Safety Science 130 (2020) 104874

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

# Time windows: The key to improving the early detection of fuel leaks in petrol stations

Silvia Alayón*, Marta Sigut, Rafael Arnay, Pedro Toledo

*Department of Computer Science and Systems Engineering, University of La Laguna, Avda. Astrofísico Francisco Sánchez, s/n. Edif. Física y Matemáticas, La Laguna, 38200 Santa Cruz de Tenerife, Spain*

ARTICLE INFO

*Keywords:*
Machine learning
Two-class classifiers
Time windows
Fuel leaks
Inventory reconciliation

ABSTRACT

In this paper, the authors propose the use of time windows to improve the detection of fuel leaks in petrol stations. They employ two-class supervised classifiers that work with feature sets containing representative variables taken from station inventory books that indicate the presence of leaks. Fuel leaks in petrol stations with underground tanks pose a serious problem from an environmental standpoint. Large leaks are very evident, and are therefore detected quickly without the need to use a specific procedure. Small leaks, however, tend to go unnoticed, and if no detection techniques are employed, they are only identified once environmental damage has been done. This makes detecting the leak in the shortest time possible as important as ascertaining when the leak started. The authors show how the use of time windows, which entails having the classifier work with information accumulated over several days, can be used to efficiently resolve the proposed problem, fully complying with the applicable regulation.

## 1. Introduction

There can be no doubt of the concern over protecting the environment and over the consequences that ignoring this question could entail. Proof of this is the contributions by Chofreh et al. (2014), Collin and Melloul (2003), Erkman (1997), Khan and Husain (2003), Maxwell and van der Vorst (2003) and Robert (2000). Specifically, the importance of detecting fuel leaks at petrol stations early is not in question, as reflected in (United States Environmental Protection Agency Pacific Southwest/Region 9, 2003). The main problem with this type of facility, in which the fuel tanks are buried in the ground, is that the leak is often detected only after irreversible environmental damage has been caused. This explains why such an effort had been made over several years to minimise the negative impact of underground fuel storage systems, as explained in (Gurr and Homann, 1996). In this regard, having an automatic system that is able to detect leaks as quickly as possible offers a significant advance in the area of environmental safety.

Many of the systems currently in use for detecting fuel leaks rely on the so-called 'inventory reconciliation' (Health and Safety Authority, 2013), as in Gorawska and Pasterak (2016), Gorawski et al. (2015a) and Li et al. (2011). Inventory logs, also called tank logs, provide a record of the amounts of fuel sold, purchased and stored in each tank. In particular, statistical inventory reconciliation (SIR) is a technique that allows detecting fuel leaks in a way that is more sensitive and accurate than that offered by classical inventory controls, as explained in (Guide EPA, 1995). By using one-month data for tanks, and by employing various statistical tools, an SIR system can identify leaks that are normally not detected by simple inventory checks.

There are also various products on the market for detecting fuel leaks in petrol stations that are based on SIR. Of note among these is the one marketed by the English company Fairbanks, which is certified in several countries, including Spain. Petrol stations are required to send to Fairbanks their tank logs, that is, their daily sales records, so it can process the information and detect any potential leaks. Some companies are hesitant to provide their sales data. This, along with the fact that the Fairbanks product is completely closed, with no details on the technique used to detect leaks, has compelled the authors of this paper to consider this problem from another perspective, one based on Machine Learning.

In Sigut et al. (2014), the authors offer an initial approach to solving the problem at hand. In that paper, the authors present a series of two-class classifiers and a series of feature sets that adequately represent the objects to be classified. These objects are the days during which the petrol station is operational, which are classified as 'day without leak' and 'day with leak'. For the leak detection to work properly, the system implemented must be able to distinguish between the differences in the inventory log due to leaks and those due to other factors, such as evaporation, volume change due to temperature, and factors like the

* Corresponding author.
*E-mail addresses:* salayon@ull.edu.es (S. Alayón), marsigut@ull.edu.es (M. Sigut), rarnayde@ull.edu.es (R. Arnay), petode@ull.edu.es (P. Toledo).

https://doi.org/10.1016/j.ssci.2020.104874
Received 14 February 2020; Received in revised form 1 June 2020; Accepted 7 June 2020

**Table 1**
Extract of a service station's daily inventory records. The different rows contain the quantities (in litres) measured every day.

| Tank | Date | Initial (litres) | (+) Receipts (litres) | (+/-) Adjustments (litres) | (-) Sales (litres) | Theoretical inventory available (litres) | Current available (litres) |
|---|---|---|---|---|---|---|---|
| e + Diesel | | | | | | | |
| ZT01 | 01.06.2016 | 15219.49 | 5998.00 | 0 | −10846.96 | 10370.53 | 10389.83 |
| | 02.06.2016 | 10389.83 | 9997.00 | 0 | −10997.71 | 9389.12 | 9409.51 |
| | 03.06.2016 | 9409.51 | 9999.00 | 0 | −10637.28 | 8771.23 | 8807.70 |
| | 04.06.2016 | 8807.70 | 12000.00 | 0 | −7694.53 | 13113.17 | 13141.52 |
| | 05.06.2016 | 13141.52 | 0 | 0 | −5434.20 | 7707.32 | 7725.19 |
| | 06.06.2016 | 7725.19 | 15005.00 | 0 | −11555.22 | 11174.97 | 11192.25 |
| | 07.06.2016 | 11192.25 | 9998.00 | 0 | −10743.08 | 10447.17 | 10465.53 |
| | 08.06.2016 | 10465.53 | 10000.00 | 0 | −9323.05 | 11142.48 | 11154.40 |
| | 09.06.2016 | 11154.40 | 15002.00 | 0 | −11170.37 | 14986.03 | 15007.53 |

ones indicated in Gorawski et al. (2015a); Gorawski et al. (2015b). This is why in Sigut et al. (2014), the authors opted for a pattern recognition system in which the two-class classifier can learn to differentiate between these two situations. That paper provided satisfactory classification results both for constant and variable leaks for certain combinations of classifiers and feature sets. In order to decide when a classification result was satisfactory, the authors relied on the European EN 13160 family of standards (UNE-EN 13160-5, 2005), which specifies a maximum time for detecting the leak with an error below a certain threshold.

The best results in (Sigut et al., 2014) were obtained with the groups of characteristics involving variables in which the effect of the leak builds up over time. Specifically, the combination of the 'variation/real volume' characteristic with certain classifiers yields adequate results that are fully compliant with the rule contained in the EN 13160 standard. That paper does not make use of the time factor; rather, data associated with individual days are analysed. The problem with this is that when a leak of X litres is identified when analysing the data for a specific day, it is impossible to know whether it is in fact an X-litre leak that occurred on that day or a smaller leak (X/N litres per day) that has been accumulating over the N previous days. In order to evaluate the gravity of the problem at hand and be able to determine the urgency of the action required, it is essential to be able to differentiate between these two situations.

To achieve this, in this paper we resort to the use of time windows. The use of variable length time windows (whose size is the number of days during which the leak is simulated) not only solves this problem, it will also reveal how many days must elapse from the time when the leak starts until the system is able to detect it. Already in Sigut et al. (2014) the authors considered including the trend over time of the system's variables as one way to improve the classification results.

To the best of our knowledge, the inclusion of time information in leak detection is quite novel. A paper was published Gorawski et al. (2017) in which time-series data are used to identify trends that allow for the early detection of fuel leaks in underground storage tanks. In that paper, the authors use one-month data batches, meaning they start from zero each month and accumulate the data associated with the subsequent days. This makes it very difficult to detect a leak in the first few days of the month. What we propose in this paper, however, is the use of time windows that slide over the total number of days available in the inventory logs. By sliding the window, we can account for the same cumulative effect regardless of the specific day in question.

In what follows, we describe the classification methods and the sets of variables used in this paper, including the construction of the time windows. We conclude by considering the validity of our results in terms of their compliance with the applicable regulation.

## 2. Material and methods

This work relies on data taken from petrol station inventories, the goal being to identify those days on which fuel leaks occurred. This section explains the inventory reconciliation method and the nature of the actual data that are available.

