Journal of Cleaner Production 80 (2014) 262–270

Contents lists available at ScienceDirect

Journal of Cleaner Production

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

Applying pattern classification techniques to the early detection of fuel leaks in petrol stations

Marta Sigut*, Silvia Alayón, Eladio Hernández

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

ARTICLE INFO

Article history:
Received 3 March 2014
Received in revised form
21 May 2014
Accepted 21 May 2014
Available online 7 June 2014

Keywords:
Fuel leaks
Petrol stations
Inventories reconciliation
Classification methods

ABSTRACT

Leaks below the surface are one of the most serious problems in service stations with underground fuel storage tanks. These leaks result in pools of fuel, which flows into both the ground and the aquifers, polluting ecosystems and damaging them severely. In this paper, pattern classification techniques are used to carry out the early detection of fuel leaks in petrol stations. Early detection is crucial from an environmental point of view. The use of these classification methods requires properly selecting those variables that suitably represent the objects to classify, which in this case are the days when the petrol station is operative. In our study we use actual data provided by Repsol (a Spanish energy company) to construct these objects, which are then classified into two possible categories: "day without leaks" or "day with leaks", applying both supervised and unsupervised classifiers. Finally, three different combinations of "classifier + feature group" are proposed as possible solutions for the problem of the early detection of fuel leaks in service stations.

1. Introduction

Sustainable production is a topic of great interest for researchers. This is due to the fact that it is one of the most important challenges industry faces, as reflected in (Chofreh et al., 2014; Erkman, 1997; Labuschagne et al., 2005; Maxwell and van der Vorst, 2003; Robèrt, 2000; Veleva and Ellenbecker, 2001). In particular, the achievement of a balance between ecological and engineering approaches to land-use alteration as one of the keys for maximizing human-life quality is stated in (Collin and Melloul, 2003).

In modern service stations, the problem of fuel leaks is mitigated in part by using systems equipped with double-walled tanks and pipes. However, the great majority of service stations are old and, consequently, do not enjoy the full guarantees offered by current mechanical systems.

The problem of detecting generic fuel leaks nowadays is handled using different approaches. One involves using some kind of sensing system in an attempt to detect fuel leaks, as in (Morisawa and Muto, 2012).

Another way of detecting fuel leaks in petrol stations is by means of the regulatory leak tests that these kinds of systems must regularly comply with, as reflected in (United States Environmental Protection Agency Pacific Southwest/Region 9, 2003). The drawback of leak tests is that they are not frequent enough. Typically, this type of test must be able to determine at least every 30 days whether or not the tank and piping are leaking by using approved release detection methods. As a consequence, by the time a leak is detected it is often too late to avoid the environmental damage and the threat to human health and safety, requiring difficult and costly clean-ups. Therefore, the early detection of the fuel leaks is crucial, as is clearly explained in (United States Environmental Protection Agency, 2005).

Another possibility involves using different methods, such as inventory reconciliation analysis, to detect unaccounted fuel in the tank, as in (Health and Safety Authority, 2013). In any case, these detection procedures are not as simple as they seem to be. This is due to the existence of several factors like evaporation, volume variations due to temperature changes and so on, that lead to differences between the theoretical and actual inventories that are not caused by fuel leaks. Consequently, to be truly efficient leak detection procedures must distinguish the inventory differences caused by normal product behaviour from those due to a leak of product to the ground. In other words, the final goal of an inventory

Acronyms: LPG, Liquefied Petroleum Gas; FG1, Feature group 1; FG2, Feature group 2; FG3, Feature group 3; FG4, Feature group 4; FG5, Feature group 5; FG6, Feature group 6; LDA, Linear Discriminant Analysis; K-NN, K-Nearest Neighbour; FCM, Fuzzy c-means.

* Corresponding author.
E-mail addresses: marsigut@ull.es, marta@isaatc.ull.es (M. Sigut), salayon@ull.es (S. Alayón), eladiohd@ull.es (E. Hernández).

http://dx.doi.org/10.1016/j.jclepro.2014.05.070
0959-6526/© 2014 Elsevier Ltd. All rights reserved.

reconciliation system is to determine the behaviour pattern of a tank and its corresponding pipes, pumps, etc., to determine which differences can be considered normal and which cannot.

This method can be applied in two different ways: by using a system like the one made by the company Veeder Root that continuously analyses the level of fuel in the tank, or by using other systems that analyse the service station's daily closing inventory. One example of the latter is the one made by the English company Fairbanks, which markets a product that detects fuel leaks in petrol stations. This product is officially approved in various countries, including Spain. As is to be expected, this is a closed product in the sense that the user does not know the detection method used. In addition to this, the petrol station must send Fairbanks its daily inventory records, that is to say, its sales data, before it can detect a possible fuel leakage. Even in the online version of this product, Fairbanks has access to this information. For those petrol stations reluctant to share their sales data with an external company, this drawback could be avoided by using the fuel leak detection system described here.

The authors of this article make use of inventory reconciliation by applying pattern classification theory to classify the days when the service station is operative as "day without leaks" or "day with leaks". There are countless pattern recognition applications available, both in fields completely unrelated to the one considered in this article and also in leakage detection. In (Gabrys and Bargiela, 2000) these techniques are used for leak detection and identification in water distribution systems. In (Sato and Mita, 2007) leak detection in water supply systems relies on pattern recognition methods applied to sound signals, while a method based on waveform correlation delta-T detection in a buried gas pipeline is used in (Jiao et al., 2009). Finally, in (Da Silva et al., 2005) a method for pipeline leak detection that uses a combination of fault detection clustering and classification tools is presented. In this case, a classifier based on a fuzzy system is applied to a small-scale LPG (Liquefied Petroleum Gas) station. Even though pattern recognition is commonly used in a very wide range of applications, even in leak detection, to the best of our knowledge this method has not been applied to the detection of fuel leaks in petrol stations.

In this paper, four different types of classifiers (two supervised and two unsupervised) are applied to the detection of fuel leaks in petrol stations. As explained below, two different experiments were conducted for the purpose of testing the performance of these four classifiers in different situations. Current regulations applicable to this field were used to determine said performance.

