Research Article

Leak Detection and Localization in Water Distribution Networks Using Conditional Deep Convolutional Generative Adversarial Networks

Mohammad Mahdi Rajabi[1*], Pooya Komeilian[2], Xi Wan[3], Raziyeh Farmani[3]

[1] *Civil and Environmental Engineering Faculty, Tarbiat Modares University, PO Box 14115-397, Tehran, Iran.*

[2] *Department of Civil Engineering, Sharif University of Technology, Tehran, Iran.*

[3] *Centre for Water Systems, Department of Engineering, University of Exeter, Exeter, Devon EX4 4QF, UK.*

* Corresponding author E-mail: mmrajabi@modares.ac.ir

First Submitted: January 2023

Revision Submitted: April 2023

Abstract

This paper explores the use of '*conditional convolutional generative adversarial networks*' (CDCGAN) for image-based leak detection and localization (LD&L) in water distribution networks (WDNs). The method employs pressure measurements and is based on four pillars: (1) hydraulic model-based generation of leak-free training data by taking into account the demand uncertainty, (2) conversion of hydraulic model input demand-output pressure pairs into images using kriging interpolation, (3) training of a CDCGAN model for image-to-image translation, and (4) using the structural similarity (SSIM) index for LD&L. SSIM, computed over the entire pressure distribution image is used for leak detection, and a local estimate of SSIM is employed for leak localization. The CDCGAN model employed in this paper is based on the pix2pix architecture. The effectiveness of the proposed methodology is demonstrated on leakage datasets under various scenarios. Results show that the method has an accuracy of approximately 70% for real-time leak detection. The proposed method is well-suited for real-time applications due to the low computational cost of CDCGAN predictions compared to WDN hydraulic models, is robust in presence of uncertainty due to the nature of generative adversarial networks, and scales well to large and variable-sized monitoring data due to the use of an image-based approach.

***Keywords*:** Leak; anomaly detection; generative adversarial networks; image-to-image translation; structural similarity index.

1. Introduction

With the rapid development of supervisory control and data acquisition (SCADA) technologies, real-time monitoring of hydraulic parameters is becoming increasingly commonplace in many water distribution networks (WDNs) (Zhou et al., 2019). Monitoring is often performed by installing pressure and flow sensors in the transmission mains, allowing for an improved understanding of system behavior, and diagnosis of anomalous events, the most notable of which is pipe leaks. In this context, diagnosis of leakage often includes at least two aspects: (a) 'leakage detection', which is to detect if leakage has occurred, and usually ends with binary results (i.e. leak alarm on or off), or fuzzy values and probabilities between 0 to 1 to represent the likelihood of a leakage event in the system (Mounce et al., 2010). (b) 'Leak localization' (or isolation) is the process of narrowing down the potential location of a leak to a specific area or district. The latter is a precursor to finding the leaky pipe(s) and then pinpointing the exact leakage location (Al Qahtani et al. 2020).

1.1. Literature Review

Several past studies (e.g. Kang et al., 2017; Guo et al., 2021) rely on auxiliary data such as acoustic and vibration signals, for more accurate LD&L in presence of nuisance factors. Despite the effectiveness, obtaining such data is labor intensive, and the quality of these data can be affected by background noise and other events in the system or its environment (Wu and Liu, 2017).

Leakages in pipeline systems can also be detected using transient analysis by relying on the principle of pressure transient wave reflection. A comprehensive overview of these methods was provided by Abdulshaheed et al. (2017). However, the effectiveness of transient analysis-based detection techniques is heavily influenced by pipe characteristics and external factors (Kammoun et al., 2022). Furthermore, this approach is unsuitable for detecting leaks over long pipe distances since pressure waves only propagate short distances. Additionally, implementing this technique often requires intricate mathematical algorithms and substantial computational resources (Wan et al., 2022).

Pressure and flow sensors have the advantage of being easy to install, enabling real-time monitoring, and being sensitive to leakage events. Therefore, leakage detection and localization (LD&L) using pressure/flow data has received increasing attention in the past decades. LD&L using pressure/flow data can be done through one of the following two groups of methods. The distinguishing factor between the two is whether a hydraulic model is used or not (Wan et al., 2022).

A. ‘Data-driven’ methods rely on historical monitoring data and involve spatial analysis of changes across a WDN, or temporal pattern analysis of time series data. Methods in this group may involve: (1) defining control limits by calculating some statistical characteristic (e.g. exponentially weighted moving average) of historical measurements and defining data that is outside these limits as potential leaks (e.g. Jung et al., 2015; Loureiro et al., 2016). These methods often rely on assumptions such as the Gaussianity of data distribution and the uncorrelation of errors. (2) Using historical data to train a predictor model that learns the expected behavior of the system in leak-free circumstances, and comparing the residuals between the observed and predicted value with a threshold to identify leaks. A variety of machine learning tools, such as support vector machines and classic neural networks have been employed in this context (e.g. Mounce et al., 2011; Romano et al., 2012; Tijani et al., 2022; Tariq et al., 2022). (3) Reframing the spatial/temporal pattern recognition as a classification problem (using e.g. Bayesian classifiers) (e.g. Wu et al., 2016). This method requires labeled data, i.e. an indication of whether a chosen data instance pertains to a leak or not. (4) Formulating LD&L as a clustering problem by grouping similar data into different clusters, and identifying leaks as those that are dissimilar to normal cluster(s) (e.g. Wu and Liu, 2020).

B. 'Model-based' methods, also known as physically-based or process-based methods, rely on the comparison of measured data with those of a calibrated WDN hydraulic model. Model-based methods include: (1) formulating LD&L as an inverse modeling problem using an optimization algorithm, and one or multiple objective function(s) (e.g. Sanz et al., 2016). (2) Relying on sensitivity-to-leak analysis (e.g. Perez et al., 2014), where model-generated pressure disturbances caused by all possible leak locations and magnitudes are stored in a leak-sensitivity matrix and matched against the difference between measured and simulated data (Soldevila et al., 2017). (3) Generating pressure or pressure residual maps of each leakage scenario using the hydraulic model, employing the resulting data to train a classifier, and finally utilizing the trained classifier for LD&L. The latter is sometimes referred to as a mixed model-based/data-driven approach (Soldevila et al., 2016).

Both data-driven and model-based methods have their advantages and limitations, and it is unfeasible to draw a conclusion about which method is universally superior (Wan et al., 2022). Data-driven approaches directly incorporate experimental data, and hence don't require a deep knowledge of the system operation and physical equations. Data-driven methods are best suited for WDNs with a long-term monitoring dataset and were found to be effective in leak detection. However, their performance in leak localization is debated (Wu and Liu, 2017). It is generally difficult to obtain enough correct labeled data to train a high-accuracy model in a supervised manner (Zhang et al., 2020); therefore data-driven LD&L methods often rely on unsupervised or semi-supervised training. Many commonly used tools within the context of the data-driven approach have limitations in learning complex features and using them requires the manual design of suitable feature extractors (Zhou et al., 2019). Data-driven methods are particularly vulnerable to missing or faulty data (Hu et al., 2021).

Model-based methods are preferred when limited historical data is available, but come at the cost of building and calibrating a hydraulic model. Moreover, the high computation demand of hydraulic models often hinders achieving real-time LD&L. A common challenge for both model-based and data-driven methods is that pressure and flow rate can be affected by other factors apart from leakages, such as variability and stochasticity in demand and random errors in sensor measurements. These are often difficult to distinguish from leaks. For model-based methods, the effect of input parameter uncertainty on the model output flow and pressure, adds to this complexity (Menapace et al., 2018; Sun et al., 2019). Hence model-based methods do not perform well in larger WDNs where the models' output uncertainty is often higher compared to smaller WDNs (Zhou et al., 2019). The performance of the model-based sensitivity-to-leak analysis method is particularly well-known to decrease due to the nodal demand uncertainty and noise in the measurements (Cugueró-Escofet et al., 2015).

The above-described challenges have forced previous LD&L studies to focus mainly on hypothetical burst events in simple WDNs (Wan et al., 2022). As highlighted by recent review papers (e.g. Gupta and Kulat, 2018; Chan et al., 2018; Wan et al., 2022), more effort is needed to solve these challenges by developing novel methods that are: (a) robust in presence of various forms of uncertainty, (b) can scale to very large monitoring datasets, (c) can learn complex features from raw data, (d) rely less on manual design, and (e) are well-adapted to a limitation in labeled training data.

A promising approach to achieve this objective is to exploit deep neural networks (DNNs). DNNs rely on hierarchical projections of the input space into increasingly low-dimensional latent representations (Goodfellow et al., 2016), enabling them to learn complex features from large amounts of data without the need for developing manual features by domain experts (Chalapathy and Chawla, 2019). In recent years, several DNN-based anomaly detection methods have been introduced for a variety of applications, demonstrating significantly better performance than conventional anomaly detection with an increase in the scale of data and complexity of the problem (Pang et al., 2021). Example applications of DNNs in LD&L include the use of convolutional neural networks (CNNs) for burst localization based on data from short-duration pressure observations (Zhou et al., 2019); supervised training of a CNN using hydraulic model-generated samples of all possible leaks for leak localization (Javadiha et al., 2019). These approaches, like many older methods, require a well-calibrated hydraulic model to generate synthetic data that accurately represent real-life situations for training a deep learning model. However, acquiring and maintaining a well-calibrated hydraulic model can be challenging for water companies.

Regression analysis has been used as an alternative method to learn the data patterns representing the healthy state of a WDN from historical monitoring data instead of a hydraulic model. As an example, Ye and Fenner (2011) utilized the Kalman filter to fit historical flow measurements and identify burst events by comparing predicted and observed values, while also employing the weighted least squares regression to model the data. However, these approaches treat each data point of a day independently, and thus, they do not fully consider the autocorrelation of the time series data.

Mounce et al. (2010) employed a neural network to learn the normal flow behavior from monitoring data and detect leaks based on prediction errors. A key limitation of their approach is the inability to share features across different steps of a time series, whereas temporal pattern recognition requires the ability to process evolving patterns. To address this problem, recurrent neural networks such as long short-term memory (LSTM) networks have been explored for leakage detection (Wang et al., 2020; Xu et al., 2020). While the results of these LSTM-based methods have shown high detection accuracy, there are still some limitations that need to be addressed. Specifically, these methods are often restricted to univariate time series analysis, which only allows for the analysis of data from a single sensor at a time. However, it's known that data from sensors in the same network are spatially correlated, and considering topology information and the spatial relation of the sensors is essential for accurate leak detection.

1.2. Study Objectives

Recent research progress has demonstrated the effectiveness of DNNs in solving LD&L problems, and it is becoming increasingly clear that DNNs present an opportunity for improving current practice. In this work, we will be leveraging some of the most recent advances in DNNs, to develop a robust methodology for achieving accurate real-time LD&L in complex settings. Our methodology is based on the use of '*conditional convolutional generative adversarial networks*' (CDCGAN) for image-based anomaly detection. Partially similar methods have been used in other applications for anomaly detection (e.g. Ezeme et al., 2020; Luo et al., 2021; Qiu et al., 2022), but to the best of our knowledge, this method has not been applied to leak detection in WDNs.