The leak detection problem is still regarded as a classification problem; as a result, we will also present the new features studied and the basic tenets of the new classifiers employed.

### 2.1. The inventory reconciliation method. Real data available and simulated data

Nowadays, inventory logs are used in every petrol station. These logs contain a record of the fuel available in each tank, as well as of petrol sales and purchases.

The basic concepts in the inventory reconciliation method include the *theoretical inventory* (quantity of fuel that is expected to be inside the tank at the end of the day or shift), *difference in inventory* or *variation* (difference between the measured and expected amount of fuel in the tank) and *variance with respect to sales.* An extract of the daily inventory records for a service station is shown in Table 1.

The main goal of the inventory reconciliation method is to detect the abnormal operation of a service station by comparing the theoretical and actual inventories. It is important to keep in mind, however, that a lack of correlation between the theoretical and actual inventories at the end of the day does not necessarily indicate a leak or excess fuel. These discrepancies could be due to a combination of various factors that are difficult to account for, such as a change in the fuel volume due to temperature, the evaporation of fuel, reading errors in the tank's sensors and so on. All of this makes applying the inventory reconciliation method a fairly complex task.

The data used for the experiments in this paper are actual data provided by Spanish gas company Repsol. These data differ from those used in Sigut et al. (2014), as they involve more modern service stations. These inventories contain the data from the station when it is operating without leaks (see Table 1). When no data from abnormal operations (i.e. when there is a leak in the station) are available, said data were simulated by altering the original data, in keeping with the recommendations in the current applicable standard (UNE EN13160-5, 2017):

- Simulation of constant leaks: the standard specifies that in this type of leak, the loss of a constant amount of fuel every day from the tank is simulated. When simulating the constant leak, the figure that represents the volume $v_i$ of the ith reading in the tank log is replaced by $v_i^{'}$, as shown in Eq. (1):

$$v_i^{'} = v_i \bullet \sum_{j=1}^{i} (t_j - t_{j-1})R \tag{1}$$

where $R$ is the simulated leak rate, $t_j$ is the time instant for the jth reading, and the summation represents the number of days elapsed since the start of the leak.

**Table 2**
Actual data, no leaks, for a ten-day sequence taken from an inventory log.

| Initial without leak (litres) | Theoretical inventory without leak (litres) | Real inventory without leak (litres) | Variation (litres) |
|---|---|---|---|
| 20272.46 | 15770.27 | 15764.53 | −5.75 |
| 15764.53 | 15413.11 | 15427.66 | 14.56 |
| 15427.66 | 16347.61 | 16366.34 | 18.73 |
| 16366.34 | 21296.34 | 21354.97 | 58.63 |
| 21354.97 | 18120.29 | 18096.09 | −24.21 |
| 18096.09 | 12893.56 | 12884.14 | −9.42 |
| 12884.14 | 15580.34 | 15601.77 | 21.43 |
| 15601.77 | 15737.26 | 15749.39 | 12.13 |
| 15749.39 | 15534.95 | 15556.35 | 21.41 |
| 15556.35 | 16053.68 | 16067.33 | 13.64 |

- Simulation of variable leaks: the standard specifies that a variable leak is one that falls when the amount of fuel stored in the tank drops. Therefore, in this case the simulated leaks depend on the volume of fuel stored in the tank. Specifically, the leak rate associated with the jth reading is obtained by using Eq. (2):

$$r_j = \frac{n\sqrt{v_j}}{\sum_{k=1}^{n}\sqrt{v_k}}R \tag{2}$$

where $n$ indicates the number of readings, which matches the size of the time window, $v_j$ is the volume of fuel stored in the tank on the jth day and $R$ is the nominal simulated leak rate.

When estimating the variable leak, the figure representing the volume of the ith reading in the tank log is replaced by $v_i^{'}$, as shown in Eq. (3):

$$v_i^{'} = v_i \bullet \sum_{j=1}^{i} (t_j - t_{j-1})r_j \tag{3}$$

where the sum represents the same thing as in Eq. (1), while the flow rate is obtained from Eq. (2).

Starting with an actual 10-day data series, like the one given in Table 2, shows the result of simulating a constant leak (Table 3) and a variable leak (Table 4). In the first case, the constant leak, notice how the values in the "Real inventory with leak" column (Table 3) match the original values (same column, Table 2) minus a constant value that is subtracted each day and is equal to the leak rate. In the second case, the variable leak, notice the differences between the values in the third column in Tables 2 and 4, which correspond to a variable leak rate that is dependent on the tank volume.

### 2.2. Detection of fuel leaks: A classification problem

In our previous paper (Sigut et al., 2014), we considered the problem of detecting leaks using inventory reconciliation analysis as a classification problem. In this approach, the goal is to use the inventory data to classify the station's days of operation into two possible classes: "day with leak" and "day without leak". In order to be able to train the classifiers, we had to simulate data to represent the "day with leak" class.

**Table 3**
Data from a simulated constant 50 L/day leak rate (see Eq. (1)) based on the original data shown in Table 2.

| Initial with leak (litres) | Theoretical inventory with leak (litres) | Real inventory with leak (litres) | Variation with leak (litres) |
|---|---|---|---|
| 20272.46 | 15770.27 | 15714.53 | −55.74 |
| 15714.53 | 15363.11 | 15327.66 | −35.45 |
| 15327.66 | 16247.61 | 16216.34 | −31.27 |
| 16216.34 | 21146.34 | 21154.97 | 8.63 |
| 21154.87 | 17920.29 | 17846.09 | −74.2 |
| 17846.09 | 12643.56 | 12584.14 | −59.42 |
| 12584.14 | 15280.34 | 15251.77 | −28.57 |
| 15251.77 | 15387.26 | 15349.39 | −37.87 |
| 15349.39 | 15134.95 | 15106.35 | −28.6 |
| 15106.35 | 15603.68 | 15567.33 | −36.35 |

**Table 4**
Data from a simulated variable leak with a nominal 50 L/day leak rate (see Eqs. (2) and (3)) based on the original data shown in Table 2.

| Initial with leak (litres) | Theoretical inventory with leak (litres) | Real inventory with leak (litres) | Variation with leak (litres) |
|---|---|---|---|
| 20272.46 | 15770.27 | 15709.31 | −60.95 |
| 15582.07 | 15357.89 | 15323.75 | −34.13 |
| 15104.96 | 16243.70 | 16214.27 | −29.43 |
| 15907.70 | 21144.27 | 21153.29 | −9.02 |
| 20753.16 | 17918.61 | 17837.74 | −80.86 |
| 17307.50 | 12635.21 | 12573.62 | −61.58 |
| 11939.78 | 15269.82 | 15247.23 | –22.58 |
| 14549.96 | 15382.72 | 15346.42 | −36.30 |
| 14566.63 | 15131.98 | 15104.71 | −27.26 |
| 14242.49 | 15602.04 | 15567.33 | −34.71 |

The selection of features is a very important step in the classification process, as it can increase the classifier's efficiency (Duda et al., 2001). In Sigut et al. (2014), the actual data taken from the inventories were used to build these features, and several alternatives were tested when designing them. The classifier's internal structure is also decisive. As a result, we tested several classification methods by applying them to different sets of features. The results showed that certain combinations of classifiers and types of features can be useful and applicable to this problem.

In this paper, we again apply classification techniques to the leak detection problem, but we attempt to improve the classifier system by varying two main aspects:

1. Design of new features based on time windows that are able to collect information on the time trend of the variables. With these new features, we hope to not only train a classifier that is able to detect the smallest possible leaks, but also to determine how many days it takes for the classifier to detect them.
2. Study of new classifier systems, in addition to the two that are considered in Sigut et al. (2014), in order to determine which classifier learns the time trend in the data better during the training.

These aspects are explained in greater detail in the sections that follow.