2. Material and methods

In this section, the inventory reconciliation method, on which this work is based, is briefly described. The most relevant aspects concerning the use of classification algorithms for detecting fuel leaks in petrol stations are also presented.

2.1. The inventory reconciliation method

In ordinary usage, the daily inventory records are used in petrol stations. These records provide an inventory control of the quantities that are sold, bought and stored in each tank, in the same way as this would be done with any other product. These records can be kept manually or by using a computer.

The basic concepts in inventory reconciliation applied to service stations are the following:

- Theoretical inventory: quantity of fuel that it is expected to be inside the tank at the end of the day or shift. It is obtained using the expression below:

$$Theoretical\ inventory = Initial\ inventory + Receipts - Sales + Adjustments$$

where:

Initial inventory: inventory at the start of the day or shift.

Receipts: delivery made to the service station by the tanker-truck on the current day. The data can be obtained from the tanker-truck delivery note or from the tank's gauges.

Sales: total amount sold at the service station's petrol pumps.

Adjustments: correction due to certain maintenance operations that can produce an error. For example, the extraction of product through the petrol tank that appears as a sale but has instead been returned to the tank.

- Difference in inventory or variation: difference between the quantity of fuel measured in the tank and the one expected, defined as follows:

$$Variation = Theoretical inventory - Real inventory$$

- Accumulated difference in inventory or accumulated variation: sum of the variations over a period of time.
- Variance with respect of sales, defined as follows:

$$Variance = Variation/Sales$$

- Accumulated variance, given by the following expression:

$$Accumulated\ variance = Accumulated\ variation/Accumulated\ sales$$

An extract of the daily inventory records for a service station is shown in Table 1.

The main objective of the inventory reconciliation method is to detect the abnormal operation of the service station by comparing the theoretical and real inventories. However, a lack of correlation between the theoretical and real inventories at the end of the day does not necessarily imply a real leak or an excess of fuel. Any such mismatch can be due to the combination of several factors that are not easily accounted for, such as volume variation with temperature, fuel evaporation, errors in the metres, etc. Because of this, using the inventory reconciliation method to detect leaks is not a trivial task.

With respect to the rules that are applicable to leak detection, each country develops its own regulations. These include the minimum requirements with which the systems that comprise a service station must comply in order to detect and avoid fuel leaks below the surface. In an effort to obtain the appropriate references for this article, the authors have resorted to the family of European regulations EN 13160, which govern the fuel leak detection systems in service stations; more specifically, regulation EN 13160-5 (2005), which governs level indicator fuel leak detection systems, includes the reference used in this paper. This reference states that in order to consider a daily analysis as acceptable, the system must be able to detect leaks of up to 96 L/day with an error (both in false positives and false negatives) lower than 5%.

2.2. Detection of petrol leaks: definition of the classification problem

As noted in Section 1, the main goal of our research is to demonstrate that pattern classification methods can be useful for detecting petrol leaks. A general classification problem consists of N objects and

Table 1
Extract of a service station's daily inventory records. Columns 3 to 6 correspond to the different items (initial inventory, receipts, sales and adjustments) which are used to calculate the theoretical inventory shown in column 7 using the expression in Section 2.1. The different rows contain the quantities (in litres) measured every day as well as the subtotals for tanks 1 and 2, which contain diesel and diesel e+10 (a product of higher quality than diesel that is stored in a different tank and has its own inventory pattern), respectively.

Tank: fuel	Date	Initial	(+) Receipts	(−) Sales	(+/−) Adjustments	(=) Theoretical inventory available	Current available
1: DIESEL							
	22/01/2007	21,214.00	0.00	0.00	0.00	21,214.00	21,214.00
	23/01/2007	21,214.00	16,998.00	8553.13	0.00	29,658.87	29,564.00
	24/01/2007	29,564.00	10,002.00	12,119.26	0.00	27,446.74	27,448.00
	25/01/2007	27,448.00	12,000.00	9355.93	0.00	30,092.07	30,117.00
	26/01/2007	30,117.00	0.00	10,810.98	0.00	19,306.02	19,269.00
	27/01/2007	19,269.00	17,000.00	7038.59	0.00	29,230.41	29,258.00
	28/01/2007	29,258.00	0.00	1589.27	0.00	27,668.73	27,217.00
	29/01/2007	27,217.00	17,004.00	13,137.14	0.00	31,083.86	31,127.00
	30/01/2007	31,127.00	9984.00	11,523.93	20.00	29,607.07	29,379.00
	31/01/2007	29,379.00	0.00	13,823.09	0.00	15,555.91	15,204.00
Subtotal		265,807.00	82,988.00	87,951.32	20.00	260,863.68	259,797.00
2: DIESEL e+10							
	22/01/2007	9700.00	0.00	0.00	0.00	9700.00	9700.00
	23/01/2007	9700.00	4999.00	298.21	0.00	14,400.79	14,504.00
	24/01/2007	14,504.00	0.00	610.53	0.00	13,893.47	13,932.00
	25/01/2007	13,932.00	0.00	750.13	0.00	13,181.87	13,164.00
	26/01/2007	13,164.00	0.00	612.04	0.00	12,551.96	12,532.00
	27/01/2007	12,532.00	0.00	872.07	0.00	11,659.93	11,619.00
	28/01/2007	11,619.00	0.00	405.71	0.00	11,213.29	11,158.00
	29/01/2007	11,158.00	5004.00	686.24	0.00	15,475.76	15,537.00
	30/01/2007	15,537.00	0.00	595.37	20.0	14,961.63	14,977.00
	31/01/2007	14,977.00	0.00	921.00	0.00	14,056.00	14,106.00
Subtotal		126,823.00	10,003.00	5751.30	20.00	131,094.70	131,229.00

M classes. The classifier system assigns each object to a certain class after analysing the representative features of the object, as explained in (Duda et al., 2001). This classification process is shown in Fig. 1.