Relying on flow measurements for LD&L is often easier than pressure measurements, because flow data allows for the use of simple mass balance relations, and pressure data is less sensitive to leak events, especially when the pressure sensor is located far from the leak location (Ye and Fenner, 2011). However, pressure meters are easier to install and less costly than flow meters (Zhou et al., 2019; Sun et al., 2019), provide instantaneous data, and are better for LD&L in WDNs where there is a dense mesh of pipes with only flow measurements at the entrance of each DMA (Soldevila et al., 2017). Hence, this work focuses on the use of pressure data for LD&L.

The rest of the paper is organized as follows. Section 2 presents the background of semi-supervised anomaly detection (SSAD) and CDCGANs. Section 3 presents the proposed methodology for LD&L. Section 4 details the application of the method to a WDN, followed by a discussion of the main findings. Finally, Section 5 draws the main conclusions of the work and introduces some potential extensions.

2. Theoretical Background

2.1. Semi-Supervised Deep Anomaly Detection

Leaks result in anomalous pressure observations, and most notably, a contextual anomaly. Therefore, LD&L in WDNs can be described as a particular case of the general problem of anomaly detection and isolation in dynamic systems (Soldevila et al., 2016). Most problems in this context suffer from the limited availability of labeled anomalous data (Schlegl et al., 2019). For LD&L, this is mostly because (a) historical monitoring data on leak events is often scarce and may be unreliable, (b) full-scale physical experiments (using e.g. fire hydrants) are very costly, and (c) model-generated leak data are limited to the specifications of the model, and comprehensive evaluation of all possible leak locations and magnitudes is often computationally unfeasible. Supervised anomaly detection is limited to already known anomalies, hence lack of labeled anomalous data severely limits its value (Wan et al., 2022). SSAD is a way to employ a large set of unlabeled data alongside limited labeled data to construct a classifier with good generalization ability (Tu et al., 2018). In SSAD, easier-to-obtain samples of normal (i.e. leak-free) data are given, and the model learns a discriminative boundary around the normal instances. New data instances that don't belong to the normal class are identified as anomalous. SSAD methods rely on the assumption that points which are close to each other in the learned feature space are more likely to share the same label (Chalapathy and Chawla, 2019).

2.2. Generative Adversarial Networks

GANs trained in a semi-supervised manner have shown great promise, even with few labeled data (Mu et al., 2021). The GAN architecture, in its basic form, is composed of a generator (G) and a discriminator (D) neural network (Goodfellow et al., 2014). A random input latent space $z \sim p_z$ (often sampled from a normal or uniform probability distribution, e.g. $z \sim \mathcal{N}(0,1)$) is mapped to the data space ϕ by the generator, which tries to generate new examples that are ideally identical to the training dataset p_{data}. The discriminator is responsible for classifying a given generator output ϕ as either real (i.e. indistinguishable from the training dataset p_{data}) or fake (non-identical to the training dataset) (Wang et al., 2017). These models are trained together in a zero-sum manner, also called min-max and adversarial, such that improvements in the discriminator come at the cost of reduced capability of the generator, and vice versa (Arjovsky and Bottou, 2017; Gui et al., 2021). This can be represented by the following objective function (Zheng et al., 2020):

$$\min_G \max_D \mathbb{E}_{\phi \sim p_{data}}[logD(\phi)] + \mathbb{E}_{z \sim p_z}[\log(1 - D(G(z)))] \tag{1}$$

GANs have several key advantages: (a) GANs are well-adapted to SSAD, eliminating the need for the often time-consuming, cumbersome, and sometimes unfeasible task of providing labeling for anomalous data (Singh and Raza, 2021), (b) Markov chains and inference are not needed during learning, and only backpropagation is used to obtain gradients (Goodfellow et al., 2014; Mirza and Osindero, 2014), and (c) GANs are capable of learning many noise types and mimicking complex (including very sharp) data distributions (Wunderlich and Sklar, 2022).

GANs allow for the synthesis of novel images, videos, numeric and audio data, or text from a random input (Al Qahtani et al., 2021). Hence, GANs are commonly used for data synthesis to facilitate the training of other ML models when data is insufficient or to correct the overtraining of a DNN (Zamora et al., 2021). For image data, the generator and discriminator frequently take the form of deep CNNs (see Radford et al., 2015 as a pioneering work). The resulting architecture is often referred to as deep convolutional GAN (DCGAN).

2.3. Conditional Training

GAN models generate new plausible examples of a given dataset, but their outputs are practically random and uncontrollable. There is no way to control the outputs other than discovering the complex relationship between the input latent space and the GAN-generated outputs, which is generally difficult and often unfeasible (Wang et al., 2018). However, GANs can be conditioned on auxiliary inputs, allowing for the targeted generation of outputs. This can be done by hotencoding the conditioning data and concatenating it to the input of the generator (noise input) as well as the discriminator (generated data) (Denton et al., 2015; Qasim et al., 2020). Conditional GANs are considered semi-supervised models (Zhang et al., 2019). In a conditional GAN, both the generator and the discriminator models are conditioned, and hence the trained generator model can be used as a standalone model to generate data in the domain of interest. The most common strategy is to condition GANs on class labels (i.e. class-conditional GANs), but GANs can also be conditioned on the auxiliary image(s) in the context of *image-to-image translation* tasks (Brownlee, 2019). The conditional training of a DCGAN-based model is referred to as CDCGAN.

3. Methodology

The proposed methodology employs a CDCGAN in the context of SSAD to identify and locate leaks in a WDN. The methodology is based on four pillars: (1) hydraulic model-based generation of training data, (2) conversion of hydraulic model input-output pairs into images, (3) semi-supervised training of a CDCGAN model for image-to-image translation, and (4) using the structural similarity (SSIM) index for LD&L. These pillars are described in the following subsections. A flowchart of the proposed methodology is provided in **Fig. 1**.

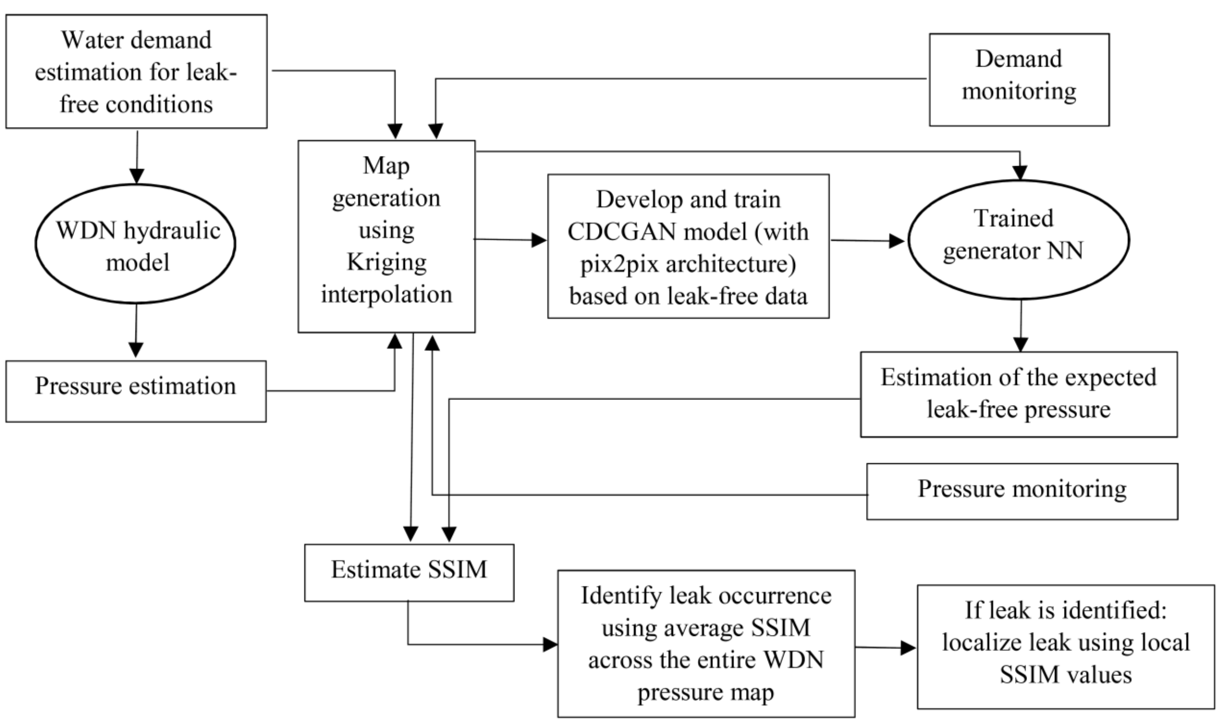


Fig. 1. Flowchart of the proposed methodology.

3.1. Hydraulic Model-Based Generation of Training Data

In a hypothetical leak-free condition, it is subjectively assumed that the uncertain demand in node i and time step t (denoted by $d_{i,t}$) is represented by a normal distribution with a mean ($\mu(d_{i,t})$) equal to the field data (or estimates) of demand, and a standard deviation of $\sigma(d_{i,t}) = 0.1\mu(d_{i,t})$ (Soldevila et al., 2017). We employ Latin hypercube sampling (LHS) to produce a variety of random, plausible values for pipe parameters and demand time series and feed them to the EPANET model to obtain the resulting pressure time series in the observation nodes. We refer to data obtained from hydraulic simulations as 'model-generated data'.

3.2. Transforming Demand and Pressure Data to Images

Demand and pressure values at various observation points (OPs) in a single time step are employed to create images of demand and pressure. This results in the creation of two '*image time series'* for the entire duration of interest. These images are created by interpolating pointwise values using the kriging method with a Gaussian variogram model. Kriging is a popular choice for interpolating data points to a continuous spatial field (Kleijnen, 2017), and has been previously used in several studies including Javadiha et al. (2019) for pressure and/or demand interpolation in WDNs. The kriging method is implemented here using the PyKrige python library (Murphy, 2014). The resulting interpolated values are then scaled into the interval 0 and 255 to generate rectangular greyscale images with 256 × 256 pixels.

3.3. CDCGAN for Image-to-image Translation

The demand-pressure image pairs are subsequently employed for training a CDCGAN, to learn how certain demand distribution maps to the associated leak-free pressure distribution. The CDCGAN model employed in this work is based on the pix2pix architecture (Isola et al., 2017). Pix2pix is composed of a generator and a discriminator network, as portrayed in **Fig. 2**. The generator is an encoder-decoder CNN (ED-CNN) based on a modified U-Net architecture (Ronneberger et al. 2015). The ED-CNN consists of (a) an encoder subnetwork that receives the input image, passes it through a contracting process in which features of increasing semantic depth and decreasing spatial resolution are learned from the input image, and outputs feature maps, and (b) a decoder which receives the feature maps and employs deconvolution and up-sampling to constructed an output image with the same spatial resolution as the original input (Rajabi et al., 2022). The encoder is made of several blocks, where each block consists of a convolutional layer proceeded by batch normalization and leaky Rectified linear unit (ReLU) activation function. Blocks of the decoder consist of a transposed convolutional layer, followed by batch normalization, dropout (applied only to the first 3 blocks), and ReLU activation function. Skip connections are employed between the encoder and decoder. The discriminator in pix2pix is a convolutional PatchGAN classifier that maps each generator output to a 70 × 70 square patch of the input image. The patches overlap to cover the 256 × 256-pixel image. The discriminator is made of several blocks with each block consisting of a convolution layer, batch normalization, and Leaky ReLU activation function. The objective function in pix2pix training is the sum of the GAN loss, a binary cross-entropy, and an L_1 norm between the generated image and the ground truth (Isola et al., 2017). In this work, the pix2pix-based CDCGAN is developed using the open-source python library TensorFlow and is trained on Google Colab.