### 2.3. New features for the classification problem in question: The use of time windows

The goal of this paper is to analyse if the inventory data for the service stations contain enough information to construct useful and representative features that allow a classifier to learn to identify days of normal operations at the station ("no leak") and days when a leak had occurred ("leak"). In Sigut et al. (2014), the most appropriate features were unknown, hence the reason why different groups of features were tested: variation (theoretical inventory - real inventory), sales variation (variation/sales), theoretical volume (real volume from the previous day + receipts - sales + adjustments), variation/real volume, daily sales and variation, daily sales and sales variation. These features were designed following the advice of an expert in the domain.

In the experiments conducted, we built these feature vectors by including only the values associated with each day independently. A novelty in this paper is the inclusion of the time trend in the station's operation as an intrinsic feature. Thus, each feature vector will include the relevant variable being studied, but not for just one day, but rather for several consecutive days. The number of days considered will be determined by the size of the time window used.

The features considered in this approach are as follows:

- Feature set 1 - FS1 (one feature): *variation* = theoretical inventory - real inventory.
- Feature set 2 – FS2 (one feature): *cumulative variation* = sum of the variation values for each day in the time window.
- Feature set 3 – FS3 (one feature): *variation over sales* = variation/sales.
- Feature set 4 – FS4 (one feature): *cumulative variation over sales* = sum of the variation values over the sales in each day of the time window.
- Feature set 5 – FS5 (one feature): *variation over real volume* = variation/real volume.
- Feature set 6 – FS6 (one feature): *cumulative variation over real volume* = sum of the variation/real volume values for each day in the time window.
- Feature set 7 – FS7 (two features): *variation over sales* and *variation over real volume*.
- Feature set 8 – FS8 (two features): *variation* and *sales*.

For example, FS1 consists of one feature, the variation, which is defined as the difference between the theoretical and real inventories. Fig. 1 shows the values for these two variables taken from the general inventory provided in Table 1. The figure also shows the value of the variation variable for each day considered. Suppose we use a four-day time window. FS1 will consist of the variation values for each day, while FS2, which is the cumulative variation, will feature a single value that is the sum of those four variation values since the start of the window. The remaining feature vectors are similarly constructed.

### 2.4. Classification algorithms used in this work

For this paper, we tested six supervised classifiers using the feature sets mentioned in the previous section. For this purpose, the software tool MatLab was used. In a supervised classification, the classes are predetermined. A subset of the available data is labelled in some of these classes to comprise the so-called “training set”. In the training stage, each classifier has to learn the underlying patterns in the training set data and adapt its internal structure to distinguish data of different classes in the most efficient way possible. When the training is complete, the resulting classifiers must be evaluated against a new data set, this one labelled the “test set” (Duda et al., 2001).

The following methods are used in this paper: Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), Naive Bayes Classifier (NBC), Decision Tree Classification (DTC), Support Vector Machine (SVM) and K-Nearest Neighbour algorithm (K-NN). The LDA and KNN methods were used in the previous paper (Sigut et al., 2014), while the other methods were tested for the first time in this work.

They consist of the following:

- LDA maximises the ratio of between-class variance to the within-class variance in any particular data set, thereby guaranteeing maximum separability. This method finds a linear combination of features that characterises or separates two or more classes of objects (Duda et al., 2001).
- QDA separates measurements of two or more classes of objects by a quadric surface. It is a more general version of the linear classifier (Jain et al., 2000).
- NBC is a simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions between the features. The use of Bayes' theorem in the classifier's decision rule determines the probability of an object belonging to a class, based on prior knowledge of conditions that might be related to the class (Niculescu-Mizil and Caruana, 2005). In this paper, we use two variants of this algorithm: the first assumes that the probability density function follows a normal distribution (NBC), and the second applies a box kernel to smooth this probability density function (NBC-kb):

Naive Bayes Gaussian Kernel: $f(x) = \frac{1}{\sqrt{2\pi}}\exp(-0.5x^2)$
Naive Bayes Box (Uniform) Kernel: $f(x) = 0.5I\{|x| \leq 1\}$

- DTC is a predictive model that goes from observations about an object (represented in the branches) to conclusions about the object's class (represented in the leaves). In these tree structures, leaves represent class labels and branches represent conjunctions of features that lead to those class labels (Pang-Ning and Steinbach, 2013).
- SVM considers each data object as a point in n-dimensional space (where n is the number of features), with the value of each feature being the value of a particular coordinate. The classification is then carried out by finding the hyper-plane that best separates the two classes in the space (Hsu and Lin, 2002). Two different kernels have been used in the experiments: linear (SVM-l) and Gaussian (SVM-g).

SVM Gaussian Kernel: $G(x_j, x_k) = \exp(-\|x_j - x_k\|^2)$
SVM linear Kernel: $G(x_j, x_k) = x_j^{'} x_k$

- K-NN is a well-known method for classifying objects based on the closest training examples in the feature space, as explained in (Cover and Hart, 1967). An object is classified by a majority vote of its neighbours, with the object being assigned to the most common class amongst its K nearest neighbours (K is a positive integer, typically small). Test were performed that varied the number of neighbours from 1 to 70. The best results were obtained with 25 neighbours.