The design of a classification system requires careful consideration of the following issues: pattern class definition, pattern representation, feature extraction and selection, classifier design and learning, selection of training and test samples, and performance evaluation (Jain et al., 2000). When applied to the problem of petrol leaks, the N objects to classify are the N days when the petrol station is operative (pattern representation), and the M possible classes are two: “day with petrol leaks” and “day without petrol leaks” (pattern classes).

Regarding the feature extraction and selection process, the main problem is determining which variables can suitably represent the objects and how many features are needed to achieve good classification results. In this case, real data extracted from the inventories are used to build the possible features. These data consist of the following: the daily initial volume of the petrol tanks, the deliveries by petrol tankers, the daily petrol sales and the inventory of the petrol station, among others. The selection of the classifier's internal structure is also an important decision. There are two trends in classification procedures: supervised and unsupervised classification, as demonstrated in (Duda et al., 2001; Kotsiantis, 2007). Both strategies have been studied and four well-known classifiers are used in this work: two supervised classifiers (Linear Discriminant Analysis, see (Duda et al., 2001) and K-Nearest Neighbour algorithm, see (Cover and Hart, 1967)); and two unsupervised classifiers (K-means, see (Kanungo et al., 2002) and Fuzzy c-means, see (Bezdek, 1981)). Both the feature selection process and these four classifiers are explained in detail below.

In Fig. 2, the detection of petrol leaks is represented as a classification process. The input is an operating day at the petrol station

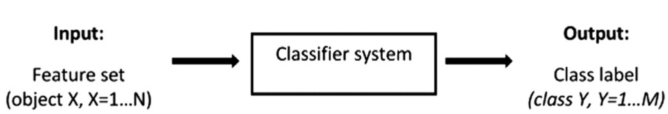

Fig. 1. Classification process.

(represented with a feature set that is composed of data extracted from the inventories), the classifier system is one of the four classification procedures used in this research work, and the output is a class label that indicates if the operating day under study is a “day with leaks” or a “day without leaks”.

2.3. Feature selection

As explained earlier, the features that represent the operating days of the petrol station are extracted from the daily inventory records.

Feature selection is a very important step in the classification process. A good feature selection can significantly improve the classifier's efficiency, as shown in (Duda et al., 2001). This is not a trivial task since the data could contain many redundant or irrelevant features. Redundant features are those which provide no more information than those currently selected, while irrelevant features provide no useful information in any context. In the problem at hand in this article, the most suitable features are unknown, hence the reason why different groups of features are tested:

- Feature group 1 (FG1) → one feature: variation (*variation* = *theoretical inventory* − *real inventory*)
- Feature group 2 (FG2) → one feature: sales variation (*sales variation* = *variation*/*sales*)
- Feature group 3 (FG3) → one feature: (*theoretical volume* = *real volume from the previous day* + *receptions* − *sales* + *adjustments*)
- Feature group 4 (FG4) → one feature: (*variation*/*real volume*)
- Feature group 5 (FG5) → two features: daily sales and variation
- Feature group 6 (FG6) → two features: daily sales and sales variation

2.4. Classification algorithms used in this work

Four well-known classifiers (two supervised and two unsupervised) have been applied in an attempt to solve the fuel leak detection problem with the feature groups mentioned above.

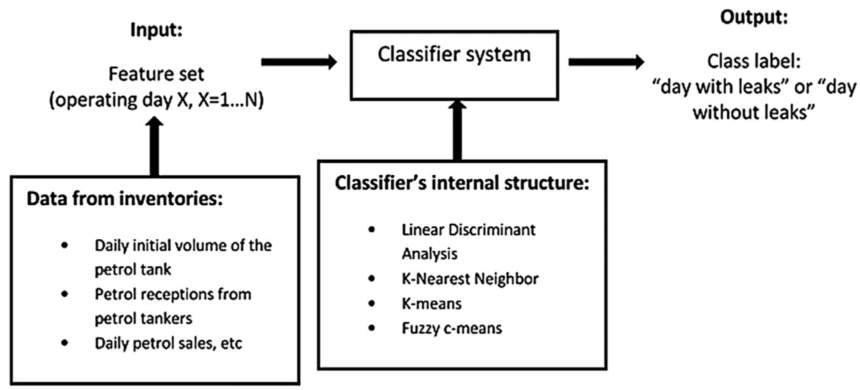

Fig. 2. The detection of petrol leaks considered as a classification process.

In supervised algorithms the classes are predetermined. A certain segment of data will be labelled with these classifications (training set). The classifier's task is to search for patterns and construct mathematical models by analysing the data from the training set. Once the training step is finished, these models are evaluated using new labelled data (test set). Two different supervised methods are used in this paper: Linear Discriminant Analysis (LDA) and K-Nearest Neighbour algorithm (K-NN). 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 maximal separability. This method fits a multivariate normal density to each group, with a pooled covariance estimate, as demonstrated in (Duda et al., 2001).
- 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). K-NN is a type of instance-based learning where the function is only approximated locally and all computations are deferred to the classification stage. This algorithm is amongst the simplest of all machine learning algorithms: 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).

For the K-NN algorithm, the value of K (number of closest neighbours) must be fixed. It is not possible to assign a value of K beforehand that will guarantee the best classification behaviour. That is why this parameter must be tuned empirically. For the selection of this parameter in this paper, all the experiments conducted with this classifier were repeated with the following values of K: 1, 3, 5, 7, 9, 11, 13 and 15. Then, the value of K was chosen based on the best classification results. For example, the classification error for a certain experiment carried out over the course of the research done by the authors with a K-NN classifier is shown in Fig. 3. As we can see, the lower classification error (10%) is obtained with $K = 3$. Therefore, this will be the value of K selected for this experiment. In the experiments described in Section 3, the K-NN classifiers perform better in most cases when $K = 5$.