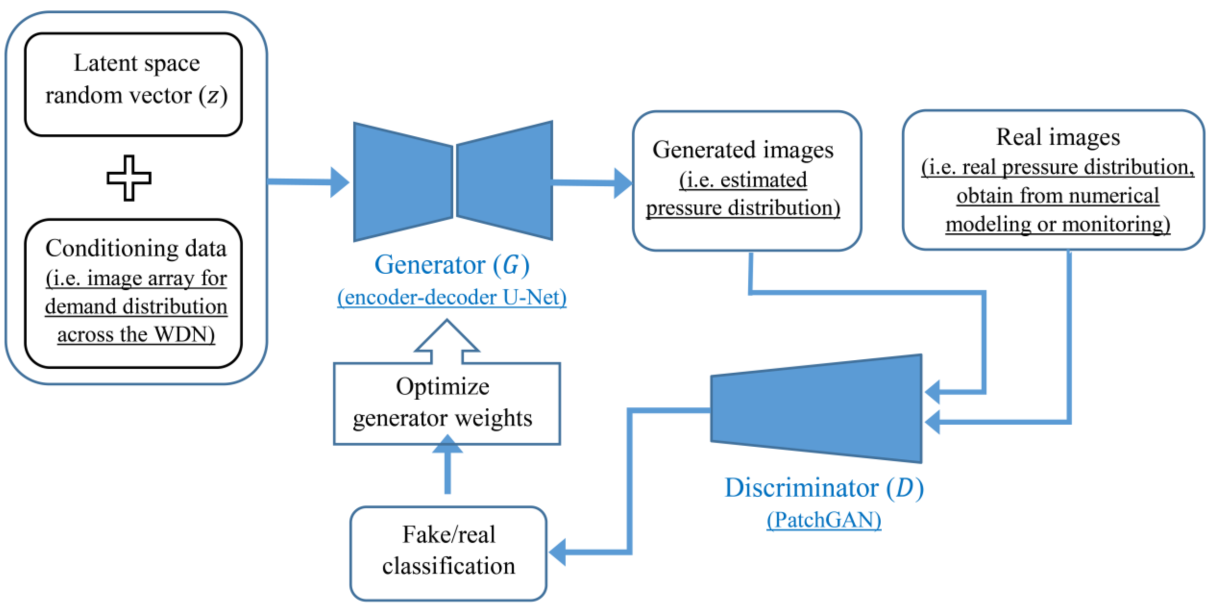


Fig. 2. The architecture of CDCGAN and how it is trained with demand-pressure image pairs

3.4. Leak Detection and Localization

After the model is trained, the CDCGAN generator is fed with new demand images (for which some are associated with leaks in the WDN) and outputs the pressure distribution image which represents what is expected in leak-free conditions. The SSIM index is then used for measuring the similarity between the ground truth (either field or model-generated data), and the CDCGAN-predicted pressure images. SSIM is a perceptual metric that quantifies the difference between two images from the same image capture and has been used in many image quality assessment applications (Chen and Bovik, 2011). SSIM is calculated on various windows of an image. The measure between two windows x and y of common size $N \times N$ is a weighted combination of three comparative measures, namely luminance ($l(x,y)$), contrast ($c(x,y)$), and structure ($s(x,y)$) which are defined as follows (Wang et al., 2004):

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \tag{2}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \tag{3}$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \tag{4}$$

Where μ_x and μ_y are the average of x and y respectively, σ_x and σ_y are their standard deviations, and σ_{xy} is their covariance. c_1, c_2 and c_3 are variables included to avoid instability when the denominator is close to zero. SSIM ranges between 0 to 1, where 1 denotes a perfect match between the reconstructed and original images. Here, SSIM is multiplied by 100 to give a percentage score. SSIM can be estimated locally (which we denote by $SSIM_l$) or computed over the entire image (represented by $SSIM_o$).

Due to factors such as demand uncertainty, the target, and CDCGAN predicted pressure images are not expected to perfectly match even in leak-free conditions. Therefore the mean and standard deviation of $SSIM_o$ in absence of a leak, is first estimated by employing an independent set of leak-free demand-pressure image pairs. The resulting values are then employed to derive a threshold that differentiates between leak-free and leak conditions. We choose this threshold (T_{SSIM_o}) to be mean minus three times the standard deviation of $SSIM_o$. This choice is based on the three-sigma rule, which is widely used in statistics and quality control to identify outliers or abnormal values in a dataset. This rule assumes that the data follows a normal distribution and that approximately 99.7% of the data will fall within three standard deviations of the mean (Panda and Khilar, 2015). However, the appropriateness of this rule depends on the characteristics of the data and the specific needs of the analysis. A more conservative threshold value, such as mean minus four or five times the standard deviation, may reduce the risk of false alarms but increase the risk of missed detections.

Furthermore, it is well-known that one abnormal data point cannot signify a leakage event, while continuous disruptive data is more suitable to indicate the occurrence of a leak (Wan et al., 2022). In this study, if the $SSIM_o$ obtained from comparing the ground truth and the CDCGAN predicted pressure is less than T_{SSIM_o} for 5 consecutive time steps (i.e. snapshots of pressure observation), we conclude that there is a leak somewhere in the WDN. The selection of the number of consecutive time steps is dependent on the sensitivity of the detection method and the noise level in the system. A smaller number of time steps may produce false alarms due to measurement noise, while a larger number of time steps may delay the detection of leaks. A balance between the risk of false alarms and missed detections is often sought in practice, and five consecutive time steps may be a suitable choice. If a leak is identified, leak localization is then carried out by using the local SSIM ($SSIM_l$). For leak localization, the area with the lowest value of $SSIM_l$ across the pressure map at each time step is identified as the most probable leak location.

4. Application

4.1. Description of the Case Study

We analyze the effectiveness and accuracy of the proposed methodology using the L-Town benchmark problem (Vrachimis et al., 2020). The L-Town problem for LD&L is founded on a transient hydraulic model that mimics the characteristics of a real-world water distribution system. This model is designed to simulate changes in flow and pressure in the pipes over time, which are influenced by a variety of factors, including variations in demand and pump operations. Additionally, the L-Town problem accounts for different types of leaks, ranging from small to large leaks, and introduces them at different times during the simulation to create transient conditions.

This case study encompasses a surface area of 3×2.6 km^2 and a total pipe length of 42 km. There are 782 junctions, 2 reservoirs, 1 tank, 905 pipes, 1 pump, and 3 pressure-reducing valves in the WDN (see **Fig. 3**). The network is monitored using 33 pressure sensors. Sensor measurements are assumed to be accurate, with no time delay, and are reduced to two decimal points. The measurement dataset is one year in length and has 5-minute time steps.

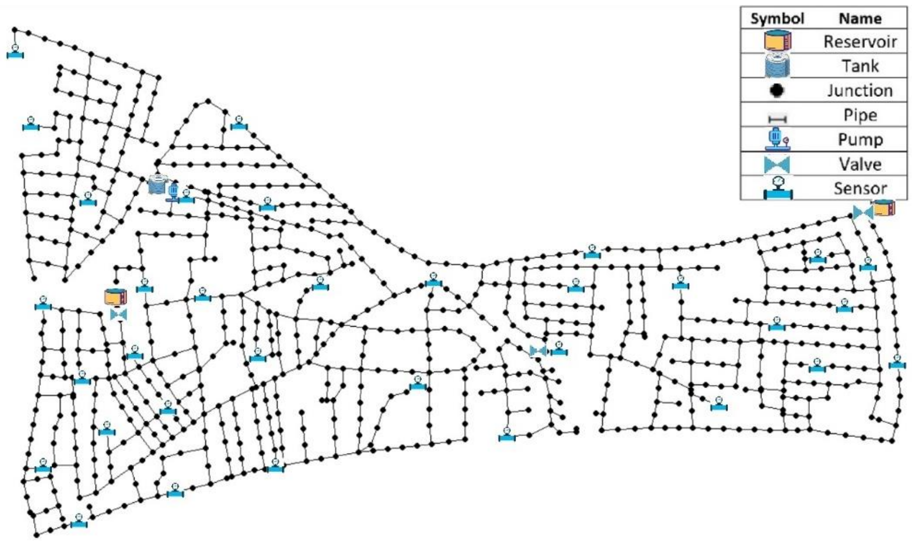


Fig. 3. Map of the WDN used as our case study.

An EPANET-based model of the L-Town benchmark problem (Vrachimis et al., 2022) is employed for data generation. This hydraulic model utilizes a pressure-driven analysis method to simulate the network. The model was calibrated based on field measurements obtained from the original water distribution system in L-Town (Steffelbauer et al., 2020). To facilitate iterative model simulations, we couple this model with an EPANET-compatible Python library, named Water Network Tool for Resilience (WNTR) (Klise et al., 2018) (available on GitHub: https://github.com/USEPA/WNTR). WNTR has an application programming interface (API) that allows for changes to the network operations and simulation of disruptive incidents such as leaks and bursts (Klise et al., 2020).

4.2. Generating the Training Data

Data for CDCGAN training is generated using the leak-free EPANET model (Vrachimis et al., 2020), following the scheme described in sub-sections 3.1 and 3.2. In this context, 365 daily demand patterns with a time step of 5 minutes have been generated, taking into account demand uncertainty. This results in 105,120 pairs of demand-pressure images, which are divided into training and testing data by an 80%-20% split. This procedure is first done using pressure observations in 780 nodes in the WDN (two nodes close to the reservoirs, which have exceptionally high pressures, have been omitted from the original 782 nodes of the case study), and then repeated by only incorporating the 33 locations for which pressure sensors are available (see **Fig. 3**).

In the process of generating pressure distribution images, since the temporal variations are small compared to the spatial variations, we divided the WDN into four zones based on the range of observed pressures (i.e. $zone_1$: $P < 33$ m, $zone_2$: $33 < P < 40$ m, $zone_3$: $40 < P < 50$ m, and $zone_4$: $P > 50$ m), and then separately scaled pressure values into the interval 0 and 255 for each zone. As an example of the leak-free images obtained from Kriging interpolation, **Fig. 4** depicts demand and pressure distributions in randomly chosen time steps, obtained from 780 (**Figs. 4a, 4b**) and 33 OPs (**Figs. 4d, 4e**). As the demand in various nodes may serve different purposes, such as residential, commercial, or industrial, reducing the number of data nodes (from 780 to 33) has a greater impact on the interpolated demand maps than on the pressure maps.

4.3. CDCGAN Training and Validation

Two separate CDCGANs are trained using images obtained from 780 and 33 OPs. We refer to these two as CDCGAN_{780} and CDCGAN_{33} respectively. For both models, 30,000 epochs are sufficient to reach a stable solution. Details of the hyper-parameter settings are presented in **Table 1**. The mean absolute percentage error (MAPE) is used to measure and quantify the prediction errors. **Fig. 5** illustrates the MAPE histogram of the trained CDCGANs. The average MAPEs are 1.42%, and 0.67% for CDCGAN_{780} and CDCGAN_{33} respectively. **Fig. 5** and the average MAPEs show that the prediction error of the CDCGAN_{33} is lower than those of CDCGAN_{780}. This can be attributed to the fact that more OPs result in more complex pressure maps that should be learned by the model. For both CDCGANs, the error distribution is not uniform or normal and is skewed to the right.