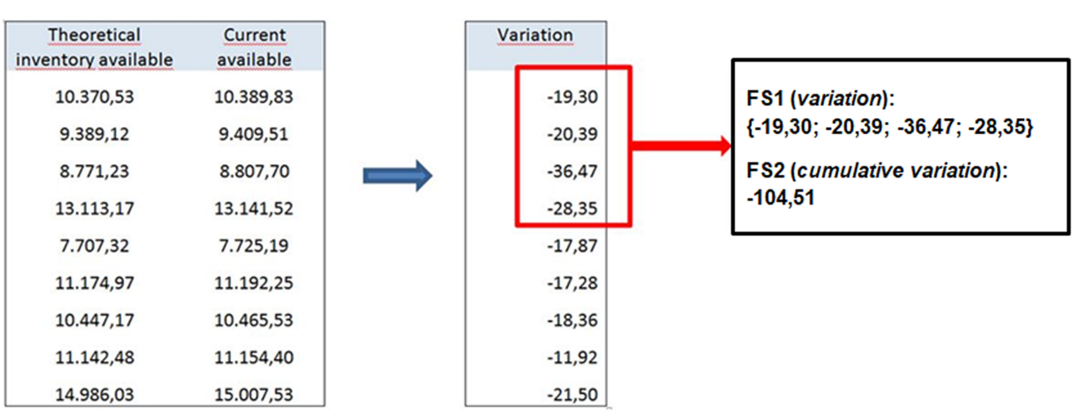


**Fig. 1.** Example of building features using time windows.

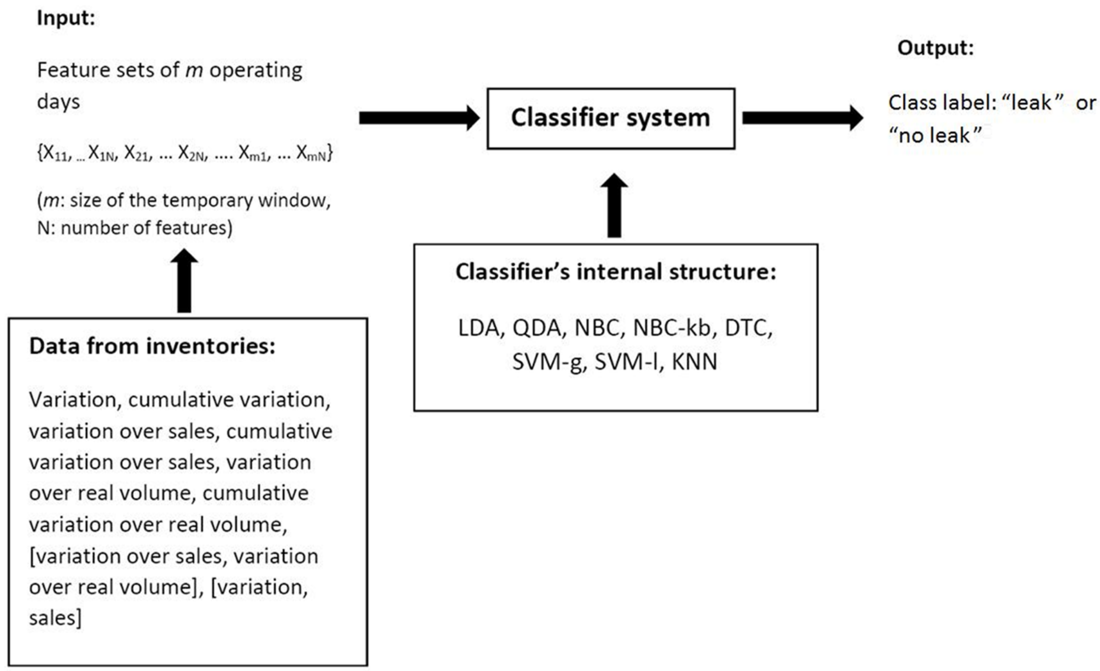


**Fig. 2.** The detection of petrol leaks considered as a classification problem.

In Fig. 2, the detection of petrol leaks is represented as a classification process. The inputs are the feature vectors representing several days of operation of the petrol station, where the number of days considered depends on the size of the time window used. These features are composed using data extracted from the inventories. The classifier system is one of the six classification procedures used in this research work, and the output is a class label that indicates if a leak occurred in the operational time period being studied or not.

## 3. Calculation

The goal of the experiments is to train classifiers that can distinguish between days with and without leaks using a technique based on time windows. This technique involves using time samples spanning at least one day to train and test the classifiers in order to provide them with more information so they can learn patterns and trends that improve the classification results.

The classifiers are trained using samples from two categories: samples with a leak and samples without a leak. The samples without a leak are obtained from the original data, selecting sequences from consecutive days for a given time window size. As explained in Section 2.1, samples without leaks are used to simulate samples with both constant and variable leak rates (see Eq. (1), and Eqs. (2) and (3), respectively). In the samples with leak, the leak is simulated from the first day in the time window and for the entire duration of the time window.

Two experiment types were conducted. On the one hand, we carried out general experiments to determine how the classifiers and features tested behave for a given range of time window sizes and leak rates. And on the other, we conducted experiments intended to determine if the classification systems tested comply with the requirements imposed by the applicable standard (UNE EN13160-5, 2017).

In the first experiment type, we trained classifiers by using time windows ranging in duration from 1 to 10 days, and we simulated both constant and variable leaks based on nominal leak rates of 5 L, 10 L, 20 L, 30 L, 50 L, 70 L and 90 L (see Eq. (1), Eq. (2)and Eq. (3)).

In the second experiment type, we used window sizes ranging from 1 to 14 days, and we simulated both constant and variable leaks based on a nominal leak rate of 19.2 L. The regulation specifies that a leak of 19.2 L/day must be detectable by day 14, with a detection error that does not exceed 5%. Note that the applicable standard (UNE EN13160-5, 2017) no longer has a requirement to detect 4 L per hour in 24 h and 2 L per hour in 7 days, which were included in the previous standard (UNE-EN 13160-5, 2005). We used this type of experiment to determine which classifier-feature combinations satisfy the standard, how many days the leak is in progress before it is detected (size of the time window) and the error rate.

We used 175 samples from the "no leak" class and 175 samples from the "leak" class. To train the classifiers and obtain a representative testing error, we used a 10-fold cross validation (Arlot, 2010). This process involves dividing the sample set into ten similarly sized sets in order to iteratively select one set for testing and use the samples in the other nine sets to train the classifier. This process is repeated ten times, with a different testing set selected each time. Finally, the classification error in all the iterations is averaged. Based on this approach, we used training sets for the cross validation with 315 samples and test sets with 35 samples.

## 4. Results and discussion

In this section, we present the main results obtained in the experiments described in the previous section, as well as the main conclusions drawn from said experiments. The detection error shown in the results is the total classification error, that is, the one that takes into account the total number of misclassified samples, both of samples with leak (false positives) and samples without leak (false negatives).

As concerns the general experiments, Section 4.1 presents the results that highlight the influence of the nominal leak rate, the size of the time window and the cumulative variables on the performance of the detection system. Section 4.2 includes the results of those experiments whose goal is to select the best classifier-feature combination.

Finally, Section 4.3 presents the results of the experiments carried out to determine if the classifiers and features selected satisfy the applicable regulation.

### *4.1. General experiments. Influence of nominal leak rate, time window size and cumulative variables*

A large number of experiments were conducted in this work. We

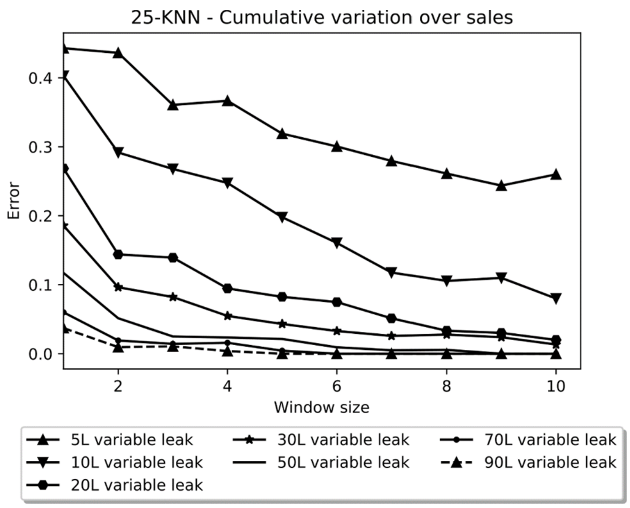


**Fig. 3.** Classification error for the combination of classifier KNN and the "cumulative variation over sales" variable for different constant leak rates.

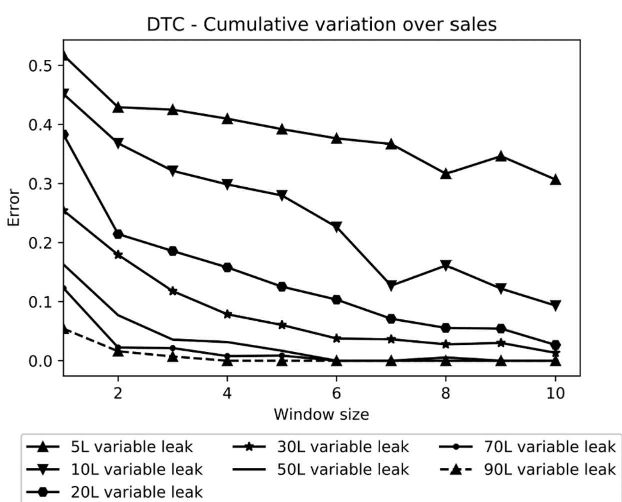


**Fig. 4.** Classification error for the combination of the DTC classifier and the "cumulative variation over sales" variable for different variable leak rates.

used eight different classifiers with eight different features, combining time windows of different lengths and analysing simulated leak data with various constant and variable leak rates. We implemented every combination possible and obtained dozens of results. It is impossible to show them all in the paper due to space constraints. We found three significant trends in these experiments involving the influence of the nominal leak rate, the influence of the time window size and the effect of using cumulative variables. We have selected some examples that are representative of these trends, as shown graphically in Figs. 3–7. However, it is important to note that the same trends that are shown in these figures are also present in the set of the various classifiers, features and leak rates considered.

The experiments conducted to determine the influence of the leak rate on the performance of the classification system show that, as expected, the higher the leak volume, the lower the classification error. The samples associated with the higher volume leaks are easy to differentiate from the samples for the days without leaks, and thus the classifiers yield better results. Figs. 3 and 4 show various examples of how, for the same combination of classifier and feature, the classification error falls as the leak rate increases.

We also noticed that the longer the window size, the lower the classification error. This result is as expected, since the classifiers have more information with which to detect the leaks. Figs. 5 and 6 show how, for different classifiers and variables, the error drops considerably when 10-day windows are considered, in comparison to using samples from a single day.