Unsupervised classifiers are not provided with labelled data. These algorithms seek out similarity between pieces of unlabelled data in order to determine whether they can be characterised as forming a group (cluster). Two different unsupervised methods are used in this paper: K-means and Fuzzy c-means (FCM). Both are briefly explained below:

- K-means is a well-known algorithm for clustering, as demonstrated by (Kanungo et al., 2002). The user has to specify the number of groups (referred to as K) to be identified. Each object can be thought of as being represented by some feature vector in an n-dimensional space, where n is the number of features used to describe the objects to cluster. The algorithm randomly chooses K points in that vector space. These points serve as the initial centres of the clusters. Then, every object is assigned to the closest centre. Usually, the distance measurement is chosen by the user and determined by the learning task. Then, for each cluster a new centre is computed by averaging the feature vectors of every object assigned to it. The process of assigning objects and recomputing centres is repeated until the process converges.
- FCM is a method of fuzzy clustering which allows objects to belong to more than one cluster, with a set of membership levels associated with each object, as demonstrated in (Bezdek, 1981). These indicate the strength of the association between that object and a particular cluster. It is an adaptation of the K-means algorithm using fuzzy logic.

2.5. Real data available for this study

The data used in the experiments described in this paper are real data provided by the Spanish energy company Repsol. The data consist of daily inventory records from actual petrol stations operating without leaks (see Table 1). These data are used to construct different feature groups, as explained in Section 2.3.

Note that since there are no data available for abnormal operating situations (with petrol leaks), these have been simulated by altering the original data. This alteration procedure is indicated in the UNE-EN 13160-3 (2004) standard (Leak detection systems – Part 3: Liquid systems for tanks). Two different forms of data alteration were tested, resulting in different data sets representing situations with fuel leaks:

- Simulation of constant leaks (DATA SET 1): the UNE-EN 13160-3 (2004) standard establishes that a constant petrol leak must be simulated by altering the original data.
- Simulation of variable leaks (DATA SET 2): the UNE-EN 13160-5 (2005) standard establishes that a variable leak is a leak that decreases when the amount of stored petrol is reduced. Thus, data set 2 includes simulated leaks depending on the volume of petrol stored in the tank.

In order to construct training and test sets for the classification process, real available data representing normal operations without leaks were combined with simulated data representing abnormal operations with leaks. For the supervised classifiers the training and test sets are composed of data corresponding to 140 operating

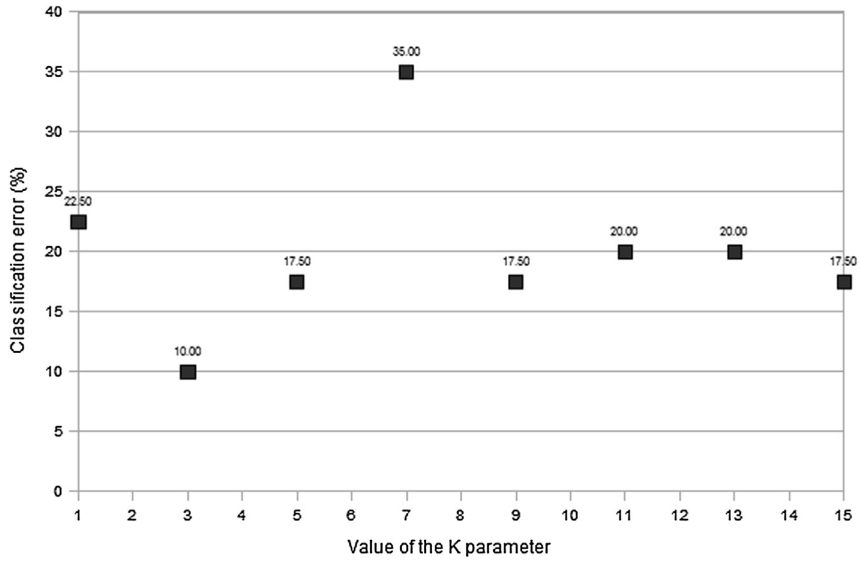

Fig. 3. Classification error (*Y*-axis) of a K-NN classifier as a function of the parameter *K* (*x*-axis).

days (70 days without leaks and 70 days with leaks). The unsupervised classifiers were applied directly to the test sets.

3. Calculation

The experiments carried out are divided into two main groups: experiments with data set 1 (simulating a constant petrol leak) and with data set 2 (simulating a variable petrol leak). The six sets of features described in Section 2.3 are used in different cases to assess those that might be of use for this classification problem. All the experiments were carried out using the Matlab (Version 7) software. The Statistics Toolbox was used for the LDA, KNN and K-means classifiers, while the Fuzzy Logic Toolbox was used for the FCM classifier.

3.1. Experiment 1. Simulation of constant leaks

Data set 1, described in Section 2.5, simulates a constant leak. This data set is composed of 13 subsets, each one simulating a constant leak with a different value: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120 and 130 L/day. The feature groups selected for these experiments are FG2, FG3, FG4 and FG6. FG1 and FG5 were discarded because both use the variable "variation", which is constant. Note that as demonstrated in (Duda et al., 2001), the LDA classifier requires a strictly positive covariance matrix, which prevents the zero covariance matrix corresponding to constant data from being used. Since the goal of the authors is to compare the results obtained with the four classifiers selected, FG1 and FG5 were not used in this experiment. The four classifiers were applied to each subset four times, each one testing a different feature group. The experimental procedure followed in experiment 1 is shown schematically in Fig. 4.

3.2. Experiment 2. Simulation of variable leaks

As noted in Section 2.5, data set 2 simulates variable leaks. This situation better reflects reality than the situation simulated in the experiment 1 because a real tank does not leak exactly the same amount of petrol every day. The magnitude of a leak depends on the part of the tank where it is located and on the pressure applied at this point. The hydrostatic pressure is higher for leaks located at the bottom of the tank than at the top. As a result, the UNE-EN 13160-5 (2005) standard establishes that a variable leak is a leak that decreases when the amount of stored petrol is reduced. In this case, the leak is simulated by altering the data in the daily inventory records corresponding to the real volume measured at the beginning of the day. The altered value is a percentage of this volume. Data set 2 consists of 12 subsets, each one simulating different variation percentages: 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2 and 1.3%. The feature groups selected in experiment 2 are FG1, FG2, FG3, FG4 and FG5. FG6 was discarded because it is redundant. As shown in the section where the feature groups are defined, FG6 provides information that is contained in FG1, FG2 and FG5. Since these three feature groups are used in this second experiment (the variable "variation" is now variable and, consequently, FG1 and FG5 can be used), it makes no sense to include FG6. In this case, the four classifiers are applied to each subset five times, each one testing a different feature group. The experimental procedure followed in experiment 2 is shown schematically in Fig. 5.

