Water leak detection using self-supervised time series classification

Ane Blázquez-García[a,*], Angel Conde[a], Usue Mori[b], Jose A. Lozano[b,c]

[a] *Ikerlan Technology Research Centre, Basque Research and Technology Alliance (BRTA). Pº. J. Mª. Arizmendiarrieta 2, 20500 Arrasate/Mondragón, Spain.*
[b] *Intelligent Systems Group, Department of Computer Science and Artificial Intelligence, University of the Basque Country UPV/EHU. Manuel Lardizabal Ibilbidea 1, 20018 Donostia/San Sebastián, Spain.*
[c] *Basque Center for Applied Mathematics (BCAM). Alameda de Mazarredo 14, 48009 Bilbao, Spain.*

Abstract

Leaks in water distribution networks cause a loss of water that needs to be compensated to ensure a continuous supply for all customers. This compensation is achieved by increasing the flow of the network, which entails an undesirable economical expense as well as negative consequences for the environment. For these reasons, detecting and fixing leaks is a relevant task for water distribution companies. This paper proposes a water leak detection method based on a self-supervised classification of flow time series. The aim is to detect the leaks in the network, providing a low false positive rate. The proposed method is applied to two water distribution networks and compared to two other methods in the literature, obtaining the best balance between the number of false positives and detected leaks.

Keywords: leak detection, anomaly detection, time series, self-supervised classification

1. Introduction

Water leaks are of special interest for water distribution companies, not only because of the economic loss they entail but also because of the environmental

*Corresponding author
Email address: ablazquez@ikerlan.es (Ane Blázquez-García)

consequences they generate. A water leak in a water distribution network consists of an accidental escape of water through a component of the network (e.g., a hole or a crack). Consequently, in order to continue to supply all customers with enough water, the flow of the water in the network needs to be increased until the leak is fixed.

An example of a water leak on a water distribution network in Yorkshire[1] is depicted in Fig. 1, where the shaded area indicates the day the leak was repaired. Indeed, a remarkable increase in the flow can be appreciated in the data between the 20^{th} and 23^{rd} of November. In particular, leaks are usually more noticeable during the night (from 1 a.m. to 5 a.m., highlighted in orange in Fig. 1), when customer demand is low [1].

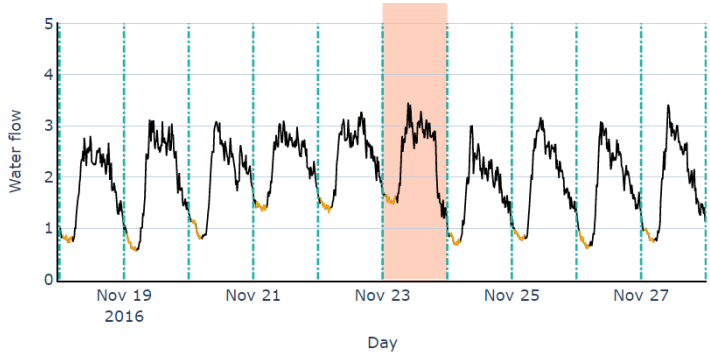


Figure 1: Example of the increase that a leak has caused in the water flow.

Until recently, leakage management procedures within the water industry tended to be predominantly resource-intensive manual processes [2]. These methods are also known as hardware-based methods [3, 4] and are based on using specialized hardware equipment, such as leak noise correlators, leak noise loggers, and gas injection. Even though these methods are very accurate, they are very expensive and not practical on a day-to-day basis.

Technological advances in recent years have brought major breakthroughs in data collection, enabling a large amount of data to be gathered. This has led

[1] https://datamillnorth.org/dataset/yorkshire-water-leakage-dma-15-minute-data

to the development of new automatic and effective data-driven leak detection techniques, which have been denominated software-based methods [3, 4, 5] and have become increasingly popular in recent years. These techniques are typically applied to hydrological data such as water flow or pressure data, although some methods use multiple variables simultaneously [6, 7, 8, 9, 10], or even include other types of data (e.g., data collected by acoustic emission sensors [11]).

Leak detection is not a trivial task, however, and software-based techniques also have their limitations. To begin with, these techniques rely on data, and so they will only be able to detect the leaks that are reflected in the data. Thus, they will usually not detect structural leaks or leaks that have existed from periods prior to the data collection. Typically, other alternatives, such as periodic inspections on the part of the company or those resulting from external customer calls, are needed to detect these leaks.

Additionally, when using hydrological data, leaks can often be mistaken for sporadic large consumption (e.g., filling pools in summer or occasional high consumption in industry) because both involve an unusual increase in the flow in the network. However, the high consumption of consumers tends to be sporadic, whereas leakage continues over time and causes a permanent increase in the flow until it is fixed.

In the literature, some of the software-based methods proposed for leak identification are based on supervised classification [7, 9, 11, 12], where the idea is to learn a classifier that discriminates between leakage and non-leakage periods of time.

The main drawback of these techniques is that leaks must be identified and labeled to form the training set. However, obtaining a sufficiently large number of leakage samples is complicated in real circumstances because leaks are typically not very common. Moreover, leak identification is normally carried out using reports from the water company. These reports do not always include all the existing leaks and are frequently diffuse and uncertain regarding the recorded leaks. All this largely complicates the learning of reliable supervised classification models for leak detection.

As an alternative to the supervised classification techniques, some authors use unsupervised software-based methods, which do not need labels at training time. These methods usually assume that the training set is only composed of normal (no leakage) data.

The most widespread unsupervised method relies on the analysis of the measured Minimum Night Flow (MNF), which is the lowest flow supplied to an area during the night (e.g., 00:00–05:00 [2, 7] or 02:00–04:00 [1, 12]). A leak alarm is generated when the MNF exceeds a threshold typically set by water utilities. Although this method uses only nightly data when the customer demand is usually low, it analyzes single points, and so leaks can still be confused with occasional high night consumption or sensor failures. In this context, despite the simplicity and intuitiveness of this method, it provides many false leak alarms, which involves additional effort in the search for leaks that do not exist. Additionally, false alarms burden workers in water companies and there is a risk that workers will start ignoring alerts. Therefore, since leaks rarely occur in water distribution networks, it is desirable that the detection methods maintain a small number of false alarms even at the expense of reducing the number of detected leaks to within an acceptable range [5].

More sophisticated unsupervised methods to detect leaks rely on fitting a prediction model to normal data to predict values over time and identify the observations that significantly deviate from their predicted values. Even if the data have an evident temporal nature, most methods do not take this aspect into account or partially consider it using time windows [2, 6, 10, 13, 14, 15, 16, 17].

The prediction-based techniques proposed in the literature typically make one-step-ahead (point) predictions [8, 10, 14, 15, 16, 17, 18, 19] rather than predicting full time windows (e.g., one-day-ahead prediction) [2, 6, 13], and use threshold values to make comparisons and identify large deviations from normality. In particular, some of these methods build a model for each time step, thus generating a large number of models [8, 14, 18, 19]. Also, since point analysis produces false alarms more easily due to sudden high water usage or sensor faults, some prediction-based techniques analyze consecutive residuals

[14, 15, 16, 19] to avoid confusing leakage with occasional high consumption, or apply a denoising approach before building the model to remove the noise caused by sensor faults [15, 16].