In order to determine the effect that one variable's accumulation has on the leak detection, we compared the results obtained with a given variable and with that same variable but using its cumulative value. In these cases, we identified a general trend toward lower classification errors using cumulative variables. This result is also to be expected, since, a priori, the most discriminatory property for classifying leaks in time windows is the accumulation of this leak. Summing the variables before training the classifiers facilitates the classification process.

Fig. 7 shows a comparison of the errors made by a classifier using a variable (left) and its cumulative version (right). Although using cumulative variables, as shown in Fig. 7, generally improves the classification results, the error reduction depends on the type of classifier. This may be because using cumulative variables is offset by the fact that the pattern to be classified always has size 1. As the window size increases, if cumulative variables are not used, the patterns have more information on how the variable evolves, which also aids in the classification process.

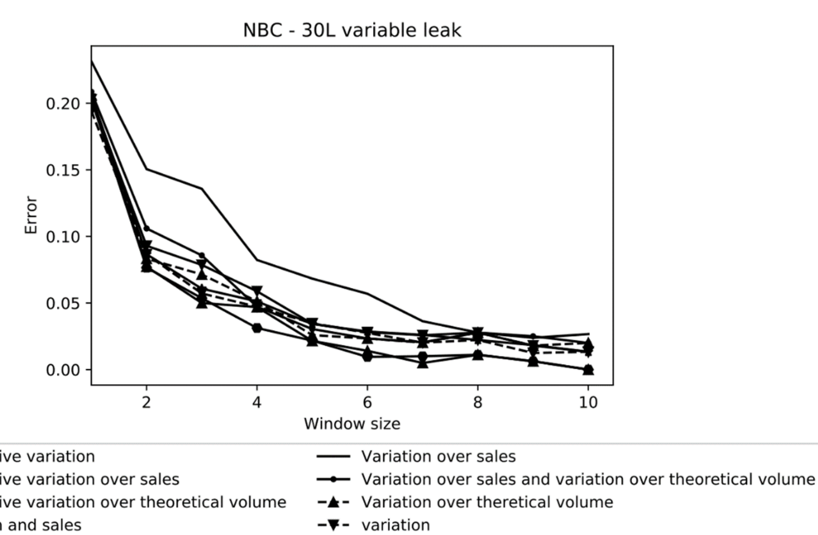


**Fig. 5.** Classification error based on window size for the NBC classifier, with the different features considered for a variable leak.

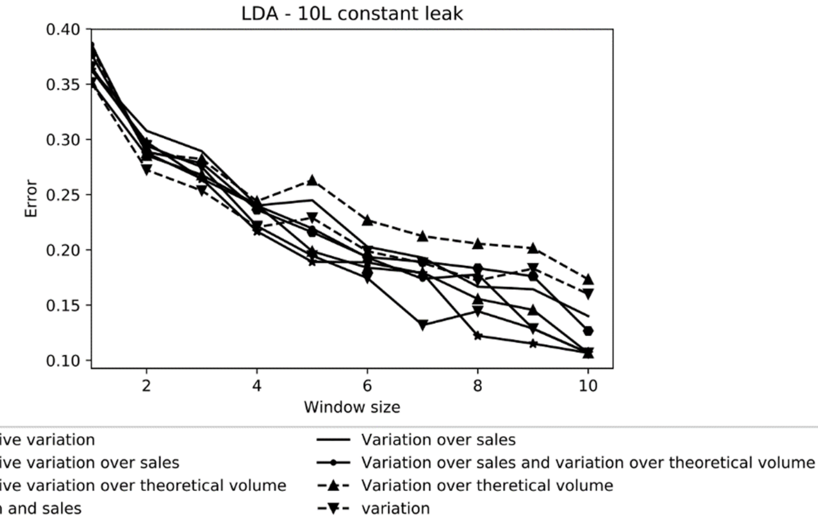


**Fig. 6.** Classification error based on window size for the LDA classifier, with the different features considered for a constant leak.

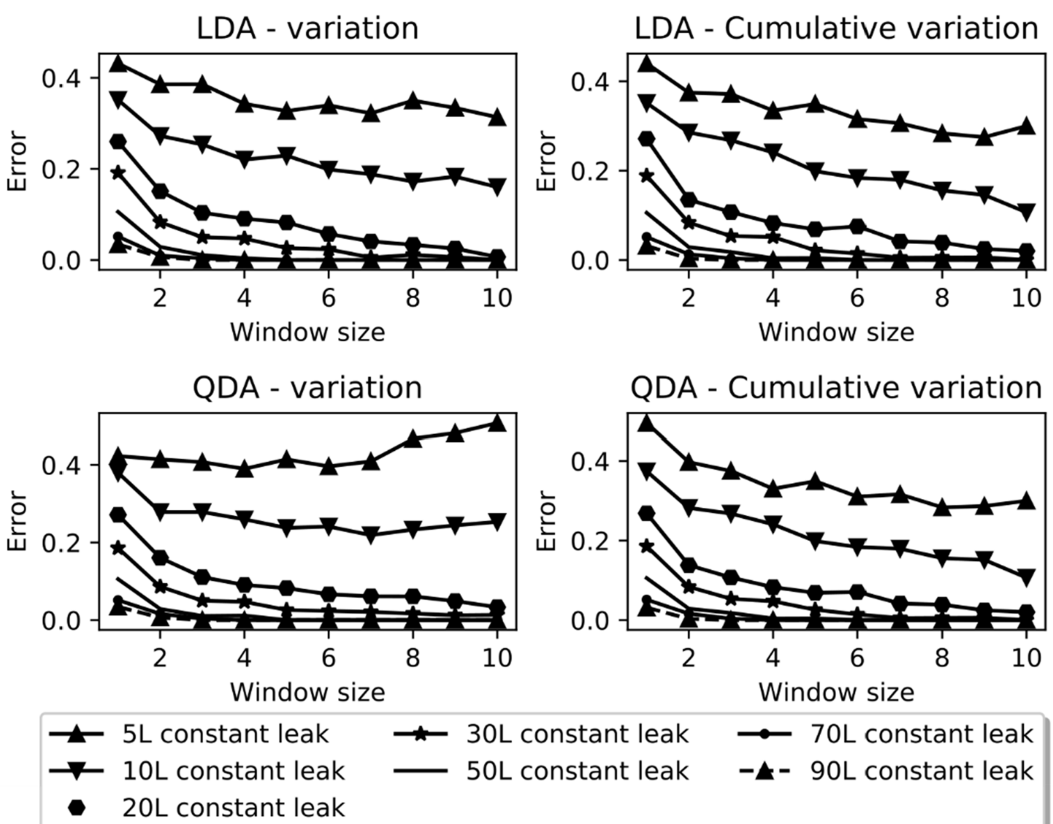


**Fig. 7.** Classification errors based on window size for the LDA and QDA classifiers, for different constant leak rates, for the "variation" feature over the days in the window (left), and for its cumulative version (right).

*4.2. General experiments. Selecting the best classifier and the best feature*

In order to determine the classifier and variable that yield the best results, we conducted an analysis in different stages. First, we attempted to determine if any of the variables exhibit behaviour that is significantly better than the rest. To this end, we proceeded to calculate four indicators.

- Acc Error: Calculated as the sum of the classification errors using a given feature. Obtained for each possible window size, classifier and leak type. Allows comparing the general behaviour of the feature for any value of the remaining analysis parameters.
- Min-Max Error: Obtained by first calculating the maximum error obtained for all window sizes for a leak type and classifier. Then, the minimum error from amongst all these maximum errors is determined. This indicator allows comparing the best behaviour in the features for the worst-case window size selected. It is primarily intended to determine the features that behave best when little time is available to obtain an answer.
- Avg Error: Average of the aggregated classification errors for all window sizes made and leak type for this feature.
- Std Error: The standard deviation of the calculation done for the

**Table 5**
Values of the 'Acc Error', 'Min-Max Error', 'Avg Error' and 'Std Error' for the different feature sets (FS1: variation, FS2: cumulative variation, FS3: variation over sales, FS4: cumulative variation over sales, FS5: variation over real volume, FS6: cumulative variation over real volume, FS7: variation over sales and variation over real volume, FS8: variation and sales).