4. Results and discussion

In this section we present the results obtained in the experiments carried out to simulate two different kinds of leaks and our main conclusions from these experiments.

4.1. Experiment 1. Simulation of constant leaks

The results shown in this section stem directly from the classification of the test sets. The error was studied separately: a classification error for data from days with no leaks (false positives) and a classification error for data from days with leaks (false negatives). Both errors are important. Table 2 shows the classification errors for the 10 first subsets of data set 1 when the classifiers act separately on the feature groups FG2, FG3, FG4 and FG6. The errors in the left part of the cells correspond to days without leaks (false positives), while those in the right part of the cells (numbers in bold) correspond to days with leaks (false negatives).

As Table 2 shows, when FG2 is used, both errors are high when the constant leak rate is low. This is understandable because a low leak rate is more difficult to detect than a high leak rate. But when the leak rate is 80 L/day (or higher) the classification results are acceptable with every classifier (<15% in the worst case and 0% in the best cases). These results indicate that feature FG2 can be useful for the classification problem in question.

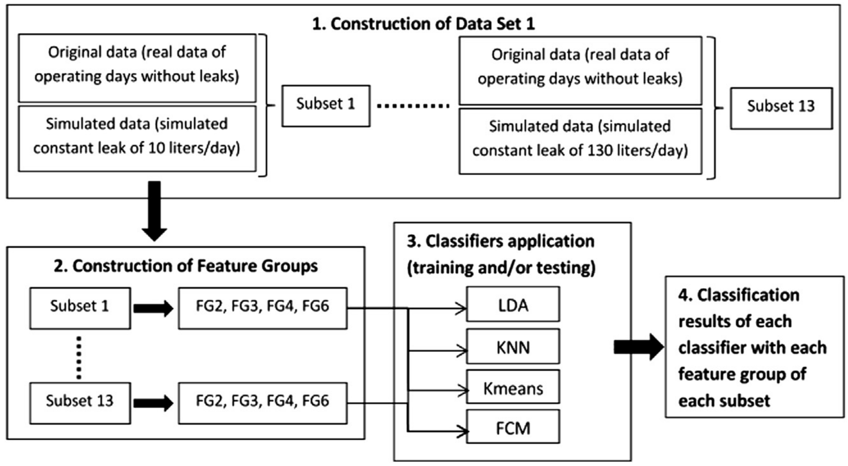

Fig. 4. Flow chart for the design and development of experiment 1 (simulation of a constant leak).

When FG3 is used, the LDA algorithm exhibits the worst behaviour (an error of 90% even with leak rates greater than 70 L/day) when classifying data from days without leaks and a 0% error for days with leaks. That means that according to this classifier, there were leaks almost every day. In contrast, FCM and K-means classify days without leaks very well (a 0% error with leak rates greater than 60 L/day) but has errors of 25–30% for days with leaks. Only the K-NN algorithm seems to exhibit balanced behaviour, with an error of around 20% in both cases with leaks higher than 60 L/day. Compared to FG2, the FG3 feature has a worse classification efficiency.

LDA does not perform acceptably with FG4 in any case with a constant leak rate, though the other three classifiers perform very well even with a low, constant leak rate (errors of 0–15% using both data sets with leak rates from 40 L/day). Consequently, FG4 has demonstrated to be a good feature for the classification problem.

Finally, when FG6 is used, the only algorithm that performs acceptably is the LDA classifier (for leak rates higher than 70 L/day: errors of 5–10% on days without leaks and a 0% error on days with leaks). Thus, the combination in this group of features (FG6) and this classifier (LDA) seems to be adequate.

As mentioned earlier, the UNE-EN 13160-5 (2005) standard specifies than an acceptable leak detection system must be able to detect constant leak rates of 90 L/day with an allowable error of 5%. A summary comparison between the four classifiers (LDA, K-NN, K-means and FCM) operating with the different feature groups analysed above (FG2, FG3, FG4 and FG6) is shown in Table 3.

4.2. Experiment 2. Simulation of variable leaks

The results obtained in this experiment are shown in the table below.

Table 4 contains the same information as Table 2 for this second experiment. Once more, the errors corresponding to days with leaks are in bold, and only the 10 first subsets of data set 2 are shown. As we can see, when FG1 (variation) is used, the LDA algorithm exhibits the worst behaviour: an error of 25–65% on days without leaks and a 0% error on days with leaks. That means that this algorithm identified almost every day as a day with a leak. In contrast, K-means and FCM classify the data with an error of 0–5% in both cases (days with and without leaks) from volume percentages of 0.6% and higher. With K-NN, the error values are worse, but better than those obtained with the LDA classifier. The unsupervised classifiers perform better with FG1.

When FG2 (sales variation) is used, the error in the unsupervised classifiers is low for days without leaks but very high and irregular for days with leaks. This means that they classify almost every day as a day without leaks. The K-NN algorithm yields an error of 15% on days without leaks and a 0% error on days with leaks when the volume percentage is 0.7% or higher. The LDA algorithm exhibits an error of 15% on days without leaks and a 0% error on days with leaks when the volume percentage is 1.1% or higher. Therefore, K-NN is better than LDA with the FG2 feature group. In general, the supervised classifiers in this experiment offer better performance with FG2, although their efficiency is not very high (errors are higher than 5%).

When FG3 (theoretical volume) is used, all the classifiers perform unacceptably. Errors are high in both situations (days without and with leaks). Thus, FG3 is not a suitable feature group in this experiment.

When FG4 (variation/real volume) is used, the behaviour of LDA is very inefficient, while the other classifiers perform better. Specifically:

- K-NN: an error of 5% on days without leaks and a 5% error on days with leaks when the volume percentage is 0.6%. These errors decrease with higher leak rates.
- K-means: an error of 5% on days without leaks and a 5% error on days with leaks when the volume percentage is 0.5% or higher.
- FCM: an error of 5% on days without leaks and the same percentage on days with leaks when the volume percentage is between 0.5% and 0.9%.