Less commonly, other unsupervised techniques which are not based on prediction models have also been proposed. Some of them identify leakage days by projecting the flow data into a space of lower dimension to detect the projections that lie outside control limits [20]. Others attempt to learn the normal behavior of flow subsequences using one-class classification [17]. Similarity-based techniques have also been used to identify flow points [21] or subsequences [22, 23, 24] with low similarity with respect to a reference of normality. These types of techniques do not usually consider the temporal correlation of the observations beyond the inclusion of temporal features (e.g., the hour of the day) or the use of features extracted within time windows [17, 20, 22, 23]. Moreover, as with prediction-based techniques, some of these methods also build a model for every time step [21, 24], thus generating a large number of models.

In this article, we propose a novel water leak detection technique based on the self-supervised classification of time series, which, to the best of our knowledge, has not been used in this context before. Self-supervised learning refers to learning methods that exploit the structure of unlabeled data to provide appropriate supervision signals and thus define a new problem that can be solved from a supervised perspective. Although this general learning approach has been previously used in the literature, mainly in the field of computer vision, the methodology proposed in this paper is the first to address the general problem of anomaly detection and, in particular, the problem of leak detection, based on the philosophy of self-supervision and using time series data.

The aim is to detect nights with water leakage by learning a model of the normal behavior using only flow data. This approach does not require a statistical or hydraulic model to be fitted to the data, nor does it require leakage labels in the training phase. Moreover, the proposed method considers the contextual information by analyzing full time series rather than point observations, and consequently, it does not require a model for every time step as other methods

in the literature do. Additionally, the temporal nature of the data is taken into account by using specific classifiers for time series, instead of using time windows or temporal features, as most of the methods in the literature do. Consequently, the proposed method succeeds in detecting a high number of leaks while providing a low number of false alarms.

The rest of the paper is organized as follows. In Section 2, the proposed methodology is described. Section 3 provides the experimental results. Finally, in Section 4, the conclusions and future research lines are outlined.

2. Methodology

The proposed method, *Self-Supervised Leak Detector (SSLD)*, aims to detect water leaks by learning the normal behavior of the water flow in a water distribution network. Self-supervised learning involves training a model that does not require external class labels and instead uses labels that have been assigned to artificially generated data. This is especially useful in problems where there are very few or no available labels, as is the case of water leak detection.

In this paper, we use the self-supervised approach within the anomaly detection framework [25] (see Fig. 2). The idea is to define a self-labeled training dataset to learn a classifier that will be used to detect the anomalies. The self-labeled dataset is built by applying a set of different transformations to the initial training instances, which are supposed to represent the normality, and by assigning a label, usually denominated *pseudo-label*, to the data instance obtained from each transformation. This self-labeled dataset is then used to learn a classifier that discriminates between the different transformations applied. Finally, the learned classifier will be used to determine whether new samples are anomalous based on a decision rule.

The rationale behind this approach is that learning to distinguish between the different transformations applied to the data also helps to learn features that are likely to be unique to normal samples: since these features do not appear in the same way in anomalous samples, the classifier will fail to discriminate

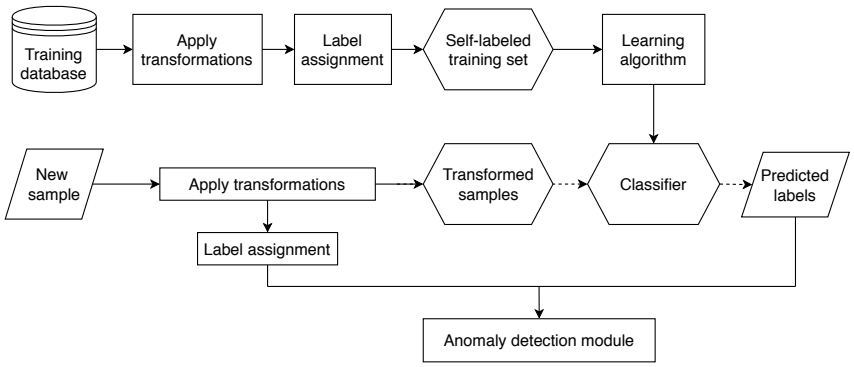


Figure 2: Illustration of the self-supervised approach for anomaly detection.

between the applied transformations. Indeed, the type of data at hand and the characteristics of the anomalies to be detected will be determinant when defining the transformations in this self-supervised framework. For example, in the context of images, self-supervision has recently been used to identify abnormal images [26, 27, 28] by applying transformations, such as geometrical transformations, to normal images.

2.1. Generation of the self-labeled dataset

In the first step of our self-supervised framework, a self-labeled training set is generated. To do this, for each univariate time series of nightly flow data $X = \{x_t\}_{t=1}^{n}$ of length n in the initial training set, K time series are artificially generated, each one with an associated pseudo-label. These artificial time series are obtained by applying K different linear transformations (g_i where $i \in \{1, ..., K\}$), which are defined by:

$$g_i : \mathbb{R}^n \to \mathbb{R}^n$$
$$X \mapsto (p_1 + ... + p_i)X$$

where $p_1 = 1, p_2, ..., p_K \in \mathbb{R}^+$, and $g_i(X) = X^i = \{(p_1 + ... + p_i)x_t\}_{t=1}^{n}$. Note that the first pseudo-label corresponds to the original non-transformed time series. An example of this procedure is shown in Fig. 3, where it can be seen that the different linear transformations increase the night water flow at different levels.

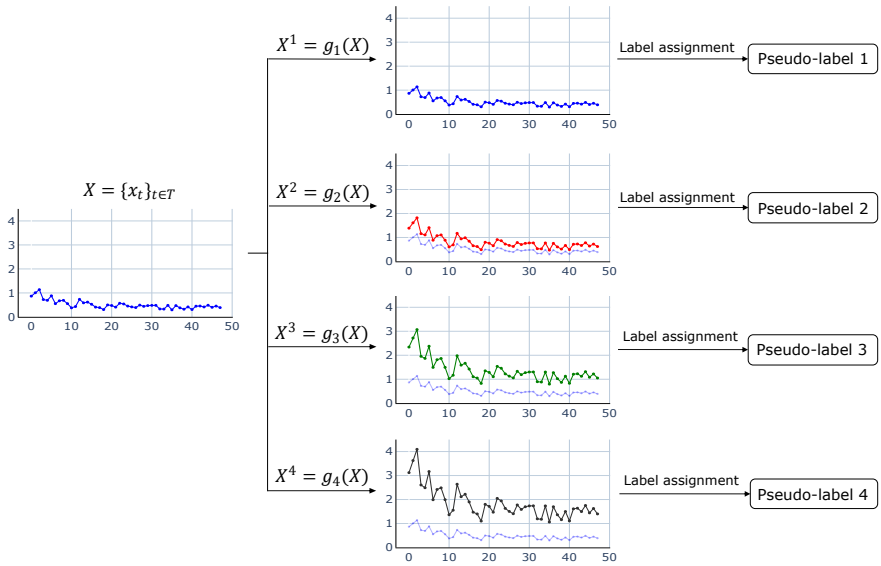


Figure 3: Example of the generation of the self-labeled dataset, where $K = 4$, $p_1 = 1$, $p_2 = 0.6$, $p_3 = 1.1$, $p_4 = 0.9$.

