Dehni, A. Soil Erodibility Mapping Using the RUSLE Model to Prioritize Erosion Control in the Wadi Sahouat Basin, North-West of Algeria. *Environ. Monit. Assess.* **2018**, *190*, 210. [[CrossRef](#)] [[PubMed](#)]
44. Yahia, M.; Gawai, R.; Ali, T.; Mortula, M.M.; Albasha, L.; Landolsi, T. Non-Destructive Water Leak Detection Using Multitemporal Infrared Thermography. *IEEE Access* **2021**, *9*, 72556–72567. [[CrossRef](#)]
45. Di, W.Y.; Li, S.P.; Liang, X. Analysis of Chinese Media Reports on Water Pipe Burst Events in 2010. *Appl. Mech. Mater.* **2013**, *316–317*, 727–731. [[CrossRef](#)]
46. Mortula, M.; Ali, T.; Sadiq, R.; Idris, A.; Mulla, A.A. Impacts of Water Quality on the Spatiotemporal Susceptibility of Water Distribution Systems. *Clean-Soil Air Water* **2019**, *47*, 1800247. [[CrossRef](#)]
47. Lu, H.; Li, S.P.; He, Y.; Zhou, W.; Zou, J. Statistical Analysis of Domestic Web News Reported Burst Events on Municipal Water Distribution System in 2011. *Appl. Mech. Mater.* **2012**, *212–213*, 619–627. [[CrossRef](#)]
48. Ayad, A.; Khalifa, A.; Fawy, M. A Model—Based Approach for Leak Detection in Water Distribution Networks Based on Optimisation and GIS Applications. *Civ. Environ. Eng.* **2021**, *17*, 277–285. [[CrossRef](#)]
49. Wols, B.A.; Van Thienen, P. Impact of Weather Conditions on Pipe Failure: A Statistical Analysis. *Aqua* **2013**, *63*, 212–223. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.