Finally, when FG5 (daily sales and variation) is used, all the classifiers perform unacceptably. Errors are high in both situations (days without and with leaks). Only the LDA algorithm can classify days without leaks with an error of 10% and days with leaks with a 0% error, but only when the leak is very high (volume percentages of 1.2% and 1.3%). Thus, it can be concluded that FG5 is not a suitable feature group in this experiment.

Note that the classification results obtained in this second experiment (variable leak simulated) are worse than those obtained in the first one (constant leak simulated). This is due to the fact that the leak values are now more dispersed since they depend on the tank volume, which changes every day.

The rule in the UNE-EN 13160-5 (2005) standard was used in experiment 2 for the purpose of establishing a comparative analysis between this one and experiment 1. In the available real data, the

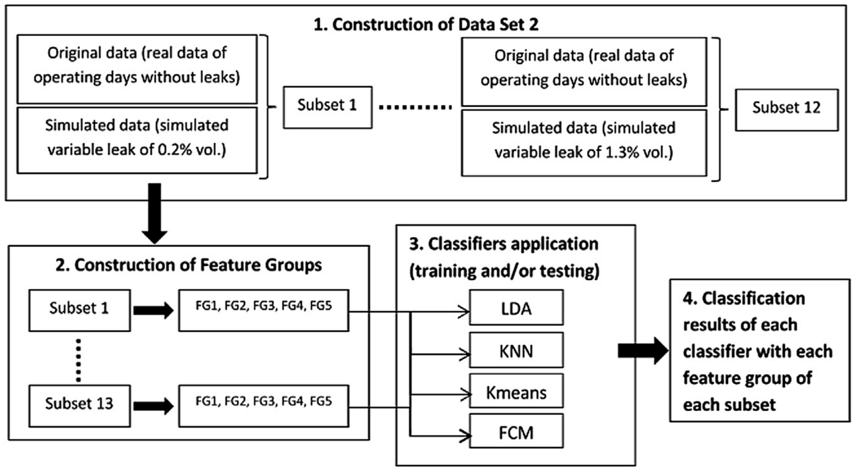

Fig. 5. Flow chart for the design and development of experiment 2 (simulation of a variable leak).

highest petrol volumes are around 15,000 L. Therefore, a leak of 90 L is approximately equivalent to a 0.6% of this volume. So this percentage value is now the reference for establishing whether a classifier is "good" (performs with an error of 5% or less for this percentage), "fair" (performs with an error of 5% or less for percentages higher than 0.6%) or "bad" (higher error at any percentage).

A summary comparison between the performances of the four classifiers operating with the feature groups mentioned above is shown in Table 5. These experimental results allow choosing the most suitable combination of "classifier + feature group" for this classification problem. Before doing so, some observations must be taken into account:

- All of the experiments were carried out with different leak values. It is clear that low leak rates are more difficult to detect than elevated ones, so the classification error of a good classifier is expected to decrease as the leak rate increases. A classifier that does not comply with this requirement must be rejected.
- When comparing the behaviour of two different classifiers, one is better than the other if it is able to correctly classify lower leak rate situations with a lower error value.

Based on these criteria and on the results shown in Tables 2–5, the FG3 feature set can be rejected. All of the classifiers tested perform incorrectly when applied to this feature set because their classification errors do not decrease when the leak rate increases.

For the remaining feature sets, the classification results are variable. For example, LDA never performs well with the FG4 feature set. This may be interpreted in one of two ways:

a) This feature set is not useful for the classification problem considered: this is not true because FG4 is very useful for the other three classifiers in all of the experiments performed.
b) The LDA algorithm is not suitable for this classification problem. This must be analysed in depth since LDA clearly performs better than the other classifiers in the first experiment with FG6 and in the second experiment with FG5.

Table 2
Results from experiment 1. The different rows correspond to the four classifiers used (LDA, K-NN, K-means and FCM) for the feature groups FG2, FG3, FG4 and FG6. For each classifier and feature group, the classification error percentages both for a day without leaks (on the left) and for a day with leaks (on the right, in bold) are shown for a constant leak rate of 10, 20, … 100 L/day.

FG2	10		20		30		40		50		60		70		80		90		100	
LDA	65	**0**	85	**0**	80	**0**	80	**0**	65	**0**	55	**0**	40	**0**	5	**0**	15	**0**	5	**0**
K-NN	25	**0**	30	**35**	20	**0**	15	**5**	10	**25**	5	**0**	5	**0**	5	**0**	5	**0**	5	**0**
K-means	85	**0**	55	**0**	35	**0**	20	**0**	15	**0**	5	**25**	5	**0**	5	**0**	0	**0**	0	**0**
FCM	55	**0**	55	**0**	30	**0**	15	**0**	15	**0**	5	**0**	5	**0**	0	**0**	0	**0**	0	**0**
FG3	10		20		30		40		50		60		70		80		90		100	
LDA	85	**0**	90	**0**	90	**0**	90	**0**	90	**0**	90	**0**	90	**0**	90	**0**	90	**0**	60	**0**
K-NN	30	**35**	30	**45**	25	**10**	10	**20**	10	**30**	10	**20**	10	**20**	10	**20**	10	**15**	10	**20**
K-means	90	**25**	35	**25**	10	**25**	5	**25**	5	**25**	0	**25**	0	**25**	0	**25**	0	**25**	0	**25**
FCM	45	**25**	35	**25**	10	**30**	5	**30**	5	**30**	0	**30**	0	**30**	0	**30**	0	**30**	0	**30**
FG4	10		20		30		40		50		60		70		80		90		100	
LDA	75	**0**	80	**0**	80	**0**	85	**0**	90	**0**	95	**0**	100	**100**	95	**0**	40	**100**	100	**100**
K-NN	30	**15**	25	**20**	15	**10**	5	**15**	5	**10**	10	**10**	5	**0**	5	**0**	0	**5**	5	**5**
K-means	85	**0**	55	**0**	30	**0**	10	**0**	10	**0**	5	**0**	5	**0**	5	**0**	5	**0**	5	**0**
FCM	55	**0**	45	**0**	10	**0**	10	**0**	5	**0**	5	**0**	5	**0**	5	**0**	5	**0**	5	**0**
FG6	10		20		30		40		50		60		70		80		90		100	
LDA	45	**0**	30	**0**	25	**0**	20	**0**	20	**0**	10	**0**	10	**0**	5	**0**	5	**0**	5	**0**
K-NN	35	**65**	35	**65**	35	**65**	35	**65**	35	**65**	35	**65**	35	**65**	35	**65**	35	**65**	35	**65**
K-means	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**
FCM	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**