The premise behind applying these linear transformations is that water leaks are characterized by a flow increase. Consequently, a night with a leak will resemble a normal night with an increase in flow. As such, since it is assumed that the classifier will be learned using only normal flow data, leak nights are expected to be assigned with pseudo-labels from 2 to K, depending on the severity of the leak, instead of being labeled with pseudo-label 1, which indicates that no transformation has been applied. Along the same line, classification errors will also be committed with the transformed time series obtained from a night with a leak because it is expected that the classifier will incorrectly assign higher pseudo-labels, which represent higher levels of flow.

Once the transformations have been defined, the training set that will be used to learn the self-supervised classifier has to be generated, as shown in Fig. 2. Let $\Omega = \{X_1, ..., X_m\}$ be the initial training database consisting of m normal time series (with no leakage). Then, $\mathbf{T} = \{(\mathrm{X}_1^1, \mathrm{y}_1^1 = 1), ..., (\mathrm{X}_1^K, \mathrm{y}_1^K =$

$K), ..., (\mathrm{X}_m^1, \mathrm{y}_m^1 = 1), ..., (\mathrm{X}_m^K, \mathrm{y}_m^K = K)\}$ is the generated set of $m \times K$ time series with their associated class labels, where X_i^j is the j^{th} transformed time series of X_i and j is its associated label indicating the applied transformation (see Fig. 3). $\mathbf{T}$ is the training set that will be used to learn the classifier.

It should be noted that, in practice, it is difficult to ensure that the initial training set is only composed of normal flow data. In principle, we can discard identified leaks from this set if the information is available, but there may remain leaks that have gone unnoticed. However, leakage occurs very infrequently, so the training set will be predominated by non-leakage days.

2.2. Construction of the classifier

The aim of this step is to train a classifier $\mathcal{F}$ that learns to discriminate between the K linear transformations applied in the previous step. In other words, the purpose is to learn the mapping between each input time series and its corresponding label. In order to consider the temporal nature of the data at hand, and contrary to other leak detection methods in the literature, in this paper we adopt a time series classification approach. This allows the contextual information to be considered rather than single observations, which is particularly helpful when detecting leaks because they remain over time.

Within the existing time series classifiers, we have chosen the Random Interval Spectral Ensemble (RISE) [29] due to the accuracy shown in other problems in the literature and also due to its robustness to noise, which makes it especially useful for the problem of leak detection in water distribution networks given that flow data are particularly noisy. RISE is an ensemble time series classification algorithm that consists of building a set of trees, each of which focuses on a randomly chosen time interval (subsequence) of the data. In particular, this method extracts spectral features over each random interval to learn the time series forest. It should be noted that our methodology is not limited to the RISE classifier, and thus, other classifiers could also be used.

2.3. Deployment of the model for anomaly (leak) detection

As shown in Fig. 2, to apply the learned classifier to a new time series X_{new}, first, the K transformations have to be applied: $\{X^1_{new} = g_1(X_{new}), ..., X^K_{new} = g_K(X_{new})\}$. Then, the classifier is used to predict the label of each generated time series ($\{X^1_{new} \xrightarrow{\mathcal{F}} \hat{y}^1_{new}, ..., X^K_{new} \xrightarrow{\mathcal{F}} \hat{y}^K_{new}\}$).

Leaks are characterized by higher levels of flow, and so, it is expected that their pseudo-labels will not be correctly predicted by the self-supervised classifier. However, the severity and typology of the leak could have an influence on the output of the classifier. For example, small leaks could result in fewer misclassified pseudo-labels or different misclassification patterns in comparison to very severe leaks. Similarly, nights with no leaks could also present a few misclassified labels due to errors in the classifier or variations in the normal patterns. In this context, a decision rule that will determine whether a night has suffered a leak must be defined based on the outputs of the self-supervised classifier in all its transformations.

In our method, the new time series X_{new} is flagged as an anomaly if the classifier assigns pseudo-label K to at least two of the transformed time series of X_{new}. That is,

$$|\{i \mid \hat{y}^i_{new} = K, i \in \{1, ..., K\}\}| \geq 2. \tag{1}$$

Note that the K^{th} transformation provides the highest increase in the flow data. Since leaks represent an increase in the flow, we expect the predictions of the labels to be displaced upwards, so the classifier should assign the largest pseudo-label to at least two of the generated time series.

3. Experimentation

The evaluation of the proposed method has been performed in both a private (scenario A) and a public (scenario B) dataset of different water distribution networks. The data used in the experiments are presented in Section 3.1, the implementation details and parameter selection of the proposed methodology

in Section 3.2, the evaluation framework and metrics in Section 3.3, the methods used to compare the results of our method in Section 3.4, and finally the experimental results are outlined in Section 3.5.

3.1. Data

The available datasets in both scenarios consist of water flow measurements recorded over time at a certain granularity and divided into different zones of a particular locality. From all these data, the time series that will be used as input for the model consists of the flow measurements collected during the night period from 1 a.m. to 5 a.m., when water consumption is usually the lowest.

In addition to the flow measurements, both scenarios also contain information about the leaks, indicating the date on which a work report associated with the repair of each leak has been generated. Some of the recorded leaks refer to structural leaks or those that are not reflected in the data. Thus, from all the leaks, we only focus on those that correspond to the *detectable leaks* or, in other words, the leaks that generate a noticeable increase in the flow data. In particular, we will only consider those leaks that make the MNF higher than usual. In this context, and following the recommendations of the experts, the data used to calculate the MNF are collected between 2 a.m. and 5 a.m.

To sum up, each zone within each scenario has both an associated time series database of nightly flow data and the work reports of the leaks. Since the flow behavior for each type of day is different, we divide the time series database into seven smaller databases, one for each day of the week. Furthermore, as shown in Fig. 4, each of these databases is divided into two new sets: the training dataset, which will be used to train the models, and the leakage dataset, which will be used to validate the models. To build the leakage dataset, the nights with a leakage record are separated from those with a non-leakage record. Since not all the leakage records represent detectable leaks, a predefined threshold value is used to select the nights on which the MNF exceeds it. These nights form the leakage dataset.

Regarding the training dataset, it should be composed only of normal days

because our approach aims to learn the normal behavior of the water flow. However, even after separating the nights with a leakage record from the training dataset, there still might be unusual days that may misguide the classifier and should be discarded. Some of these unusual days are related to the leakage records: since a leakage report refers to the date when the leak is fixed, and not when it appeared, we also remove the three days prior to the leakage report, in accordance with expert guidance. Holidays might also show a different behavior, and thus, they have also been excluded from the training sets. The nights after removing both those unusual days and the nights with a leakage record form the final training dataset.

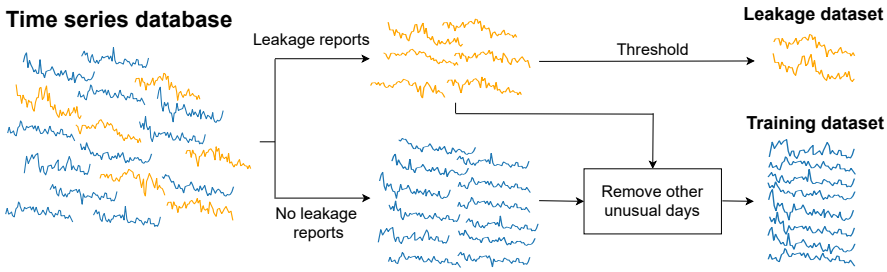


Figure 4: The training and leakage datasets considered in the experimentation.