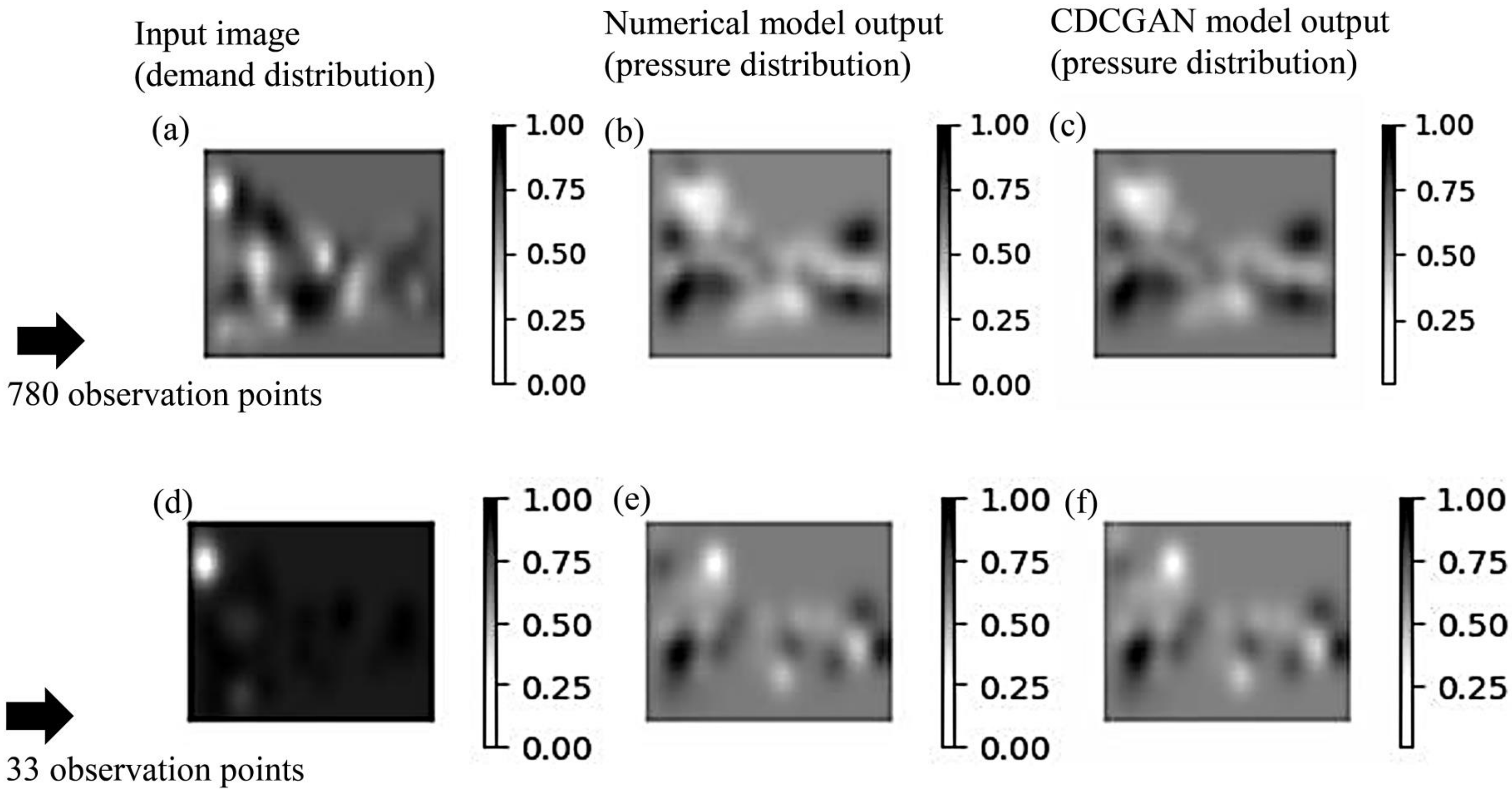


Fig. 4. Comparison of the numerical model and CDCGAN metamodel outputs (i.e. pressure distribution) for the same input image (i.e. demand distribution), randomly chosen from the test dataset, for (a, b, c) 780, and (d, e, f) 33 observations points. In these images, the demand and pressure are unitless and normalized through division by the respective maximum values.

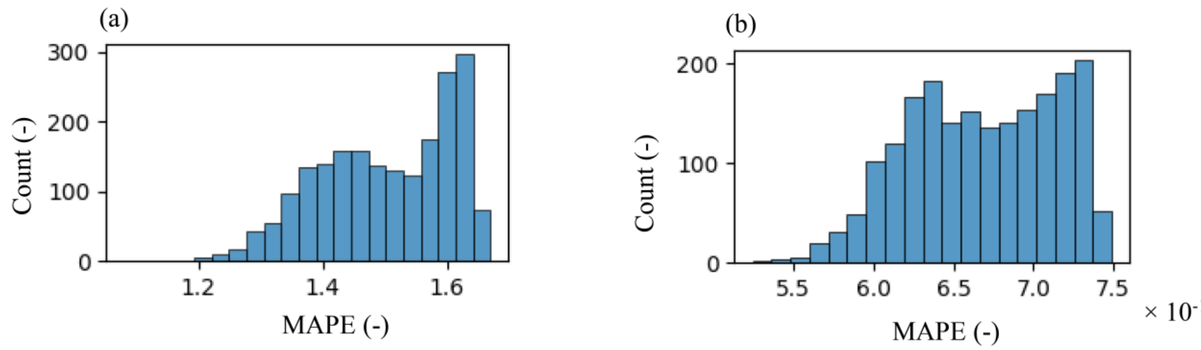


Fig. 5. Histogram of MAPE for the CDCGANs trained based on pressure images obtained from (a) 780, and (b) 33 observation points.

Table 1. Hyper-parameters values for the CDCGAN.

Name	**Optimal choice**
Optimizer	Adam
Learning rate	2×10^{-4}
The exponential decay rate for the 1st-moment estimates	0.5
Epochs	30,000
Normalization type	Batch
Weight of L_1 loss in the generator objective	100
Weight initialization method	Normal
Number of generator filters in the last conv layer	256
Number of discriminator filters in the first conv layer	256

Based on the trained CDCGANs, the SSIM is estimated for the test dataset using the method described in sub-section 3.4. To account for diurnal variations in demand and pressure, the leak detection threshold (T_{SSIM_o}) is then calculated on an hourly basis for 24 hours. The resulting thresholds are presented in **Fig. 6**. As demonstrated, hourly T_{SSIM_o} values are consistently larger for $CDCGAN_{33}$ compared to $CDCGAN_{780}$, because the prediction errors are also lower and the model-generated pressure maps more closely resemble the data.

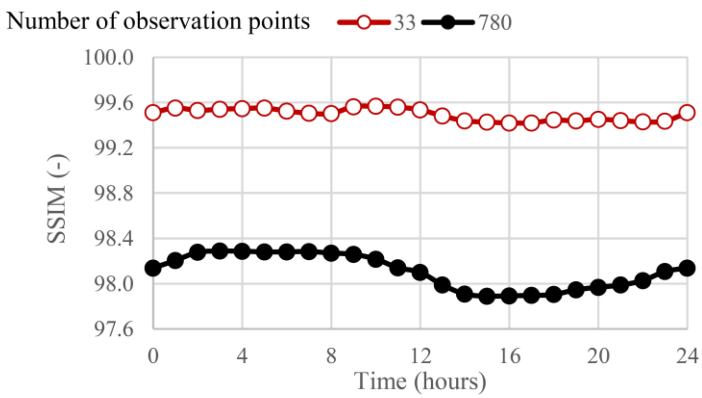


Fig. 6. Hourly thresholds of SSIM for leak detection.

4.4. Leak Detection and Isolation

Two sets of leakage scenarios, described in the following sub-sections, are employed to analyze the performance of the trained CDCGAN model in LD&L. In all scenarios, if leakage occurs and data points during the leak period are identified as anomalies by the CDCGAN model, the outcome is identified as a true positive (TP). On the other hand, if there is no leak in the WDN, and the model doesn't classify data points during the leak as anomalous, the outcome is a true negative (TN). The detection method can also fail. In this case, if the model fails to identify leaks, it results in a false negative (FN), and if it identifies leak-free conditions as anomalous, the outcome is a false positive (FP). Based on these concepts, several key metrics are calculated to evaluate the effectiveness of leak detection (Wan et al., 2022):

1. The true positive rate (TPR), is defined as:

$$TPR = \frac{TP}{TP + FN} \tag{5}$$

2. The true negative rate (TNR), is estimated as:

$$TNR = \frac{TN}{TN + FP} \tag{6}$$

3. Accuracy (ACC) is defined as:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

For a TP, the detection time (DT) is employed to express the time interval between the actual start time of leakage and the time when a method successfully identifies an anomaly (i.e. CDCGAN predicted pressure is less than T_{SSIM_o} for 5 consecutive time steps). In practice, it is highly favorable to have the least DT to minimize water loss and its collateral damages. To quantify the accuracy of leak localization, the graphical distance between the estimated leak location (i.e. location with the minimum $SSIM_l$), and the actual leak location, is employed. We denote this by GDRL.

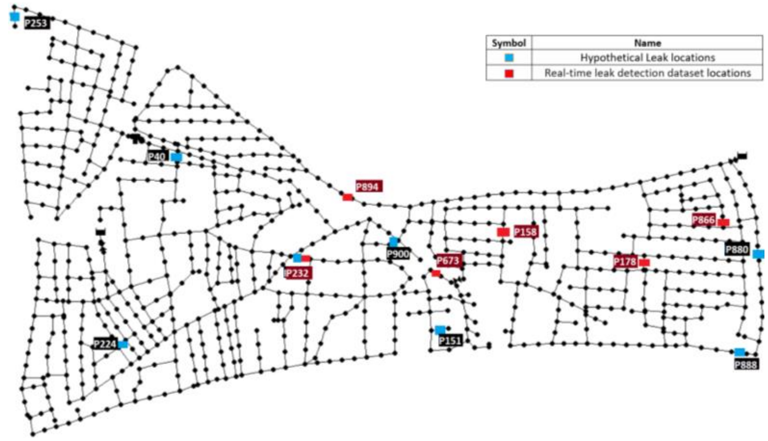
4.4.1. Hypothetical Single Leak Scenarios

For the hypothetical single leaks, EPANET simulations are performed assuming that a single leak has occurred in the WDN. Three different leak rates (with leak areas (LAs) equal to 0.0005, 0.005, and 0.05 m^2), and eight alternative leak locations are simulated, resulting in a total of 24 hypothetical single leak scenarios. Two leaks are selected to represent each of the four leak-free pressure intervals defined in **Fig. 7**: one at the junction of 2 and the other at the junction of 3 pipes. The leak area (A_L) is related to the leak flow rate (q_L) through the following equation (Crowl and Louvar, 2001):

$$q_L = C_L A_L \sqrt{2gh_L} \tag{8}$$

Where C_L is the discharge coefficient with a default value of 0.75, h_L is the head, and g is the acceleration of gravity. We assume that the leaks occur in a stepwise manner (with 10 equally spaced steps) as demonstrated in **Fig. 8**, and are hence expected to cause a similarly stepwise pressure drop in parts of the network. The simulation period pertains to the WDN conditions in the first week of 2018, and the time steps are 5 minutes. For each scenario, this results in the generation of 2,016 pairs of demand-pressure images. The generation of pressure images is once done using 780 observations, and then again for 33 OPs, to assess how the performance of the proposed approach is affected by the number of observation locations. For all hypothetical scenarios, the leak starts at 0:00 of day five and ends at 24:00 of the same day.