Table 3
Compliance with the UNE-EN 13160-5 standard by the four classifiers in experiment 1. The classifiers' performance is rated as follows: "good" means that the classifier reaches the condition required by the standard (full compliance), "fair" means that the classifier fulfils the error condition only when the leak rate is higher than 90 L/day (partial compliance), and "bad" means that the classifier does not reach the condition with any leak rate (zero compliance).

	FG2	FG3	FG4	FG6
LDA	Fair	Fair	Bad	Good
K-NN	Good	Bad	Good	Bad
K-means	Good	Bad	Good	Bad
FCM	Good	Bad	Good	Bad

Table 5
Compliance with the UNE-EN 13160-5 standard by the four classifiers in experiment 2. The classifiers' performance is rated as follows: "good" means that the classifier reaches the condition required by the standard (full compliance), "fair" means that the classifier fulfils the error condition only when the leak rate is higher than 90 L/day (partial compliance), and "bad" means that the classifier does not reach the condition with any leak rate (no compliance).

	FG1	FG2	FG3	FG4	FG5
LDA	Fair	Fair	Bad	Bad	Fair
K-NN	Good	Fair	Bad	Good	Bad
K-means	Good	Bad	Bad	Good	Bad
FCM	Good	Bad	Bad	Good	Bad

Therefore, in light of the results shown in Tables 3 and 5, selecting the best feature set and the best classifier is not a trivial task for this classification problem. Feature sets FG5 and FG6 each consist of two features. It could be supposed that these feature sets could help any classifier to improve its efficiency since they apparently provide more information. However, in both experiments we found that K-NN, K-means and FCM classifiers perform better with FG4, which consists of a single feature. As a result, the scheme for the final solution proposed is as shown in Fig. 6.

Notice that it is not possible to compare the results obtained using the method proposed in this paper with those of other researchers for two main reasons. Firstly, the results mentioned in Section 1 involving the work done by other researchers who make use of classification methods to detect leaks are not comparable to those presented here because they correspond to very different applications. For example, in Gabrys and Bargiela (2000), Sato and Mita (2007) and Jiao et al. (2009) the authors try to detect water or gas leaks. It is clear that the nature of these problems differs fundamentally from the nature of the problem considered in this paper. In particular, the tolerance thresholds for what is considered an acceptable leak rate are completely different, meaning the results are not at all comparable. Secondly, in the case of the results subject to comparison with those presented in this article, which are the ones obtained by the two companies mentioned in Section 1 (Veeder Root and Fairbanks), the problem is that the data are not available since they belong to the energy company that contracts them to provide this service. As a consequence of this, no comparison is possible.

5. Conclusions

Comparing the results of experiments 1 and 2 (Tables 3 and 5), we see that combining FG4 with K-NN, K-means or FCM yields adequate results that are fully compliant with the rule contained in the UNE-EN 13160-5 (2005) standard. As a consequence of this, the authors propose any of these combinations (K-NN with FG4, K-means with FG4 or FCM with FG4) as the classification solution for detecting fuel leaks in service stations (see Fig. 6).

Finally, even though we believe that the results presented in this article comply with expectations, it is clear that there exist several aspects that are open to improvement. Our findings thus point the way toward the following future areas of research: the design of a pre-processing method for the classifier input data that can remove irrelevant outliers; the study of the possibility of using new features for the classification, paying special attention to the inclusion of the features' evolution in time; the study of the behaviour of more advanced classifiers, such as neural networks; and the application of "one-class" classifiers to the problem in question. Since these classifiers are capable of learning the data distribution associated with only one class, they are useful in problems where data from

Table 4
Results obtained in experiment 2. The different rows correspond to the four classifiers used (LDA, K-NN, K-means and FCM) for the feature groups FG1, FG2, FG4 and FG5. For each classifier and feature group, the classification error percentages both for a day without leaks (on the left) and for a day with leaks (on the right, in bold) are shown for a variable leak with a percentage in variation (Theoretical inventory - Actual inventory) of 0.2%, 0.3%, … 1.1%.