3.1.1. Scenario A

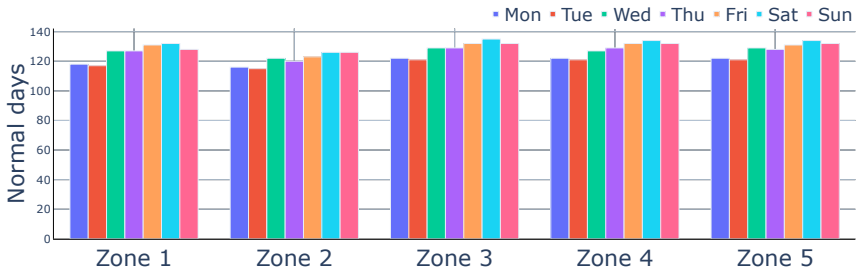
The dataset in this scenario consists of water flow measurements collected every 5 minutes over three years (2017–2019) in a private water distribution network located in Azkoitia, a town in the Basque Country. This network, which is managed by *Gipuzkoako Urak S.A.*[2], has five zones that measure the flow in different areas of the town, with at least one detectable leakage record in the study period. Each zone has very distinct flow patterns, and so we apply the proposed method to each zone individually.

As additional useful information, this dataset contains information about

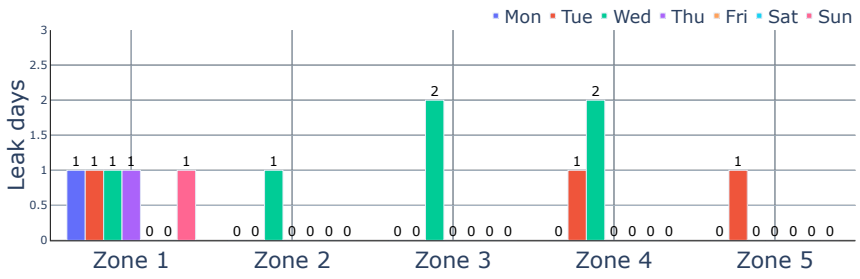
[2] https://www.gipuzkoakour.eus/

the threshold value established by the water utility company, which is used to raise alarms in the MNF approach and is based on the annual flow values and other objectives set by the company (e.g., reducing the cost of the water loss every year). Recall that this threshold is also used to select the detectable leaks and thus define the leakage dataset.

Fig. 5a shows the number of days that are considered to be normal and thus belong to each training set. Similarly, Fig. 5b presents the number of reported leakage days used to test the models. Note the challenge of addressing this as a supervised problem due to the small number of detectable leakage records per training set.



(a) Number of normal days in each training set.



(b) Number of detectable leaks in each leakage dataset according to the provided threshold value.

Figure 5: Number of samples in the training and leakage sets in scenario A.

3.1.2. Scenario B

Yorkshire Water is a water supply and treatment utility company in England that has released its data related to different domains, such as pollution, consumption, and leakage. In this paper, we use the leakage dataset[1], which contains the records of the water flow of more than 2000 Distribution Management Areas (DMAs) organized in 20 regions, with a 15-minute granularity for one year (April 2016–April 2017).

In contrast to scenario A, the water utility company does not explicitly provide any threshold for the analysis of the MNF. Thus, we define this threshold as a percentile value ($80, 85, 90$, and 95) of the MNF data of the training sets.

Due to the high number of DMAs, a single region (called E1 in the original database) has been chosen for the purpose of this paper. Even if this region is composed of 117 DMAs, we have specifically selected the DMAs where:

a) there are no "invalid" flow records[3];

b) data are available for the whole year.

c) all flow records are non-negative.

d) there is at least one day with a detectable leakage record.

Note that the chosen threshold value will directly influence the number of detectable leaks and thus the size of the leakage dataset. In particular, the lower the threshold value, the greater the number of detectable leaks and the more DMAs that will be analyzed.

Taking this into account, Table 1 shows the final number of DMAs to be analyzed for the different threshold values. This table also describes the training and leakage datasets for each percentile and type of day, presenting the mean number of time series per DMA, together with the standard deviation. As an example, 24 DMAs have been analyzed with the 80^{th} percentile, and on

[3]The dataset contains an extra column called 'Flow Validity Code' that indicates if the record is valid ('V'), invalid ('I') or missing ('M').

average, these DMAs have 43.83 series in the training set used for Mondays and 0.67 series in the leakage set used for the same type of day. As the standard deviation is small, the actual number of elements in each DMA is close to these average values.

For all types of day, the number of nights in the training sets is quite uniform. Regarding the number of leaks, the DMAs contain very few leaks per type of day, and, in total, between 2 and 3 detectable leaks on average. Note that the percentiles 90 and 95 provide the same training datasets but different leakage datasets.

Percentile	Analyzed DMAs	Set	Mon	Tue	Wed	Thu	Fri	Sat	Sun
80	24	Training	43.83 (2.43)	47.83 (2.37)	49.17 (1.79)	49.46 (1.98)	50.38 (2.14)	49.17 (1.99)	48.25 (1.89)
		Leakage	0.67 (0.92)	0.42 (0.72)	0.54 (0.72)	0.50 (0.78)	0.29 (0.55)	0.33 (0.48)	0.25 (0.53)
85	23	Training	43.78 (2.47)	47.70 (2.32)	49.09 (1.78)	49.35 (1.94)	50.30 (2.16)	49.13 (2.03)	48.17 (1.90)
		Leakage	0.61 (0.94)	0.35 (0.57)	0.57 (0.73)	0.52 (0.79)	0.30 (0.56)	0.35 (0.49)	0.22 (0.52)
90	22	Training	43.64 (2.42)	47.55 (2.26)	49.05 (1.81)	49.27 (1.96)	50.23 (2.18)	49.09 (2.07)	48.09 (1.90)
		Leakage	0.59 (0.91)	0.36 (0.58)	0.45 (0.51)	0.41 (0.73)	0.32 (0.57)	0.27 (0.46)	0.18 (0.39)
95	22	Training	43.64 (2.42)	47.55 (2.26)	49.05 (1.81)	49.27 (1.96)	50.23 (2.18)	49.09 (2.07)	48.09 (1.90)
		Leakage	0.55 (0.86)	0.27 (0.46)	0.41 (0.50)	0.41 (0.73)	0.23 (0.43)	0.27 (0.46)	0.18 (0.39)

Table 1: Description of the training and leakage datasets for each threshold value. The mean number of samples per DMA and day is shown, together with the standard deviation between parentheses.

3.2. Implementation details

The number of transformations that will be applied to build the self-labeled training sets has been set to $K = 4$ (four possible pseudo-labels), and the transformation parameters initially to $p_1 = 1$, and $p_2 = p_3 = p_4 = 0.7$. Note that very small transformation parameters would hinder the differentiation between these pseudo-labels, thus complicating the learning of the classifier, while very large parameters would make this difference too obvious, making it difficult to detect less severe leaks.