Leak location*	Pressure Range	Junction of	
		2 pipes	3 pipes
P-40	<30 m	■	
P-151			■
P-151	[30-40) m	■	
P-253			■
P-880	[40-50) m	■	
P 900			■
P-224	≥50 m	■	
P-888			■



* The numbers represent the node numbers in the L-Town WDN (Vrachimis et al., 2020)

Fig. 7. Location of hypothetical and real-time leaks in the WDN.

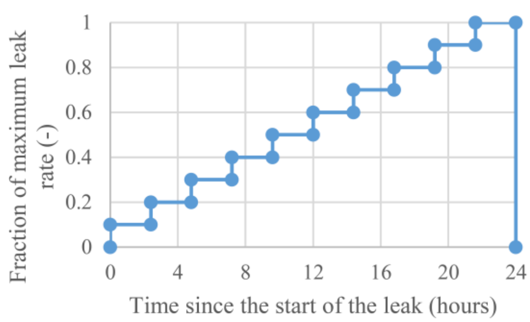


Fig. 8. Stepwise increase in leak rate for the hypothetical single leak scenarios.

As an example of the images acquired through Kriging interpolation, **Fig. 9** compares pressure maps of a no-leak condition with one of the hypothetical single leakage scenarios ($LA = 0.05\ \mathrm{m}^2$) in the same time step. Both images are obtained using the EPANET numerical model using 780 Ops. This exemplary figure demonstrates that the pressure maps significantly differ between the two cases, especially in the vicinity of the leak location. In our proposed leak identification methodology, this difference is quantified using the SSIM index.

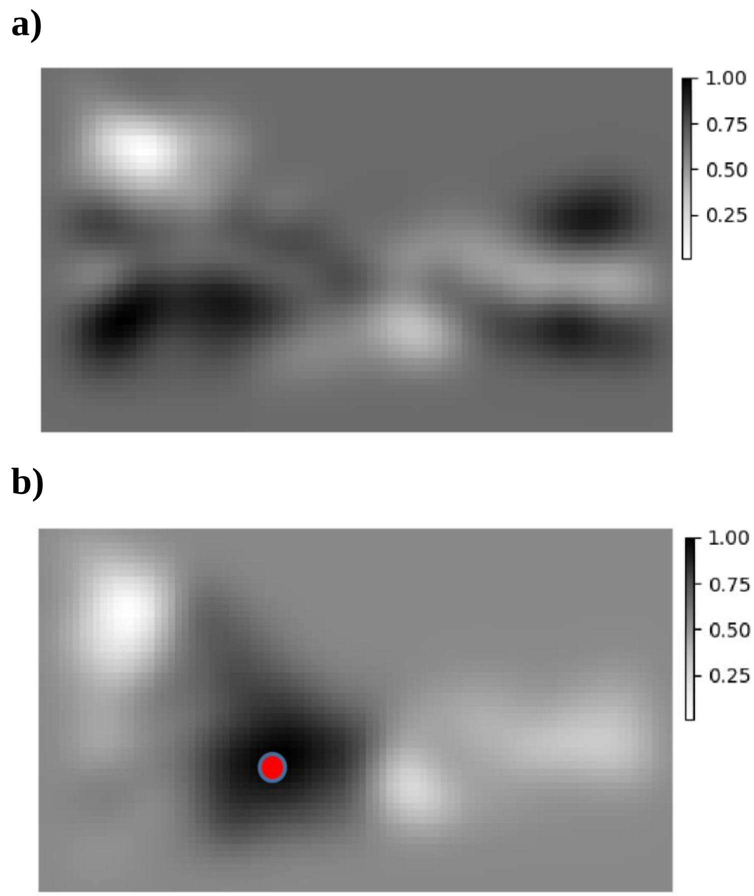


Fig. 9. Examples of how leaks appear in pressure image: (a) no-leak, and (b) leaky conditions. The red circle has been overlaid on the right image to show the location of the leak.

Fig. 10 illustrates the $SSIM_o$ times series obtained from CDCGAN$_{780}$ for the hypothetical leak at P-40, based on three different leak rates. For the highest leak rate ($LA = 0.05$ m^2), $SSIM_o$ drops below the hourly leak detection threshold (T_{SSIM_o}) as soon as the leak starts, and increases back to its normal range as the leak ends (**Fig. 10a**). Hence, the model is perfectly capable of identifying this burst with minimal DT.

To demonstrate how the proposed method identifies the leak location, **Fig. 11** shows the $SSIM_l$ map for three different time steps during the leak. The location of the minimum $SSIM_l$ and the actual leak location is also shown in the figure. During the leak, $SSIM_l$ significantly drops in the vicinity of the leak location, forming a zone of low $SSIM_l$ values around the leak location. The extent of this zone increases with increasing leakage rates. The estimated location of the leak (pertaining to the minimum value of $SSIM_l$) gradually converges to the actual leak location with an increased leakage rate. **Fig. 12a** shows how GDRL varies after the identification of the leak. GDRL starts from about 12m and decreases with increasing leakage rate until it converges to 4 m. This is a relatively highly accurate estimation compared to previous studies (see the review paper, by Wan et al., 2022).

For the middle value of leak rate ($LA = 0.005$ m^2), $SSIM_o$ becomes smaller than T_{SSIM_o} with delay (DT ≅ 16 hours), and after the leak rate has increased beyond 0.002 m^2 (**Fig. 10b**). **Fig. 12b** demonstrates how GDRL varies after the leak is identified. GDRL starts from around 140 m, and then sharply drops to about 12 m after a few hours. For the smallest value of leak rate ($LA = 0.0005$ m^2), DT equals about 21 hours, and the leak is identified only when the leak rate nears its maximum value (**Fig. 10c**). GDRL is approximately 96 m in this scenario.

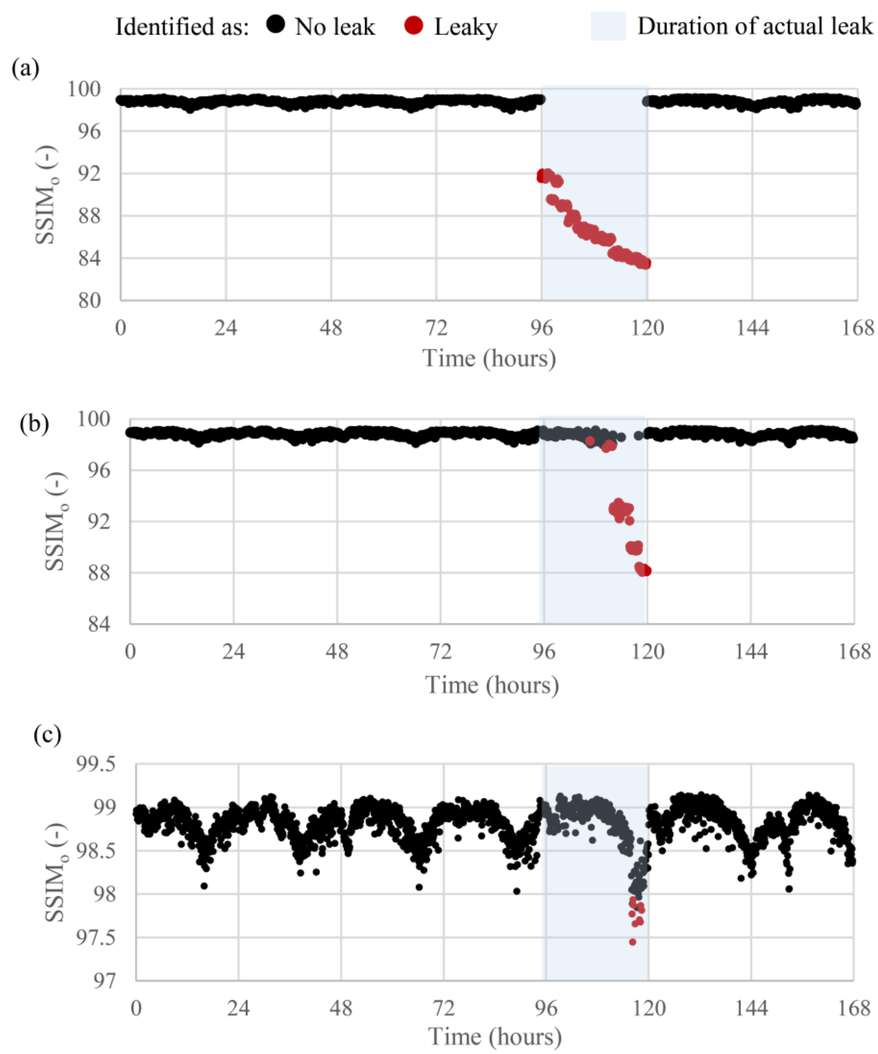


Fig. 10. SSIM$_o$ time series the for the hypothetical leak at P-40, based on: (a) $LA = 0.05$ m^2, (b) $LA = 0.005$ m^{2,} and (c) $LA = 0.0005$ m^2.

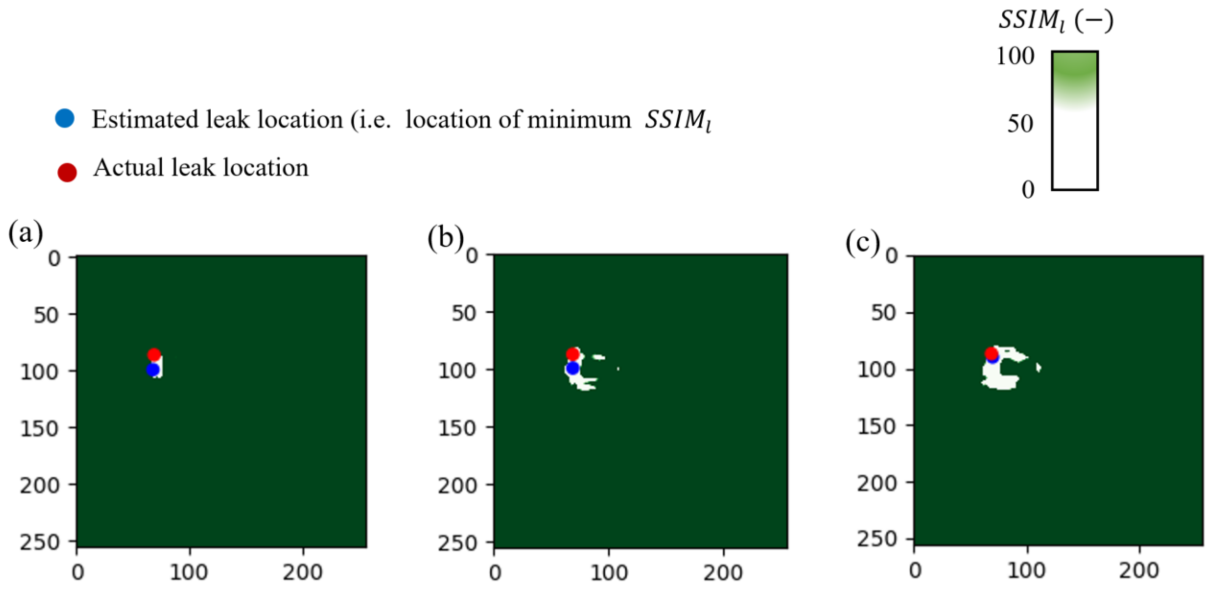


Fig. 11. The $SSIM_l$ map for the hypothetical single leak at times: (a) $t = 100$, (b) $t = 110$, and (c) $t = 120$ hours since the start of the synthetic leak simulations.