FG1	0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9		1.0		1.1	
LDA	65	**0**	65	**0**	60	**0**	55	**0**	55	**0**	50	**0**	45	**0**	30	**0**	30	**0**	25	**0**
K-NN	30	**20**	15	**0**	10	**5**	10	**0**	10	**0**	15	**5**	10	**5**	10	**0**	5	**0**	5	**0**
K-means	55	**0**	30	**0**	15	**5**	15	**5**	5	**5**	5	**5**	5	**5**	5	**5**	5	**5**	5	**5**
FCM	55	**0**	25	**0**	15	**5**	15	**5**	5	**5**	5	**5**	5	**5**	5	**5**	5	**5**	5	**5**
FG2	0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9		1.0		1.1	
LDA	85	**0**	85	**0**	85	**0**	85	**0**	85	**0**	85	**0**	80	**0**	50	**0**	25	**0**	5	**0**
K-NN	40	**20**	30	**15**	20	**25**	20	**0**	20	**5**	15	**0**	15	**0**	15	**0**	15	**0**	15	**0**
K-means	55	**0**	15	**30**	5	**35**	5	**35**	0	**55**	0	**40**	0	**30**	0	**40**	0	**40**	0	**30**
FCM	35	**15**	20	**20**	15	**20**	5	**25**	5	**25**	0	**30**	0	**30**	0	**30**	0	**30**	0	**25**
FG3	0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9		1.0		1.1	
LDA	95	**0**	95	**0**	95	**0**	95	**0**	95	**0**	90	**0**	90	**0**	80	**0**	40	**5**	5	**20**
K-NN	50	**20**	50	**40**	35	**25**	20	**25**	15	**20**	10	**20**	15	**20**	25	**20**	25	**20**	10	**20**
K-means	30	**30**	5	**45**	0	**50**	0	**50**	0	**20**	0	**50**	0	**50**	0	**40**	0	**40**	0	**50**
FCM	30	**30**	5	**40**	0	**45**	0	**45**	0	**45**	0	**45**	0	**45**	0	**45**	0	**40**	0	**40**
FG4	0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9		1.0		1.1	
LDA	50	**0**	50	**0**	45	**0**	45	**0**	50	**0**	65	**0**	85	**50**	0	**100**	0	**100**	95	**0**
K-NN	25	**20**	15	**20**	15	**20**	5	**15**	5	**5**	0	**10**	0	**10**	5	**0**	5	**0**	5	**0**
K-means	55	**0**	30	**0**	10	**5**	5	**5**	5	**5**	5	**5**	5	**5**	5	**5**	5	**5**	5	**5**
FCM	30	**10**	20	**5**	10	**5**	5	**10**	5	**5**	5	**5**	5	**5**	5	**5**	5	**10**	0	**10**
FG5	0.2		0.3		0.4		0.5		0.6		0.7		0.8		0.9		1.0		1.1	
LDA	60	**0**	55	**0**	55	**0**	50	**0**	50	**0**	45	**0**	40	**0**	35	**0**	25	**0**	25	**0**
K-NN	30	**45**	25	**40**	25	**40**	35	**50**	25	**40**	15	**35**	15	**40**	15	**40**	15	**35**	15	**35**
K-means	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**
FCM	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**	45	**50**

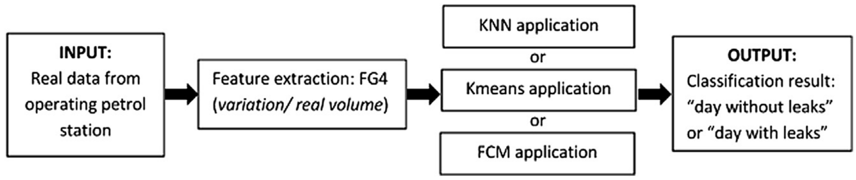

Fig. 6. Schematic view of the final solution proposed for the early detection of fuel leaks in petrol stations using the classifiers studied.

only one class is available. The "one-class" classifiers construct a boundary region around the data from the class in question and classify the test data as a function of the distance to this region. In the problem considered in this paper, the only real data available correspond to the "days without leaks" class. Consequently, the use of "one-class" classifiers would avoid the necessity of simulating the data corresponding to the "day with leaks" class.

Acknowledgements

The authors would like to thank Moisés Magdalena, a graduate student at the University of La Laguna, for his collaboration during the early stages of this research.

References

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, Norwell, Massachusetts.
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.
Da Silva, H.V., Morooka, C.K., Guilherme, I.R., Da Fonseca, T.C., Mendes, J.R.P., 2005. Leak detection in petroleum pipelines using a fuzzy system. J. Petrol. Sci. Eng. 49, 223–238.
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. 5 (1–2), 1–10.
Gabrys, B., Bargiela, A., 2000. General fuzzy min-max neural network for clustering and classification. IEEE Trans. Neural Netw. 11 (3), 769–783.
Health and Safety Authority, 2013. Wetstock Reconciliation at Fuel Storage Facilities – an Operator's Guide. http://www.hsa.ie/eng/Your_Industry/Petrol_Stations/Wetstock_Reconcilliation.pdf.
Jain, A.K., Duin, R.P.W., Mao, J., 2000. Statistical pattern recognition: a review. IEEE Trans. Pattern Anal. Mach. Intell. 22 (1), 4–37.
Jiao, Y., Li, G.H., Sun, G.K., Zhu, Y.K., Zhao, Y.B., 2009. Study on Leak Location Technology for Buried Gas Pipeline. International Conference on Wavelet Analysis and Pattern Recognition, pp. 492–495.
Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y., 2002. An efficient k-means clustering algorithm: analysis and implementation. IEEE Trans. Pattern Anal. Mach. Intell. 24 (7), 881–892.
Kotsiantis, S.B., 2007. Supervised machine learning: a review of classification techniques. Informatica 31, 249–268.
Labuschagne, C., Brent, A.C., van Erck, R.P.G., 2005. Assesing the sustainability performances for industries. J. Clean. Prod. 13, 373–385.
Maxwell, D., van der Vorst, R., 2003. Developing sustainable products and services. J. Clean. Prod. 11, 883–895.
Morisawa, M., Muto, S., 2012. Plastic optical fibre sensing of fuel leakage in soil. J. Sens. 2012, 1–6.
Robèrt, 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.
Sato, T., Mita, A., 2007. Leak Detection Using the Pattern of Sound Signals in Water Supply Systems. Proceedings of SPIE – Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems, p. 6529.
UNE-EN 13160–3, 2004. Leak Detection Systems – Part 3: Liquid Systems for Tanks. http://www.aenor.es/aenor/normas/normas/fichanorma.asp?tipo=N&codigo=N0030646&PDF=Si#.UlhjHVBLPwY.
UNE-EN 13160–5, 2005. Leak Detection Systems – Part 5: Tank Gauge Leak Detection Systems. http://www.aenor.es/aenor/normas/normas/fichanorma.asp?tipo=N&codigo=N0033325&pdf=#.UlhmrlBLPwY.
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.
United States Environmental Protection Agency, 2005. Straight Talk on Tanks – Leak Detection Methods for Petroleum Underground Storage Tanks and Piping. http://www.epa.gov/oust/pubs/STOT05_Rev4-6-09.pdf.
Veleva, V., Ellenbecker, M., 2001. Indicators of sustainable production: framework and methodology. J. Clean. Prod. 9, 519–549.