The parameters specified for the RISE classifier have been the number of trees, w, and the minimum interval length, l. While the number of trees $w \in \{10, 11, ..., 20\}$ has been chosen using a grid-search with 5-fold Cross-Validation

(5-fold CV), the minimum interval length has been set to be $l = min(16, n/2)$, where n is the length of the time series, taking [30] as a reference.

Finally, as mentioned on more than one occasion, our method is a data-based one that can only detect leaks that generate a noticeable increase in the flow data (*detectable leaks*). As such, only the nights on which the MNF is higher than usual (i.e., MNF$> \tau$, where τ is a predefined threshold) will be introduced into our self-supervised classifier. In the deployment step, the nights which do not have an appreciable increase in the flow (MNF$\leq \tau$) will be directly labeled as "non-leak nights", without undergoing further analysis. In summary, our methodology can be seen as a 2-stage approach, where in the first stage, a night is predicted as a detectable leak or not using a threshold on the MNF, and in the second stage, only the detectable leaks are classified using SSLD to eliminate some of the false positives.

3.3. Evaluation framework and metrics

The evaluation of the proposed method is performed in terms of both the False Positive Rate (FPR), which estimates the false leak alarms generated in the training dataset, and the True Positive Rate (TPR), which measures the capacity to detect leaks in the leakage dataset (both datasets defined in Fig. 4).

As shown in Fig. 6, the number of false alarms or the FPR is estimated using a nested 5-fold CV in the full training set, in which the inner loop is used to tune the w parameter of the RISE classifier, and the outer loop is used to estimate the FPR. More specifically, the training set is split into 5 folds, and for each iteration $i \in \{1, ..., 5\}$, 4 of the folds are used to perform a grid search over the w parameter of the RISE classifier, using another 5-fold CV. The metric used in the inner loop is the classification accuracy, which computes the rate of correctly classified transformations. Once w has been selected, a model is trained using those 4 folds, and the resulting model is validated on the remaining fold to compute the FPR.

Similarly, to estimate the number of detected leaks or the TPR, first, the w parameter is tuned using a 5-fold CV in the training set, also based on the

classification accuracy. Note that this accuracy is not calculated in terms of leak/no leaks, but in terms of the pseudo-labels created in the self-supervised context. Then, the final model is learned using all the training data and the selected best w parameter. This model is applied to the leakage set to compute the TPR.

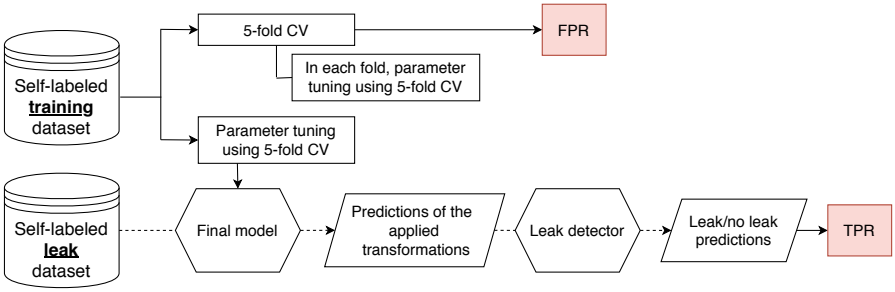


Figure 6: Evaluation framework of the proposed methodology.

3.4. Comparison with other methods

The results of our method have been compared with two other unsupervised techniques. One of these techniques, which we will consider the baseline, is the traditional MNF method [1, 2, 7, 12]. Even though this method identifies all the detectable leaks by default, it usually provides a high number of false alarms. In particular, our method aims to reduce this quantity of alarms while maintaining a high number of detected leaks. Note that the FPR of the SSLD will never be greater than that provided by the MNF method because, as explained in Section 3.2, we will automatically label as non-leakage all the nights that have no noticeable increase in the flow (MNF$\leq \tau$).

We also compare our method with the ϵ-SVR prediction-based method [14] as prediction-based methods have been widely used to identify leaks in the literature. Unlike other prediction methods that do not provide the code and do not clearly state the implementation details, this method contains all the necessary specifications to reproduce it. The procedure followed and the parameter values established are outlined in [14]. We have applied this method to the nightly

time series, as in our method, with an embedding dimension of one hour.

3.5. Experimental results

In the following sections, the results of the experiments performed in both scenario A (Section 3.5.1) and scenario B (Section 3.5.1) are presented. In particular, we first provide a comparative analysis concerning the results of other methods and then a sensitivity analysis of the transformation parameters of the self-supervised framework.

3.5.1. Scenario A

The results for each zone and for the initially fixed transformation parameters ($p_1 = 1$, $p_2 = p_3 = p_4 = 0.7$) are summarized in Table 2. From the three methods that we have compared, MNF is the only method that can perceive all the detectable leaks. However, our proposed SSLD method provides significantly fewer false positives while still being able to identify almost all the leaks: taking into account all the zones, 91.67% of the leaks are detected. The reduction in the FPR is better appreciated in the *mean FPR* column in Table 2, which represents the average value of the FPR of all the day types for each zone. Also note that although the ϵ-SVR method obtains an even lower FPR than our SSLD method in most of the zones, it detects very few leaks. Moreover, in the only zone that it can detect all the leaks (Zone 5), it has a remarkably high FPR.

In conclusion, in finding a trade-off between the detected leaks and the number of false positives, our method is the most successful.

Analysis of the transformation parameters

Once we have analyzed the ability of our method to detect leaks while maintaining low levels of false positives, in this section, we study the robustness of our method to changes in the transformation parameters. In particular, while keeping the RISE parameters previously obtained, the aim is to analyze how the FPR and TPR change with respect to the parameters p_2, p_3, p_4 (recall that $p_1 = 1$). The transformation parameter space considered is

Zone	Method	FPR							Mean FPR	Mean TPR
		Mon	Tue	Wed	Thu	Fri	Sat	Sun		
Zone 1	SSLD	0.1699	0.1283	0.1588	0.1422	0.1678	0.1808	0.2277	0.1679	**1.0000**
	MNF	0.3040	0.2915	0.3164	0.3644	0.2672	0.3088	0.3460	0.3141	**1.0000**
	ϵ-SVR	**0.0763**	**0.0950**	**0.1108**	**0.0788**	**0.0684**	**0.0527**	**0.0800**	**0.0803**	0.2000
Zone 2	SSLD	0.1717	0.2000	0.2135	0.2167	0.1722	0.2222	0.2065	0.2004	**1.0000**
	MNF	0.2841	0.2435	0.2878	0.3250	0.3074	0.4840	0.8889	0.4030	**1.0000**
	ϵ-SVR	**0.0087**	**0.0087**	**0.0167**	**0.0083**	**0.0167**	**0.0000**	**0.0000**	**0.0084**	0.0000
Zone 3	SSLD	0.2199	0.1560	0.1556	0.1614	0.2055	0.2444	0.2269	0.1957	0.5000
	MNF	0.3186	0.3220	0.3319	0.3079	0.3104	0.3926	0.3863	0.3385	**1.0000**
	ϵ-SVR	**0.0083**	**0.0080**	**0.0160**	**0.0000**	**0.0077**	**0.0074**	**0.0000**	**0.0068**	0.0000
Zone 4	SSLD	0.1301	**0.0910**	0.1399	0.1465	0.1220	0.1123	**0.1132**	0.1221	**1.0000**
	MNF	0.1551	0.1573	0.1867	0.2221	0.1758	0.1487	0.1516	0.1711	**1.0000**
	ϵ-SVR	**0.0897**	0.1153	**0.1022**	**0.0938**	**0.1148**	**0.0815**	0.1275	**0.1036**	0.6667
Zone 5	SSLD	**0.0910**	**0.1410**	**0.1774**	**0.1966**	**0.1755**	**0.1595**	**0.1132**	**0.1506**	**1.0000**
	MNF	0.1314	0.1823	0.2003	0.2651	0.2285	0.1749	0.1203	0.1861	**1.0000**
	ϵ-SVR	0.9596	0.9833	0.9920	0.9626	0.9692	0.9923	1.0000	0.9799	**1.0000**

Table 2: Results in each zone of scenario A. The FPR values for each day and method are shown, along with the mean FPR and TPR values of all the models of each method.