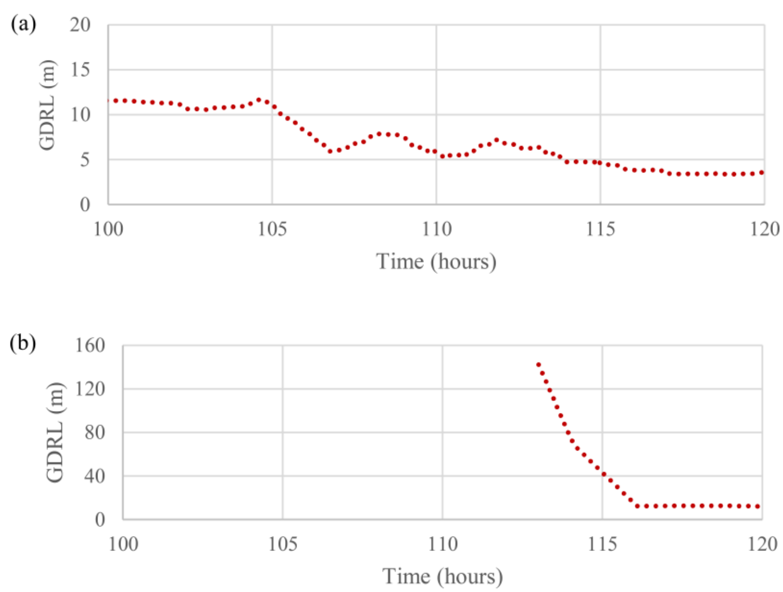


Fig. 12. GDRL time series for the hypothetical leak at P-40 based on: (a) $LA = 0.05$ m^2, and (b) $LA = 0.005$ m^2

Fig. 13 provides the TPRs and TNRs obtained from applying the proposed methodology to all hypothetical single leak scenarios. As demonstrated in **Fig. 13a, b** the TPR is close to the ideal value of one for $LA = 0.05$ m^2, highly decreases with decreasing leak rates, and approaches zero for $LA = 0.0005$ m^2. Detail analysis of the $SSIM_o$ time series for the hypothetical single leak scenarios (an example of which was presented in **Fig.10**) shows that the detectable threshold representing the minimum detectable leak rate is between 0.0001-0.0005 m^2 for the various scenarios. The maximum standard deviation of TPR is observed for the middle value of leak rate ($LA = 0.005$ m^2). The TNR is generally high for all scenarios (**Fig. 13c, d**). Comparison of TPR and TNR values obtained from incorporating 780 (**Fig. 13a, c**) and 33 (**Fig. 13b, d**) OPs, shows a minor difference between the two, with the average TPR for 780 OPs 0.02 higher than 33 OPs. The average ACC across all scenarios is 0.91 and 0.90 for 780 and 33 OPS, respectively.

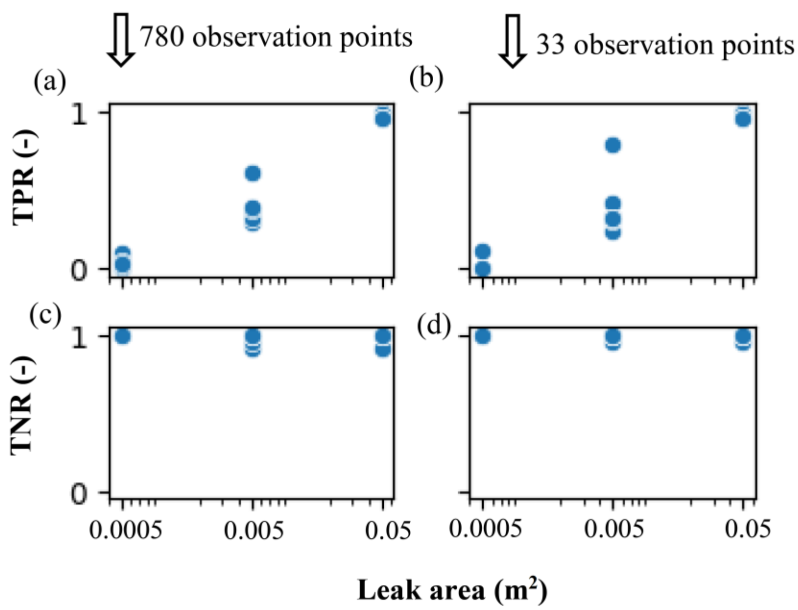


Fig. 13. TPR and TNR for various hypothetic leak scenarios.

Fig. 14 compares TPR and TNR values obtained for leaks in various locations. Based on this figure, there seems to be no correlation between the leak location or pressure range and the associated TPR and TNR. This is observed for both 780 and 33 OPs.

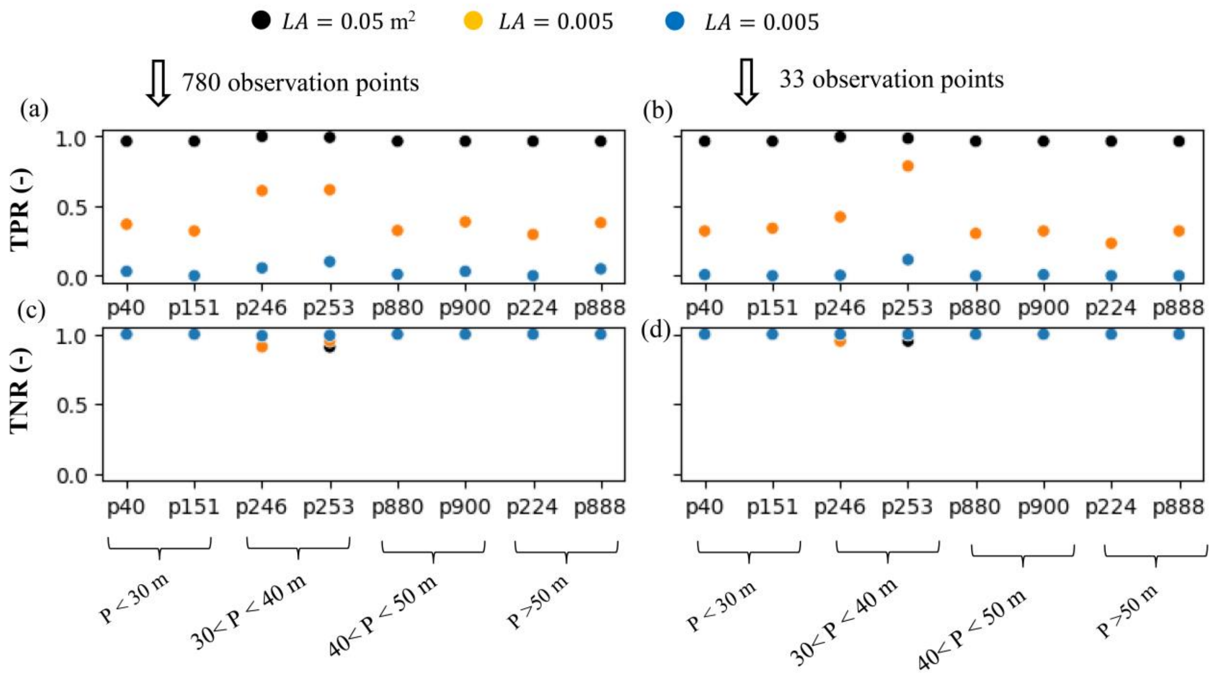


Fig. 14. TPR and TNR for leaks in various locations. The associated pressure ranges are shown for each leak location.

Fig. 15 provides the DT and minimum GDRL (GDRL$_{min}$) values obtained for various leak locations and two different leak rates. It can be seen that DT is significantly lower for the larger leak rate. For $LA = 0.005$ m^2, DT is constantly below 17 hours, and for $LA = 0.05$ m^2 DT is generally less than an hour. GDRL$_{min}$ for the larger leak rate is equal to or less than the associated values of the smaller leak, indicating that, as expected, larger leaks can more accurately be localized. For LA equal to 0.05 and 0.005 m^2, GDRL$_{min}$ is lower than 11 and 35 m, respectively. There is no correlation between leak location and either DT or GDRL$_{min}$.

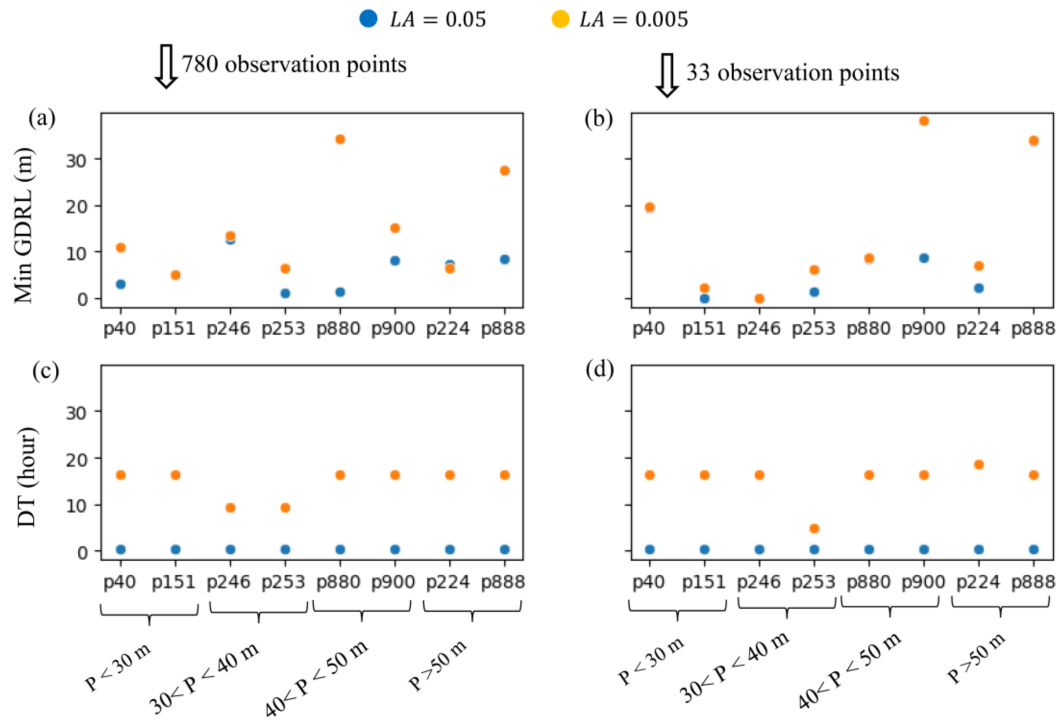


Fig. 15. GDRL and DT for leaks in various locations. The associated pressure ranges are shown for each leak location.