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 159.0009 | 0.0314 | 1.4196 | 1.7198 |
| FS2 | 118.1759 | 0.0314 | 1.0551 | 1.2233 |
| FS3 | 176.4209 | 0.0285 | 1.5751 | 1.6912 |
| FS4 | 122.4089 | 0.0285 | 1.0929 | 1.1527 |
| FS5 | 166.0588 | 0.0400 | 1.4826 | 1.7366 |
| FS6 | 124.4746 | 0.0371 | 1.1113 | 1.3051 |
| FS7 | 182.5245 | 0.0342 | 1.6296 | 1.8030 |
| FS8 | 260.2795 | 0.0342 | 2.3239 | 2.4982 |

above indicator. Together, these last two indicators measure the average behaviour when using the feature analysed.

The results of these indicators are shown in Table 5.

The best behaviour was observed for the "Cumulative Variation" feature, which very clearly exhibits the best behaviour for "Acc Error" and shows very good behaviour for the remaining indicators.

During the second phase of the analysis, we measured the behaviour of the classifiers by using reasoning analogous to that presented above for the features. We again built four indicators, whose definitions are consistent with those provided for comparing the different features. Table 6 shows a comparison of the results for different classifiers.

The LDA classifier behaves better than the remaining classifiers. Its "Acc Error" is the lowest, and the remaining indicators also confirm its better overall behaviour.

The analysis presented so far poses the multidimensional problem of selecting the features and classifier to build the final system, analysing each of the parameters separately. This entails an initial approach for the final solution that presupposes that the parameters of the analysis are independent, and that the goal might thus be to minimise the error produced in each parameter individually. However, if this premise were not satisfied, it would be possible to obtain other specific combinations of features and classifiers that would improve the system's performance, even if said features or classifiers are not the best when analysed independently.

In order to rule out the possibility of this phenomenon occurring, we also conducted a joint analysis of the feature and classifier dimensions, which is shown in detail in Appendix A. Our results show that the best classifier-feature combination corresponds to the SVM-linear classifier with the "cumulative variation" feature, which yields the following values: Acc Error = 13.9735, Min-Max Error = 0.0314, Avg Error = 0.9981 and Std Error = 1.1682. The second-best classifier-feature combination corresponds to the LDA classifier with the same feature ("cumulative variation"), with the following values: Acc Error = 14.0534, Min-Max Error = 0.0314, Avg Error = 1.0038 and Std Error = 1.1751.

**Table 6**
Values of the 'Acc Error', 'Min-Max Error', 'Avg Error' and 'Std Error' for the different classifiers.

| Classifier | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| SVM-g | 207.0043 | 0.0314 | 0.1848 | 0.2284 |
| SVM-l | 119.1311 | 0.0314 | 0.1063 | 0.1240 |
| KNN | 199.4223 | 0.0314 | 0.1780 | 0.2117 |
| QDA | 147.1472 | 0.0314 | 0.1313 | 0.1604 |
| DTC | 189.9909 | 0.0400 | 0.1696 | 0.1685 |
| LDA | 118.7585 | 0.0285 | 0.1060 | 0.1243 |
| NBC | 147.3272 | 0.0314 | 0.1315 | 0.1559 |
| NBC-kb | 180.5653 | 0.0314 | 0.1315 | 0.1592 |

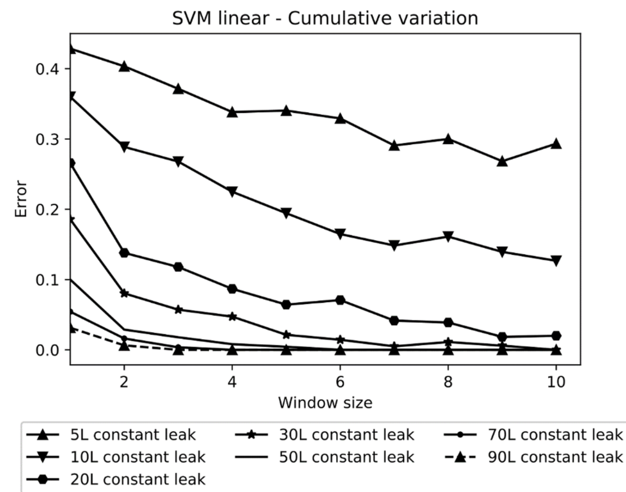


**Fig. 8.** Classification error based on window size for the SVM-linear classifier with the "cumulative variation" feature for different constant leak rates.

Fig. 8 shows how, for the best combination (SVM linear + cumulative variation), when the leak rate is very small (5 L), the classification error exhibits practically no decrease when the size of the time window is changed (from 43% to 31% with the SVM-linear classifier). We can also see how for medium leaks (30 L), the classification error is below 10% starting with a window size of 2. Likewise, for larger leaks, the error practically goes to zero starting with a window size of 5.

Figs. 9 and 10 show the rate for False Positives (FP) and False Negatives (FN) associated with the SVM-linear and cumulative variation combination. As we can see, the percentage of false positives for small, 5-L leaks varies as a function of the window size between approximately 20% and 15%. For larger leak sizes, the rate of false positives drops considerably; thus, for example, for 30-L leaks, after the second day the rate of false positives falls below 5%.

### *4.3. Classifier-feature combinations that satisfy the standard and minimise classification errors*

The results presented in the previous section show the expected behaviour of the system from the standpoint of the trends. However, in absolute terms, it is difficult to know if the data are within the margin that might be considered acceptable for deployment in a real system.

In an effort to obtain actual references on the performance of our

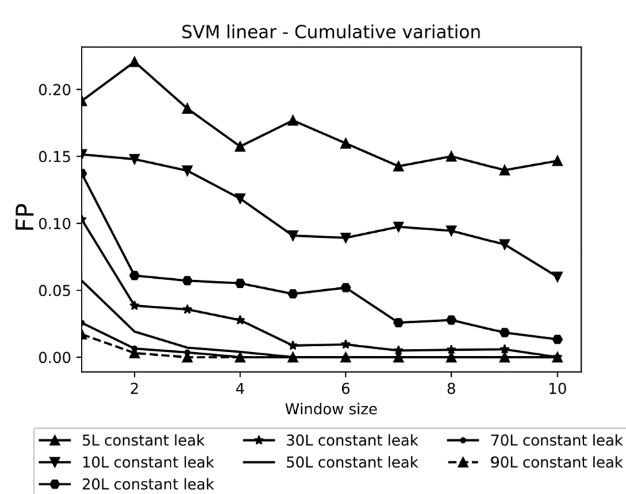


**Fig. 9.** False Positives based on window size for the SVM-linear classifier with the "cumulative variation" feature for different constant leak rates.

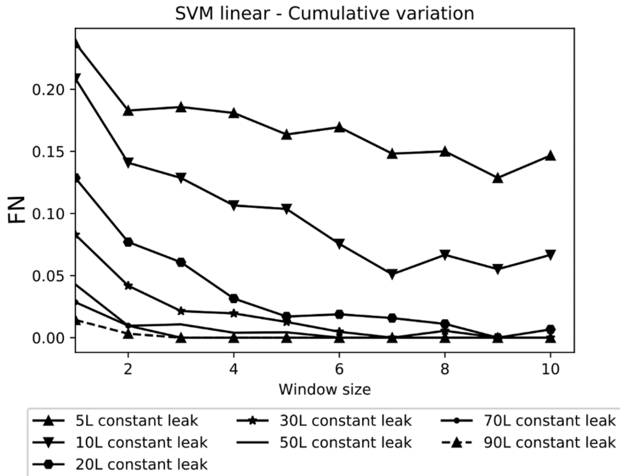


**Fig. 10.** False Negatives based on window size for the SVM-linear classifier with the "cumulative variation" feature for different constant leak rates.

system, we decided to compare our results with what the standard requires for detection systems. We did so by once again conducting the experimental tests for the leak rate specified in the applicable standard, UNE-EN 13160-5 (2017): a leak of 19.2 L/day within a maximum detection period of fourteen days. This was thus the size used for the time windows on this occasion.