$p_2, p_3, p_4 \in \{0.5, 0.7, 0.9, 1.1, 1.3, 1.5\}$. All possible combinations are considered, which total 216 combinations.

The FPR and TPR values obtained with different parameter combinations for each zone and model (type of day) are shown in Fig. 7. Due to the low number of leaks that the ϵ-SVR method detects, the results of this section are only compared with the MNF method.

As expected, our proposed method reduces the FPR of the MNF method, regardless of the value of the chosen transformation parameters (see Fig. 7). This reduction is particularly noticeable in zones and types of day where the MNF method provides many false alarms (e.g., weekend models of Zone 2), as it has more room for improvement.

Regarding the number of detected leaks, and only considering the types of days that have at least one leak, 52.96% of the parameter combinations are able to detect all of them in Zone 1, 79.63% in Zone 2, 0.46% in Zone 3, 85.42% in Zone 4, and 100% in Zone 5. The reason for obtaining worse results in Zones

1 and 3 is that both zones have small leaks that complicate their detection. If we analyze these zones in more detail, most parameter combinations in Zone 1 detect all the leaks on Mondays and Tuesdays (i.e., 81.02% on Mondays and 86.57% on Tuesdays), but the percentage is lower on the other types of days (30.09% on Wednesdays, 44.91% on Thursdays, and 22.22% on Sundays). In Zone 3, there are only two detectable leaks on Wednesdays, and although the method can hardly detect both of them, 24.07% of the combinations can detect at least one of the leaks.

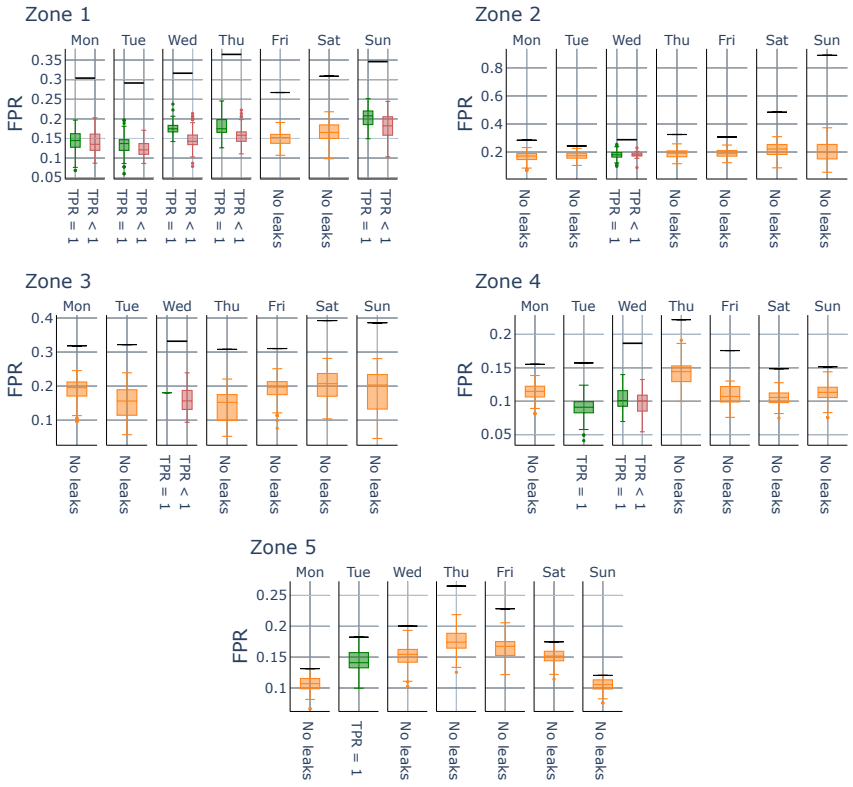


Figure 7: FPR and TPR of the models obtained with all the different transformation parameters. *NO LEAKS* indicates that there are no leaks in the leakage dataset for the given day of the week, $TPR = 1$ indicates that all the leaks have been detected, and $TPR < 1$ that the method has not identified all the leaks. The baseline FPR of the MNF method is highlighted with a black horizontal line.

In Fig. 8, the average performance of each parameter combination over all the models (day types) is studied. In particular, this figure shows the mean FPR obtained for different parameter combinations and highlights with a blue cross those that can detect all the leaks. Also, within the combinations that can detect all the leaks, the one that gives the lowest FPR is marked (e.g., $p_1 = 1$, $p_2 = 0.5$, $p_3 = 1.3$ and $p_4 = 0.7$ in Zone 1). As expected, all the transformation parameters reduce the mean FPR value in comparison to the MNF method, and even though some are not able to detect all the leaks, considering all zones together, 61.57% of combinations detect all of them.

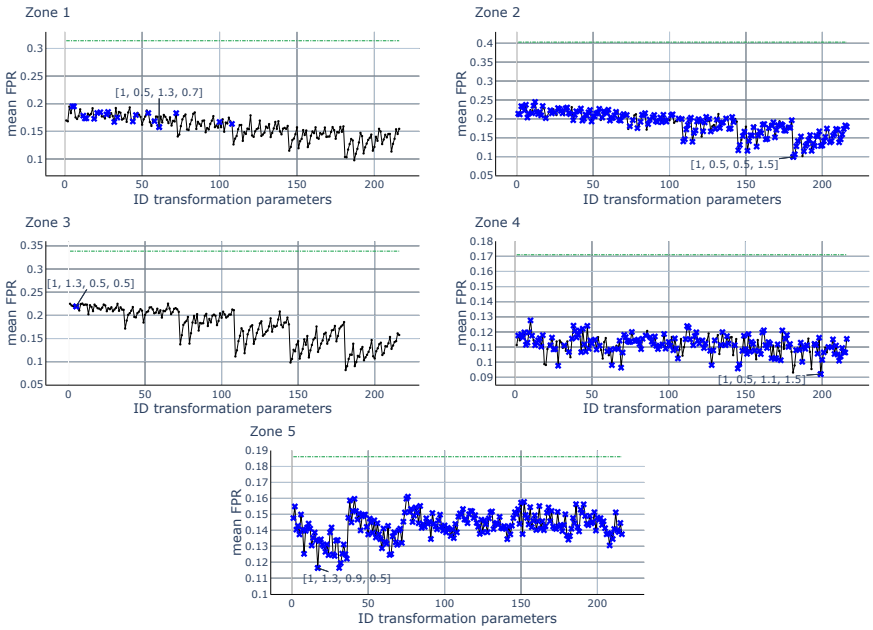


Figure 8: Mean FPR and TPR according to different transformation parameters in each zone. The IDs assigned to the transformation parameters in this figure are ordered based on a triple loop of the parameters in ascending order. That is, the first six points in the graphs indicate the parameter combinations $[1, p_2, 0.5, 0.5]$, where p_2 traverses all the parameter space in ascending order, the next six points refer to $[1, p_2, 0.7, 0.5]$, and so on.