4.4.2. Real-time Leak Detection for a One-Year Dataset

Six real-time leaks in the L-Town WDN provided by the *Battle of Leakage Detection and Isolation Methods* (BattLeDIM) (Vrachimis et al., 2022), are also assessed. These leaks occur at different times during the year 2018 (with no concurrency) and involve a gradual increase in the outflow rate. The leak locations and magnitudes are presented in **Table 2**. All six leaks have smaller peak values of leak rate compared to the hypothetical scenarios described in the previous sub-section. A single one-year EPANET simulation is performed by incorporating these leaks, and the results are used to create pairs of demand-pressure images for each 5-minute time step. Only 33 OPs (of the sensor locations) are employed to create the pressure images. These images are incorporated in the LD&L algorithm, and the resulting $SSIM_o$ time series for the entire year is presented in **Fig. 16**. Except for the leak RL_6 (which has the shortest duration among the six leaks), the other five leaks (RL_1 to RL_5) are detected by the proposed methodology. Generally, leak detection only becomes possible near the peak leak rate. For the entire year, the accuracy of leak detection is 70%. $GDRL_{min}$ ranges between 160 to 185 m for the five identified leaks.

Table 2. Summary of the real-time leak detection scenarios.

Scenario	RL_1	RL_2	RL_3	RL_4	RL_5	RL_6
Leak position	p232	p673	p866	p183	p158	p427
Leak area at the peak (m^2)	0.00032	0.00024	0.00031	0.00041	0.00028	0.00031
Leak type	incipient	incipient	abrupt	incipient	incipient	abrupt
Start (date, time, time step)	2018-05-01, 02:35, 35,167	2018-06-20, 15:45, 49,725	2018-08-03, 07:00, 62,292	2018-08-28, 10:35, 69,535	2018-10-06, 02:35, 80,671	2018-12-15, 13:00, 100,956
End (date, time, time step)	2018-05-17, 09:20, 39,856	2018-07-10, 10:25, 55,421	2018-08-03, 11:00, 62,340	2018-09-15, 17:30, 74,802	2018-11-15, 13:35, 92,323	2018-12-15, 17:00, 101,004
Peak (date, time, time step)	2018-05-12, 16:05, 38,497	2018-07-06, 15:45, 54,333	2018-08-03, 07:00, 62,292	2018-09-10, 02:45, 73,185	2018-11-10, 02:35, 88,867	2018-12-15, 13:00, 100,956
Mean daily discharge (m^3/hour)	4.11	3.01	3.98	5.27	3.6	3.99

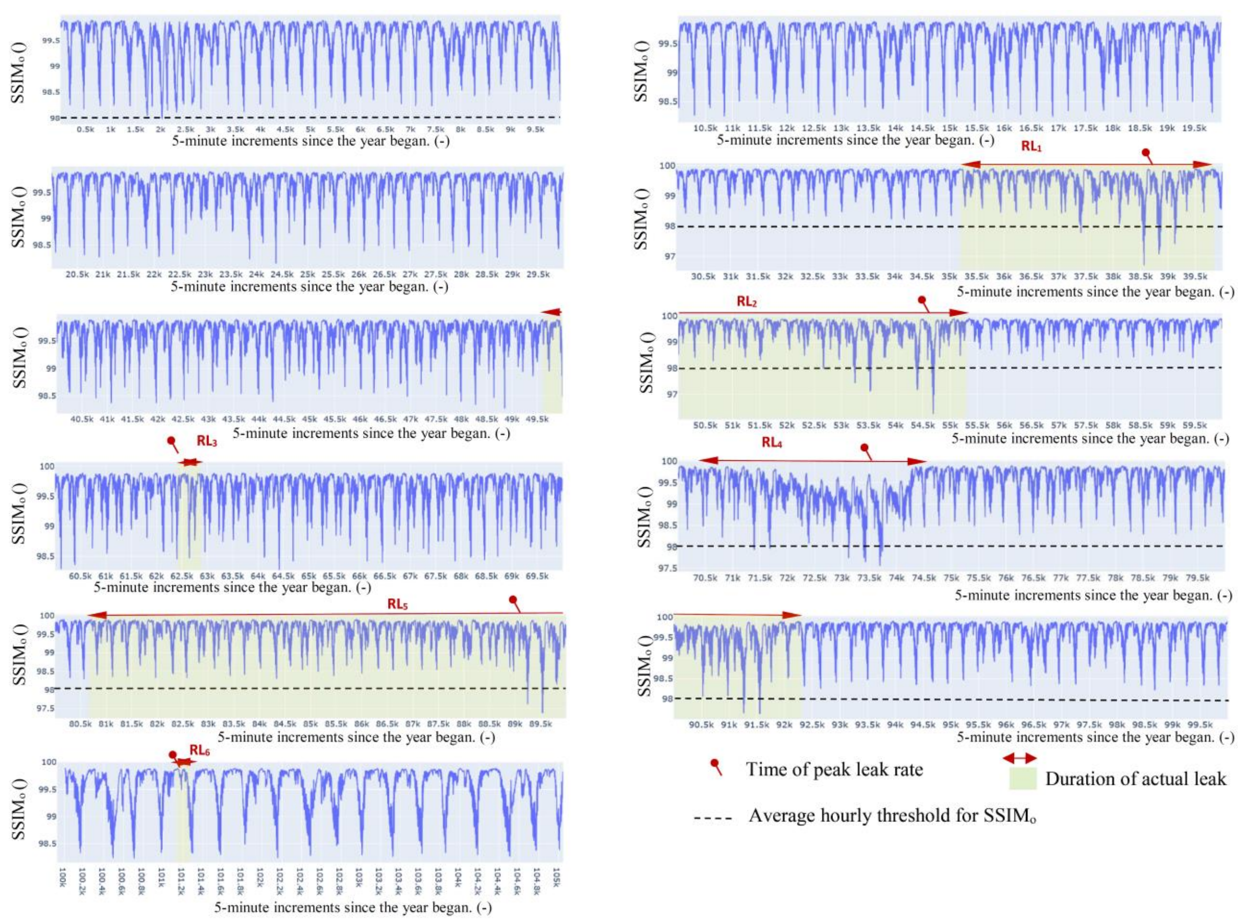


Fig. 16. $SSIM_o$ time series a one-year simulation based on the real-time leaks of L-Town

5. Conclusion

Previous studies in different settings have shown that GANs are a promising tool for anomaly detection and localization (Xia et al., 2022). Due to their ability in mapping complex non-linear relationships, robustness in presence of uncertainty, and adaptability to limited data, GANs can theoretically overcome many key challenges of LD&L. In this paper, we exploit the use of a particular GAN architecture and develop an LD&L methodology around it. Our key contributions can be summarized as follows:

- We develop a conditional GAN architecture based on pix2pix to predict the pressure distribution resulting from a known demand distribution in the context of image-to-image translation. The GAN model is trained using hydraulic model-generated data of leak-free conditions and learns a discriminative boundary around the normal, leak-free instances. New data instances that don't belong to this normal class are identified as anomalous.
- We propose the use of the SSIM index for LD&L. In this framework, the SSIM index is computed over the entire pressure distribution image for leak detection, and a local estimate of SSIM is employed for leak localization.
- We analyze the effectiveness and accuracy of the proposed methodology using the L-Town WDN case study. Besides the six real-time leaks provided by BattLeDIM, several hypothetic leak scenarios are also defined and analyzed to assess the minimum detectable leak, and the correlation between leak rate and location with key metrics such as TPR, TNR, ACC, DT, and GDRL.

The L-Town benchmark problem has several key underlying assumptions, allowing for simplified modeling and LD&L. However, these assumptions may not reflect the actual circumstances in a WDN. Specifically, the assumptions include (1) the network's homogeneity, where pipes, nodes, and other components are assumed to have similar properties and behave uniformly; (2) the absence of leaks or defects at the start of the simulation; and (3) a fully-piped network, which does not include open channels, and (4) no changes in pipe characteristics and system configuration over time. Additionally, in developing our LD&L methodology, we assume that both demand uncertainty and measurement error are normally distributed, which may or may not reflect the actual circumstances in a WDN. Although the normal distribution is commonly used in this context, there are other distribution models, such as the lognormal, gamma, and Weibull distributions, which are also widely used to model demand uncertainty or measurement error in

WDNs. The choice of distribution should be based on the data and the specific needs of the analysis.

Various architectures of GANs, such as f-AnoGAN (Schlegl et al., 2019), BiGAN (Kaplan and Alptekin, 2020), etc., have been proposed for anomaly detection and localization in other settings, and we suggest future studies consider these various architectures to improve LD&L in WDNs.

Future investigations could focus on exploring the robustness of GAN-based methods across different operational scenarios and evaluating the impact of the WDN's size and complexity on the accuracy of the method. Additionally, conducting a comprehensive analysis of the effect of simultaneous leaks and observation location on the model's accuracy and reliability may also be a valuable avenue for further research.

References

Abdulshaheed, A., Mustapha, F., & Ghavamian, A. (2017). A pressure-based method for monitoring leaks in a pipe distribution system: A Review. *Renewable and Sustainable Energy Reviews*, 69, 902-911.

Al Qahtani, H., Kavakli-Thorne, M., & Kumar, G. (2021). Applications of generative adversarial networks (gans): An updated review. *Archives of Computational Methods in Engineering*, *28*(2), 525-552.

Al Qahtani, T., Yaakob, M. S., Yidris, N., Sulaiman, S., & Ahmad, K. A. (2020). A review on water leakage detection method in the water distribution network. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, *68*(2), 152-163.

Arjovsky, M., & Bottou, L. (2017). Towards principled methods for training generative adversarial networks. *arXiv preprint arXiv:1701.04862*.

Brownlee, J. (2019). *Generative adversarial networks with python: deep learning generative models for image synthesis and image translation*. Machine Learning Mastery.

Chan, T. K., Chin, C. S., & Zhong, X. (2018). Review of current technologies and proposed intelligent methodologies for water distributed network leakage detection. *IEEE Access*, *6*, 78846-78867.

Chalapathy, R., & Chawla, S. (2019). Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*.

Chen, M. J., & Bovik, A. C. (2011). Fast structural similarity index algorithm. *Journal of Real-Time Image Processing*, *6*(4), 281-287.

Crowl, D. A., & Louvar, J. F. (2001). *Chemical process safety: fundamentals with applications*. Pearson Education.

Cugueró-Escofet, P., Blesa, J., Pérez, R., Cuguero-Escofet, M. A., & Sanz, G. (2015). Assessment of a leak localization algorithm in water networks under demand uncertainty. *IFAC-PapersOnLine*, *48*(21), 226-231.

Denton, E. L., Chintala, S., & Fergus, R. (2015). Deep generative image models using a laplacian pyramid of adversarial networks. *Advances in neural information processing systems*, *28*.

Ezeme, O. M., Mahmoud, Q. H., & Azim, A. (2020). Design and development of AD-CGAN: Conditional generative adversarial networks for anomaly detection. *IEEE Access*, *8*, 177667-177681.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*.

Gui, J., Sun, Z., Wen, Y., Tao, D., & Ye, J. (2021). A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE Transactions on Knowledge and Data Engineering*.

Guo, G., Yu, X., Liu, S., Ma, Z., Wu, Y., Xu, X., ... & Wu, X. (2021). Leakage detection in water distribution systems based on time–frequency convolutional neural network. *Journal of Water Resources Planning and Management*, *147*(2), 04020101.