Table 7 shows the classification error, the true positive rate (%), the false positive rate (%), and the number of days needed to identify the leak for each classifier and the feature with which it works best. We can see that for every classifier, the leak was detected within the error range allowed by the time limit specified in the standard. In fact, the time periods required for detecting the leak are significantly lower than those specified by the standard.

## 5. Conclusions

Sigut et al. (2014) marked the authors' first incursion into the field of fuel leak detection in service stations using classifier systems. Although the results obtained were very positive, we identified several aspects of the classification process that were subject to improvement. Specifically, we identified as one potential area of future research the possibility of using new features for the classification, paying particular attention to how said features evolved over time.

In this work, we again rely on classifiers to detect leaks, but this time we introduce time windows into the analysis to allow the effect of the leak to build up over the course of several consecutive days. This adds information involving the time trend of the features used in the classification process. As for the rest, the classification techniques used continue to be two-class classifiers, with the data to be classified corresponding to actual inventories from service stations.

The findings presented in the "Results and discussion" section show that the use of time windows not only allows detecting the leak, but also the time when the leak started. This information is of vital importance to the problem at hand.

In general terms, we conclude from the research presented in this paper that the classification error is smallest when the size of the time window is largest (when the classifiers make use of the features containing the cumulative value of the variables) and when the leak rate is highest.

The experiments conducted also allow us to state that, in general, the LDA classifier provides the best results. As concerns the study of the different features, we observed that the "cumulative variation" provides the most discriminating information. Lastly, our results show that the combination of the LDA classifier and the cumulative variation offers the second-best solution, with the first one being the SVM linear + "cumulative variation" combination, where the smallest leak can be detected in the shortest time possible.

Finally, we considered if there exist "classifier + feature" combinations capable of satisfying the UNE-EN 13160-5 (2017) standard, which requires that a leak be detected within a certain number of days and with a maximum specified error. In light of the results shown in Section 4.3, we found that the nominal leak rate of 19.2 L/day is detected well before the 14 days required in the standard for both a constant and a variable leak. Depending on the "classifier + feature" combination used, the leak is detected within a period of time that ranges between six and nine days (see Table 7).

The authors believe that using the most realistic data to train and test the classifiers is important. In this regard, as a future area of research we propose creating a set of simulated data for days with leaks that can recreate as well as possible the wide range of situations that may arise in actual practice. This would allow training the classifiers using more realistic scenarios.

The goal is to expand beyond the leak types described in the standard to include other possibilities. Along this line, it would be of special interest to conduct tests with classifiers that learn based on combined leak patterns, in which the leak does not necessarily start on the first day of the time window, this being more representative of reality.

**Table 7**
Classification error (%), true positive rate (%), false positive rate (%) and window size (days needed to detect the leak) for a nominal leak rate of 19.2 L/day for a constant leak (left) and a variable leak (right). FS2: cumulative variation, FS6: cumulative variation over real volume, FS8: variation and sales.

| Constant leak 19.2 L/day | | | | | | Variable leak 19.2 L/day | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Classifier** | **Best feature set** | **Error (%)** | **TPR (%)** | **FPR (%)** | **Time window size** | **Best feature set** | **Error (%)** | **TPR (%)** | **FPR (%)** | **Time window size** |
| LDA | FS8 | 3.03 | 99,50 | 2.53 | 7 | FS8 | 4.22 | 99,07 | 3.29 | 6 |
| KNN | FS6 | 3.58 | 98,97 | 2.55 | 7 | FS2 | 4.08 | 98,47 | 2.55 | 7 |
| NBC | FS6 | 3.58 | 98,97 | 2.55 | 7 | FS6 | 3.55 | 99,00 | 2.55 | 7 |
| SVM-g | FS6 | 3.58 | 98,97 | 2.55 | 7 | FS6 | 3.55 | 99,00 | 2.55 | 7 |
| SVM-l | FS6 | 3.58 | 98,97 | 2.55 | 7 | FS6 | 3.55 | 99,00 | 2.55 | 7 |
| QDA | FS6 | 3.58 | 98,97 | 2.55 | 7 | FS6 | 3.55 | 99,00 | 2.55 | 7 |
| NBC-kb | FS6 | 3.58 | 98,97 | 2.55 | 7 | FS6 | 3.55 | 99,00 | 2.55 | 7 |
| DTC | FS6 | 3.89 | 98,33 | 2.22 | 8 | FS6 | 4.44 | 97,78 | 2.22 | 8 |

## Acknowledgements

We want to thank Mr Eladio Hernández, an engineer at the company Repsol, for providing us with actual inventory data from service stations and for his guidance and support during our research.

## Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Appendix A

See TableA1–A8.

**Table A1**
Values of the ‘Acc Error’, ‘Min-Max Error’, ‘Avg Error’ and ‘Std Error’ for the various features with the LDA classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 14.6698 | 0.0343 | 1.0478 | 1.2467 |
| FS2 | 14.0534 | 0.0314 | 1.0038 | 1.1751 |
| FS3 | 15.7135 | 0.0286 | 1.1224 | 1.2480 |
| FS4 | 14.5846 | 0.0286 | 1.0418 | 1.1324 |
| FS5 | 15.6911 | 0.0457 | 1.1208 | 1.3348 |
| FS6 | 14.8069 | 0.0457 | 1.0576 | 1.2297 |
| FS7 | 14.8123 | 0.0457 | 1.0580 | 1.2554 |
| FS8 | 14.4243 | 0.0343 | 1.0303 | 1.3112 |

**Table A2**
Values of the ‘Acc Error’, ‘Min-Max Error’, ‘Avg Error’ and ‘Std Error’ for the various features with the QDA classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 17.54533 | 0.0314 | 1.2467 | 1.5255 |
| FS2 | 14.2353 | 0.0314 | 1.0168 | 1.1946 |
| FS3 | 20.5841 | 0.0400 | 1.4703 | 1.5709 |
| FS4 | 15.3065 | 0.0429 | 1.0933 | 1.1630 |
| FS5 | 18.7897 | 0.0429 | 1.3421 | 1.6235 |
| FS6 | 14.9819 | 0.0371 | 1.0701 | 1.2513 |
| FS7 | 23.4530 | 0.0486 | 1.6752 | 1.9337 |
| FS8 | 22.3434 | 0.0343 | 1.5960 | 2.1504 |

**Table A3**
Values of the ‘Acc Error’, ‘Min-Max Error’, ‘Avg Error’ and ‘Std Error’ for the various features with the NBC classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 17.8669 | 0.0314 | 1.2762 | 1.5544 |
| FS2 | 14.2353 | 0.0314 | 1.0168 | 1.1946 |
| FS3 | 23.8460 | 0.0400 | 1.7033 | 1.7705 |
| FS4 | 15.3065 | 0.0429 | 1.0933 | 1.1630 |
| FS5 | 19.0520 | 0.0429 | 1.3609 | 1.5916 |
| FS6 | 14.9819 | 0.0371 | 1.0701 | 1.2513 |
| FS7 | 20.5636 | 0.0343 | 1.4688 | 1.6666 |
| FS8 | 21.4751 | 0.0343 | 1.5339 | 1.9541 |

**Table A4**
Values of the ‘Acc Error’, ‘Min-Max Error’, ‘Avg Error’ and ‘Std Error’ for the various features with the NBC-kb classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 23.9973 | 0.0457 | 1.7141 | 1.8290 |
| FS2 | 14.6881 | 0.0457 | 1.0491 | 1.1764 |
| FS3 | 28.0237 | 0.1200 | 2.0017 | 1.6642 |
| FS4 | 15.5835 | 0.0429 | 1.1131 | 1.0880 |
| FS5 | 24.2219 | 0.0708 | 1.7301 | 1.7330 |
| FS6 | 15.5120 | 0.0400 | 1.1080 | 1.2689 |
| FS7 | 30.1125 | 0.1600 | 2.1509 | 1.6300 |
| FS8 | 28.4263 | 0.0571 | 2.0304 | 2.1696 |