To conclude, taking into account these results, our proposed method is robust to the transformation parameters, especially in terms of FPR.

Finally, the downward trend in some of the zones (Zone 1, Zone 2, and Zone 3) in Fig. 8 suggests that some parameters have more influence on the FPR value than others. In particular, the last parameter has a greater influence than the rest of parameters (i.e., the higher the p_4 is, the lower the mean FPR becomes). For example, the FPRs of the first 36 points, which correspond to $p_4 = 0.5$, are higher than the FPRs of the last 36 points, which correspond to $p_4 = 1.5$. Note that larger p_4 values make the difference between the magnitude of pseudo-label 3 and pseudo-label 4 higher (the labels used when identifying leaks), and thus, the classifier discriminates better between them, providing fewer false alarms.

3.5.2. Scenario B

The DMAs to be analyzed have been chosen according to the criteria established in Section 3.1.2 and the proposed SSLD method is applied to each of them individually. The results obtained for different threshold values are shown in Fig. 9, where the mean FPR values for each DMA are shown in the left column and the mean TPR values in the right column.

For any threshold obtained with the 80^{th}, 85^{th}, 90^{th} or 95^{th} percentiles, both the SSLD and the ϵ-SVR methods reduce the mean FPR of the MNF method. Note that the higher the percentile is, the lower the FPR of the MNF becomes, and therefore the lower the room for improvement is. As in scenario A, the ϵ-SVR method detects very few leaks, while our method able to detect most of them. In particular, the higher the percentile value is, the higher the mean TPR our method obtains. Consequently, the best threshold would be the one that provides the best detection rate which, in this case, is obtained using the 95^{th} percentile, because our method reduces the FPR of the MNF regardless of the chosen threshold. With this value and considering all the DMAs together, our SSLD method detects 82.35% of the leaks.

It should be noted that the training sets in the DMAs where our method reaches the lowest TPR have very variable flow values. Two examples are shown in Fig. 10. The red vertical lines indicate the reported leakage records, which, depending on the selected threshold, will be considered detectable or not. In

Figure 9: Mean FPR and TPR for each DMA and percentile value.

these datasets, there is no clear pattern of normality. In fact, the training set contains several days that are assumed to be normal but have similar flow values to the leak days. This complicates the learning of the normality, and thus additional information would be needed to differentiate the leakage days from the normal days.

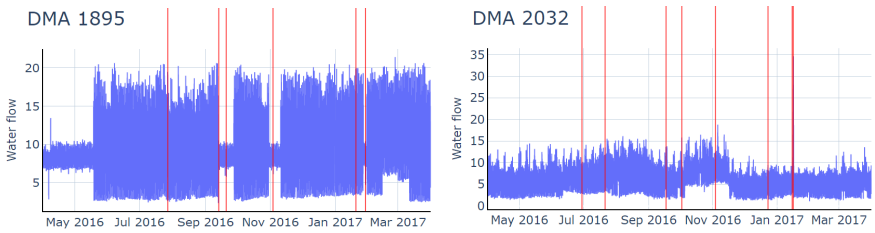


Figure 10: Example of DMAs where the proposed method obtains a low TPR.

Analysis of the transformation parameters

As with scenario A, this section analyzes the influence of the parameters $p_2, p_3, p_4 \in \{0.5, 0.7, 0.9, 1.1, 1.3, 1.5\}$ (recall that $p_1 = 1$) over the FPR and the TPR, for the threshold value obtained with the 95^{th} percentile. This threshold has been chosen for simplicity as it provides the best performance regarding the detected number of leaks. Note that a total of 216 possible combinations are considered. Also, the results in this section are only compared to the MNF method because it provides better performance than the ϵ-SVR method.

Fig. 11 shows the mean FPR and TPR obtained for each DMA with all the different transformation parameters. Taking into account all the DMAs together, the mean FPR value is significantly lower for our method than that obtained by the MNF method for all the parameter combinations, as expected. The reduction in the FPR is better appreciated in the figure on the left in Fig. 11. Regarding the number of detected leaks and taking into account all the DMAs together, most combinations of parameters (80.25%) are able to detect all the leaks.

In conclusion, our proposed method succeeds in significantly reducing the FPR in comparison to the baseline MNF method and also in detecting most of the detectable leaks.

4. Conclusions and Future Work

In this paper, we have proposed a water leak detection method based on the self-supervised classification of nightly flow time series. The classifier we have

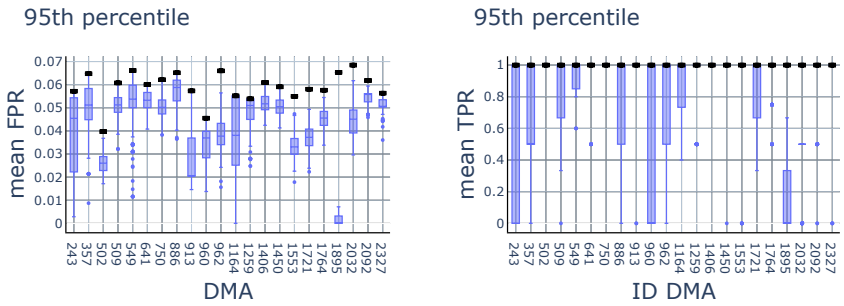


Figure 11: The mean FPR and TPR for different transformation parameter combinations in each DMA. The values obtained with the baseline MNF method are highlighted with a horizontal black line.

chosen (RISE) is specific for time series, which allows considering the temporality of the data, and it is also robust to noise. In particular, our approach builds one model per day of the week, thus it has far fewer models than other methods that require a model for each time step. In addition, the proposed method is entirely data-driven and therefore does not require in-depth knowledge about the dynamics of the series.

The results obtained from the experiments show that, in comparison to other methods in the literature, the proposed SSLD method obtains the best trade-off between detecting the majority of the detectable leaks and providing a low FPR. Specifically, the results have been compared with the MNF, the ϵ-SVR, which, contrary to our method, either detect very few leaks or provide many false alarms.

Several combinations of transformation parameters have also been considered to define the self-labeled dataset that is used as input for the classifier. In all the considered scenarios, our method significantly reduces the FPR of the traditional MNF method, so the SSLD is robust to the choice of transformation parameters in terms of FPR. In particular, the higher the FPR of the MNF, the higher the reduction provided by our method. Although the transformation parameters are more sensitive regarding the TPR in some of the zones, they are in general able to detect most of the detectable leaks.

The main avenue that the results open for future research is related to the type of transformation applied to the data. The self-labeled dataset has been formed using linear transformations, but other types of transformations could also provide appropriate or even better results. An interesting future line of research would be to develop a theory about these transformations, taking into account both the type of available data in the training set and the type of anomalies to be detected. An additional interesting research line would be to test this approach on other types of problems different to leak detection with other types of data and anomalies.