Gupta, A., & Kulat, K. D. (2018). A selective literature review on leak management techniques for water distribution system. *Water resources management*, *32*(10), 3247-3269.

Hu, X., Zhang, H., Ma, D., & Wang, R. (2021). Hierarchical pressure data recovery for pipeline network via generative adversarial networks. *IEEE Transactions on Automation Science and Engineering*.

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).

Javadiha, M., Blesa, J., Soldevila, A., & Puig, V. (2019, April). Leak localization in water distribution networks using deep learning. In *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)* (pp. 1426-1431). IEEE.

Jung, D., Kang, D., Liu, J., & Lansey, K. (2015). Improving the rapidity of responses to pipe burst in water distribution systems: a comparison of statistical process control methods. *Journal of Hydroinformatics*, *17*(2), 307-328.

Kammoun, M., Kammoun, A., & Abid, M. (2022). Leak detection methods in water distribution networks: a comparative survey on artificial intelligence applications. *Journal of Pipeline Systems Engineering and Practice*, 13(3), 04022024.

Kang, J., Park, Y. J., Lee, J., Wang, S. H., & Eom, D. S. (2017). Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems. *IEEE Transactions on Industrial Electronics*, *65*(5), 4279-4289.

Kaplan, M. O., & Alptekin, S. E. (2020). An improved BiGAN based approach for anomaly detection. *Procedia Computer Science*, *176*, 185-194.

Kleijnen, J. P. (2017). Regression and Kriging metamodels with their experimental designs in simulation: a review. *European Journal of Operational Research*, *256*(1), 1-16.

Klise, K., Hart, D., Bynum, M., Hogge, J., Haxton, T., Murray, R., & Burkhardt, J. (2020). *Water Network Tool for Resilience (WNTR). User Manual, Version 0.2. 3* (No. SAND-2020-9301R; EPA/600/R-20/185). Sandia National Lab.(SNL-NM), Albuquerque, NM (United States).

Klise, K.A., Murray, R., Haxton, T. (2018). An overview of the Water Network Tool for Resilience (WNTR), In Proceedings of the 1st International WDSA/CCWI Joint Conference, Kingston, Ontario, Canada, July 23-25, 075.

Loureiro, D., Amado, C., Martins, A., Vitorino, D., Mamade, A., & Coelho, S. T. (2016). Water distribution systems flow monitoring and anomalous event detection: A practical approach. *Urban Water Journal*, *13*(3), 242-252.

Luo, J., Huang, J., & Li, H. (2021). A case study of conditional deep convolutional generative adversarial networks in machine fault diagnosis. *Journal of Intelligent Manufacturing*, *32*(2), 407-425.

Menapace, A., Avesani, D., Righetti, M., Bellin, A., & Pisaturo, G. (2018). Uniformly distributed demand EPANET extension. *Water resources management*, *32*(6), 2165-2180.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

Mounce, S. R., Mounce, R. B., & Boxall, J. B. (2011). Novelty detection for time series data analysis in water distribution systems using support vector machines. *Journal of hydroinformatics*, *13*(4), 672-686.

Mounce, S. R., Boxall, J. B., & Machell, J. (2010). Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows. *Journal of Water Resources Planning and Management*, *136*(3), 309-318.

Mu, H., Sun, R., Yuan, G., & Wang, Y. (2021). Abnormal Human Behavior Detection in Videos: A Review. *Information Technology and Control*, *50*(3), 522-545.

Murphy, B. S. (2014, December). PyKrige: development of a kriging toolkit for Python. In *AGU fall meeting abstracts* (Vol. 2014, pp. H51K-0753).

Panda, M., & Khilar, P. M. (2015). Distributed self fault diagnosis algorithm for large scale wireless sensor networks using modified three sigma edit test. *Ad Hoc Networks*, *25*, 170-184.

Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys (CSUR)*, *54*(2), 1-38.

Perez, R., Sanz, G., Puig, V., Quevedo, J., Escofet, M. A. C., Nejjari, F., ... & Sarrate, R. (2014). Leak localization in water networks: A model-based methodology using pressure sensors applied to a real network in Barcelona [applications of control]. *IEEE control systems magazine*, *34*(4), 24-36.Qasim, A. B., Ezhov, I., Shit, S., Schoppe, O., Paetzold, J. C., Sekuboyina, A., ... & Menze, B. (2020, September). Red-GAN: Attacking class imbalance via conditioned generation. Yet another medical imaging perspective. In *Medical Imaging with Deep Learning* (pp. 655-668). PMLR.

Qiu, Y., Misu, T., & Busso, C. (2022). Driving Anomaly Detection Using Conditional Generative Adversarial Network. *arXiv preprint arXiv:2203.08289*.

Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

Rajabi, M. M., Javaran, M. R. H., Bah, A. O., Frey, G., Le Ber, F., Lehmann, F., & Fahs, M. (2022). Analyzing the efficiency and robustness of deep convolutional neural networks for modeling natural convection in heterogeneous porous media. *International Journal of Heat and Mass Transfer*, *183*, 122131.

Romano, M., Kapelan, Z., & Savic, D. (2012). Automated detection of pipe bursts and other events in water distribution systems. American Society of Civil Engineers.

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.

Sanz, G., Pérez, R., Kapelan, Z., & Savic, D. (2016, September). Leak detection and localization through demand components calibration. American Society of Civil Engineers.

Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., & Schmidt-Erfurth, U. (2019). f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. *Medical image analysis*, *54*, 30-44.

Singh, N. K., & Raza, K. (2021). Medical image generation using generative adversarial networks: A review. *Health informatics: A computational perspective in healthcare*, 77-96.

Soldevila, A., Fernandez-Canti, R. M., Blesa, J., Tornil-Sin, S., & Puig, V. (2017). Leak localization in water distribution networks using Bayesian classifiers. Journal of Process Control, 55, 1-9.

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

Steffelbauer, D., Deuerlein, J., Gilbert, D., Piller, O., & Abraham, E. (2020, September). Dual Model for Leak Detection and Localization. In *CCWI/WDSA 2020*.

Sun, C., Parellada, B., Puig, V., & Cembrano, G. (2019). Leak localization in water distribution networks using pressure and data-driven classifier approach. *Water*, *12*(1), 54.

Tariq, S., Bakhtawar, B., & Zayed, T. (2022). Data-driven application of MEMS-based accelerometers for leak detection in water distribution networks. *Science of The Total Environment*, *809*, 151110.

Tijani, I. A., Abdelmageed, S., Fares, A., Fan, K. H., Hu, Z. Y., & Zayed, T. (2022). Improving the leak detection efficiency in water distribution networks using noise loggers. *Science of the Total Environment*, *821*, 153530.

Tu, Y., Lin, Y., Wang, J., & Kim, J. U. (2018). Semi-supervised learning with generative adversarial networks on digital signal modulation classification. *Comput. Mater. Continua*, *55*(2), 243-254.

Vrachimis, S. G., Eliades, D. G., Taormina, R., Kapelan, Z., Ostfeld, A., Liu, S., ... & Polycarpou, M. M. (2022). Battle of the leakage detection and isolation methods. *Journal of Water Resources Planning and Management*, *148*(12), 04022068.

Vrachimis, S. G., Eliades, D. G., Taormina, R., Ostfeld, A., Kapelan, Z., Liu, S., ... & Polycarpou, M. M. (2020). BattLeDIM: Battle of the leakage detection and isolation methods. In *Proc., 2nd Int. CCWI/WDSA Joint Conf.*

Wan, X., Kuhanestani, P. K., Farmani, R., & Keedwell, E. (2022). Literature Review of Data Analytics for Leak Detection in Water Distribution Networks: A Focus on Pressure and Flow Smart Sensors. *Journal of Water Resources Planning and Management*, *148*(10), 03122002.

Wang, X., Guo, G., Liu, S., Wu, Y., Xu, X., & Smith, K. (2020). Burst detection in district metering areas using deep learning method. *Journal of Water Resources Planning and Management*, *146*(6), 04020031.

Wang, Y., Yu, B., Wang, L., Zu, C., Lalush, D. S., Lin, W., ... & Zhou, L. (2018). 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. *Neuroimage*, *174*, 550-562.

Wang, K., Gou, C., Duan, Y., Lin, Y., Zheng, X., & Wang, F. Y. (2017). Generative adversarial networks: introduction and outlook. *IEEE/CAA Journal of Automatica Sinica*, *4*(4), 588-598.

Wang, Z., Lu, L., & Bovik, A. C. (2004). Video quality assessment based on structural distortion measurement. *Signal processing: Image communication*, *19*(2), 121-132.

Wu, Y., & Liu, S. (2020). Burst detection by analyzing shape similarity of time series subsequences in district metering areas. *Journal of Water Resources Planning and Management*, *146*(1), 04019068.

Wu, Y., & Liu, S. (2017). A review of data-driven approaches for burst detection in water distribution systems. *Urban Water Journal*, *14*(9), 972-983.

Wu, Y., Liu, S., Wu, X., Liu, Y., & Guan, Y. (2016). Burst detection in district metering areas using a data driven clustering algorithm. *Water research*, *100*, 28-37.

Wunderlich, A., & Sklar, J. (2022). Learning Noise with Generative Adversarial Networks: Explorations with Classical Random Process Models. *arXiv preprint arXiv:2207.01110*.

Xia, X., Pan, X., Li, N., He, X., Ma, L., Zhang, X., & Ding, N. (2022). GAN-based anomaly detection: A review. *Neurocomputing*.

Xu, Z., Ying, Z., Li, Y., He, B., & Chen, Y. (2020). Pressure prediction and abnormal working conditions detection of water supply network based on LSTM. *Water Supply*, 20(3), 963-974.

Ye, G., & Fenner, R. A. (2011). Kalman filtering of hydraulic measurements for burst detection in water distribution systems. *Journal of pipeline systems engineering and practice*, *2*(1), 14-22.

Ye, G., & Fenner, R. A. (2014). Weighted least squares with expectation-maximization algorithm for burst detection in UK water distribution systems. *Journal of Water Resources Planning and Management*, 140(4), 417-424.

Zamora, Y. M., Hernández-Callejo, L., Duque-Pérez, O., & Alonso-Gómez, V. (2021). Diagnosis of Broken Bars in Wind Turbine Squirrel Cage Induction Generator: Approach Based on Current Signal and Generative Adversarial Networks. *Applied Sciences*, *11*(15), 6942.

Zhang, H., Hu, X., Ma, D., Wang, R., & Xie, X. (2020). Insufficient data generative model for pipeline network leak detection using generative adversarial networks. *IEEE Transactions on Cybernetics*.

Zhang, Y. Y., Shen, C. M., Feng, H., Fletcher, P. T., & Zhang, G. X. (2019). Generative adversarial networks with joint distribution moment matching. *Journal of the Operations Research Society of China*, *7*(4), 579-597.

Zheng, M., Li, T., Zhu, R., Tang, Y., Tang, M., Lin, L., & Ma, Z. (2020). Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification. *Information Sciences*, *512*, 1009-1023.

Zhou, X., Tang, Z., Xu, W., Meng, F., Chu, X., Xin, K., & Fu, G. (2019). Deep learning identifies accurate burst locations in water distribution networks. *Water research*, *166*, 115058.