**Table A5**
Values of the 'Acc Error', 'Min-Max Error', 'Avg Error' and 'Std Error' for the various features with the DTC classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 25.8980 | 0.0486 | 1.8499 | 1.7282 |
| FS2 | 18.2127 | 0.0400 | 1.3009 | 1.4592 |
| FS3 | 27.1728 | 0.0486 | 1.9409 | 1.7021 |
| FS4 | 18.2015 | 0.0457 | 1.3001 | 1.2989 |
| FS5 | 26.7437 | 0.0514 | 1.9103 | 1.8025 |
| FS6 | 18.9513 | 0.0543 | 1.3537 | 1.5474 |
| FS7 | 27.4602 | 0.0571 | 1.9614 | 1.7818 |
| FS8 | 27.3507 | 0.0429 | 1.9536 | 1.8738 |

**Table A6**
Values of the 'Acc Error', 'Min-Max Error', 'Avg Error' and 'Std Error' for the various features with the SVM-g classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 27.8116 | 0.0314 | 1.9865 | 2.5250 |
| FS2 | 14.1468 | 0.0314 | 1.0105 | 1.1867 |
| FS3 | 27.8195 | 0.0371 | 1.9871 | 2.3138 |
| FS4 | 14.6064 | 0.0371 | 1.0433 | 1.1229 |
| FS5 | 28.3621 | 0.0400 | 2.0259 | 2.4707 |
| FS6 | 15.1549 | 0.0400 | 1.0825 | 1.2861 |
| FS7 | 34.5541 | 0.0533 | 2.4681 | 2.5480 |
| FS8 | 44.5491 | 0.1400 | 3.1821 | 3.0228 |

**Table A7**
Values of the 'Acc Error', 'Min-Max Error', 'Avg Error' and 'Std Error' for the various features with the SVM-l classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 14.8363 | 0.0343 | 1.0597 | 1.2443 |
| FS2 | 13.9735 | 0.0314 | 0.9981 | 1.11683 |
| FS3 | 15.7072 | 0.0371 | 1.1219 | 1.2308 |
| FS4 | 14.4061 | 0.0371 | 1.0290 | 1.1142 |
| FS5 | 16.1694 | 0.0400 | 1.1550 | 1.3731 |
| FS6 | 14.8069 | 0.0400 | 1.0576 | 1.2517 |
| FS7 | 14.3661 | 0.0371 | 1.0261 | 1.1916 |
| FS8 | 14.8657 | 0.0343 | 1.0618 | 1.3251 |

**Table A8**
Values of the 'Acc Error', 'Min-Max Error', 'Avg Error' and 'Std Error' for the various features with the KNN classifier.

| Feature set | Acc Error | MinMax Error | Avg Error | Std Error |
|---|---|---|---|---|
| FS1 | 16.4678 | 0.0314 | 1.1763 | 1.4623 |
| FS2 | 14.6309 | 0.0314 | 1.0451 | 1.2590 |
| FS3 | 17.5541 | 0.0371 | 1.2539 | 1.4701 |
| FS4 | 14.4138 | 0.0371 | 1.0296 | 1.1016 |
| FS5 | 17.0290 | 0.0400 | 1.2164 | 1.4989 |
| FS6 | 15.2791 | 0.0400 | 1.0914 | 1.2994 |
| FS7 | 17.2027 | 0.0371 | 1.2288 | 1.4802 |
| FS8 | 86.8450 | 0.6222 | 6.2032 | 0.4777 |

## References


Arlot, S., 2010. A survey of cross-validation procedures for model selection. Statist. Surv. 4, 40–79.

Chofreh, A.G., Goni, F.A., Shaharoun, A.M., Ismail, S., Klemes, J.J., 2014. Sustainable enterprise resource planning: imperatives and research directions. J. Clean. Prod. 71, 139–147.

Collin, M.L., Melloul, A.J., 2003. Assessing groundwater vulnerability to pollution to promote sustainable urban and rural development. J. Clean. Prod. 11, 727–736.

Cover, T., Hart, P., 1967. Nearest neighbour pattern classification. IEEE Trans. Inf. Theory 13 (1), 21–27.

Duda, R.O., Hart, P.E., Stork, D.G., 2001. Pattern Classification, second ed. John Wiley & Sons, New York.

Erkman, S., 1997. Industrial ecology: an historical review. J. Clean. Prod. 1e2 (5), 1–10.

Gorawska, A., Pasterak, K., 2016. Anomaly detection in data streams: the petrol station simulator. Commun. Comput. Inform. Sci. 613, 727–736.

Gorawski, M., Gorawska, A., Pasterak, K., 2015a. Liquefied petroleum storage and distribution problems and research thesis. Commun. Comput. Inform. Sci. 521, 540–550. https://doi.org/10.1007/978-3-319-18422-7_48.

Gorawski, M., Skrzewski, M., Gorawski, M., Gorawska, A., 2015b. Neural networks in petrol station objects calibration. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 9532, 714–723.

Gorawski, M., Gorawska, A., Pasterak, K., 2017. The TUBE algorithm: discovering trends in time series for the early detection of fuel leaks from underground storage tanks. Expert Syst. Appl. 90, 356–373.

Gurr, T.M., Homann, R.L., 1996. Managing underground storage tanks. Pollut. Eng. 28, 40–44.

Health and Safety Authority, 2013. Wetstock Reconciliation at Fuel Storage Facilities - an Operator's Guide. http://www.hsa.ie/eng/Publications_and_Forms/Publications/Chemical_and_Hazardous_Substances/Wetstock_Reconciliation_at_Fuel_Storage_Facilities.pdf.

Hsu, C.W., Lin, C.J., 2002. A comparison of methods for multiclass support vector machines. IEEE Trans. on Neural Networks 13 (2), 415–425.

Introduction to Statistical Inventory Reconciliation for Underground Storage Tanks, 1995. US Environmental Protection Agency, Guide EPA 510-B-95-009.

Jain, A.K., Duin, R.P.W., Mao, J., 2000. Statistical Pattern Recognition: A Review. IEEE Trans. on Pattern Anal. and Mach. Intell., 22(1), 4–37.
Khan, F.I., Husain, T., 2003. Evaluation of a petroleum hydrocarbon contaminated site for natural attenuation using 'RBMNA' methodology. Environ. Model. Softw. 18 (2), 179–194.
Li, Z., Shui, A., Luo, K., Chen, J., Li, M., 2011. SIR-based oil tanks leak detection method. Proc. of the 2011 Chinese Control and Decision Conference. 1946-1950.
Maxwell, D., van der Vorst, R., 2003. Developing sustainable products and services. J. Clean. Prod. 11, 883–895.
Niculescu-Mizil, A., Caruana, R., 2005. Predicting good probabilities with supervised learning. Proc. of the 22nd International Conference on Machine Learning, Bonn, Germany. doi:10.1145/1102351.1102430.
Pang-Ning, T., Steinbach, M., 2013. Introduction to Data Mining. Pearson New International Edition.
Robert, K.-H., 2000. Tools and concepts for sustainable development, how do they relate to a general framework for sustainable development, and to each other? J. Clean. Prod. 8, 243–254.
Sigut, M., Alayón, S., Hernández, E., 2014. Applying pattern classification techniques to the early detection of fuel leaks in petrol stations. J. Clean. Prod. 80, 262–270.
UNE-EN 13160-5, 2005. Leak Detection Systems - Part 5: Tank Gauge Leak Detection Systems. https://www.aenor.com/normas-y-libros/buscador-de-normas/UNE?c=N0033325.
UNE-EN 13160-5, 2017. Leak detection systems - Part 5: Requirements and test/assessment methods for in-tank gauge systems and pressurised pipework systems. https://www.en.aenor.com/normas-y-libros/buscador-de-normas/une?c=N0058060.
United States Environmental Protection Agency Pacific Southwest/Region 9, 2003. Preventing Leaks and Spills at Service Stations - a Guide for Facilities. http:// www.epa.gov/region9/waste/ust/pdf/servicebooklet.pdf.