Another promising future work would be to consider the correlation between different zones in the network and to address the self-supervised classification approach from a multivariate perspective with hierarchically structured time series. Although leaks are usually reflected in each zone, and it is generally sufficient to analyze each zone individually, additional information from the network could help to improve the results.

Finally, the proposed method deals with regularly sampled time series of the same length, so future research could also focus on improving this method in situations in which the time series are irregularly sampled or of variable lengths. This will allow series with missing values to be handled, which could appear in this type of problem where sensor failures may occur.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Ane Blázquez-García and Angel Conde are supported by the 3KIA project of the Basque Government under Grant KK-2020/00049 from the ELKARTEK program. Jose A. Lozano and Usue Mori are partially supported by the Basque

Government (IT1244-19) and by the Spanish Ministry of Science, Innovation and Universities (PID2019-104966GB-I00). Additionally, Usue Mori is also supported by the Basque Government under Grant KK-2019/00088 and KK-2020/00049 from the ELKARTEK program and Jose A. Lozano is also partially supported through the BERC 2018-2021 program of the Basque Government and the BCAM Severo Ochoa accreditation SEV-2017-0718 of the Spanish Ministry of Science, Innovation and Universities.

References

[1] E. Farah, I. Shahrour, Leakage Detection Using Smart Water System: Combination of Water Balance and Automated Minimum Night Flow, Water Resources Management 31 (15) (2017) 4821–4833.

[2] S. R. Mounce, A. J. Day, A. S. Wood, A. Khan, P. D. Widdop, J. Machell, A neural network approach to burst detection, Water Science and Technology 45 (4-5) (2002) 237–246.

[3] R. Li, H. Huang, K. Xin, T. Tao, A review of methods for burst/leakage detection and location in water distribution systems, Water Science and Technology: Water Supply 15 (3) (2015) 429–441.

[4] P. S. Murvay, I. Silea, A survey on gas leak detection and localization techniques, Journal of Loss Prevention in the Process Industries 25 (6) (2012) 966–973.

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

[6] S. R. Mounce, A. Khan, A. S. Wood, A. J. Day, P. D. Widdop, J. Machell, Sensor-fusion of hydraulic data for burst detection and location in a treated water distribution system, Information Fusion 4 (3) (2003) 217–229.

[7] S. R. Mounce, J. Machell, Burst detection using hydraulic data from water distribution systems with artificial neural networks, Urban Water Journal 3 (1) (2006) 21–31.

[8] G. Ye, R. A. Fenner, Kalman filtering of hydraulic measurements for burst detection in water distribution systems, Journal of Pipeline Systems Engineering and Practice 2 (1) (2011) 14–22.

[9] M. Zadkarami, M. Shahbazian, K. Salahshoor, Pipeline leakage detection and isolation: An integrated approach of statistical and wavelet feature extraction with multi-layer perceptron neural network (MLPNN), Journal of Loss Prevention in the Process Industries 43 (2016) 479–487.

[10] D. Vries, B. van den Akker, E. Vonk, W. de Jong, J. van Summeren, Application of machine learning techniques to predict anomalies in water supply networks, Water Science and Technology: Water Supply 16 (6) (2016) 1528–1535.

[11] J. Kang, Y. J. Park, J. Lee, S. H. Wang, D. S. Eom, Novel leakage detection by ensemble CNN-SVM and graph-based localization in water distribution systems, IEEE Transactions on Industrial Electronics 65 (5) (2018) 4279–4289.

[12] P. Huang, N. Zhu, D. Hou, J. Chen, Y. Xiao, J. Yu, G. Zhang, H. Zhang, Real-time burst detection in District Metering Areas in water distribution system based on patterns of water demand with supervised learning, Water 10 (12) (2018) 1765.

[13] S. R. Mounce, J. B. Boxall, J. Machell, 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) (2010) 309–318.

[14] S. R. Mounce, R. B. Mounce, J. B. Boxall, Novelty detection for time series

data analysis in water distribution systems using support vector machines, Journal of Hydroinformatics 13 (4) (2011) 672–686.

[15] M. Romano, Z. Kapelan, D. A. Savić, Real-time leak detection in water distribution systems, in: Water Distribution Systems Analysis, 2010, pp. 1074–1082.

[16] M. Romano, Z. Kapelan, D. A. Savić, Automated detection of pipe bursts and other events in water distribution systems, Journal of Water Resources Planning and Management 140 (4) (2014) 457–467.

[17] M. Fagiani, S. Squartini, L. Gabrielli, M. Severini, F. Piazza, A statistical framework for automatic leakage detection in smart water and gas grids, Energies 9 (9) (2016) 665.

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

[19] X. Wang, G. Guo, S. Liu, Y. Wu, X. Xu, K. Smith, Burst Detection in District Metering Areas Using Deep Learning Method, Journal of Water Resources Planning and Management 146 (6) (2020) 1–12.

[20] C. V. Palau, F. J. Arregui, M. Carlos, Burst detection in water networks using principal component analysis, Journal of Water Resources Planning and Management 138 (1) (2012) 47–54.

[21] Y. Wu, S. Liu, X. Wu, Y. Liu, Y. Guan, Burst detection in district metering areas using a data driven clustering algorithm, Water Research 100 (2016) 28–37.

[22] K. Aksela, M. Aksela, R. Vahala, Leakage detection in a real distribution network using a SOM, Urban Water Journal 6 (4) (2009) 279–289.

[23] S. Patabendige, R. Cardell-Oliver, R. Wang, W. Liu, Detection and interpretation of anomalous water use for non-residential customers, Environmental Modelling and Software 100 (2018) 291–301.

[24] Y. Wu, S. Liu, Burst Detection by Analyzing Shape Similarity of Time Series Subsequences in District Metering Areas, Journal of Water Resources Planning and Management 146 (1) (2020) 1–12.

[25] A. Carreño, I. Inza, J. A. Lozano, Analyzing rare event, anomaly, novelty and outlier detection terms under the supervised classification framework, Artificial Intelligence Review (2019) 1–20.

[26] I. Golan, R. El-Yaniv, Deep anomaly detection using geometric transformations, in: Advances in Neural Information Processing Systems, Vol. 31, 2018, pp. 9758–9769.

[27] S. Gidaris, P. Singh, N. Komodakis, Unsupervised representation learning by predicting image rotations, in: International Conference on Learning Representations, 2018, pp. 1–16.

[28] S. Wang, Y. Zeng, X. Liu, E. Zhu, J. Yin, C. Xu, M. Kloft, Effective End-to-end Unsupervised Outlier Detection via Inlier Priority of Discriminative Network, in: Advances in Neural Information Processing Systems (NeurIPS), 2019, pp. 5962–5975.

[29] J. Lines, S. Taylor, A. Bagnall, Time series classification with HIVE-COTE: The hierarchical vote collective of transformation-based ensembles, ACM Transactions on Knowledge Discovery from Data 12 (5) (2018) 1–35.

[30] A. Bagnall, M. Flynn, J. Large, J. Lines, M. Middlehurst, A tale of two toolkits, report the third: On the usage and performance of HIVE-COTE v1.0 (2020). arXiv:2004.06069v2