A fault detection system based on unsupervised techniques for industrial control loops

Esteban Jove[1][0000−0002−0625−359X], José-Luis Casteleiro-Roca[1][0000−0001−9740−6477], Héctor Quintián[1][0000−0002−0268−7999], Juan Albino Méndez-Pérez[2][0000−0003−2004−5989], and José Luis Calvo-Rolle[1][0000−0002−2333−8405]

How to cite:
Jove E, Casteleiro-Roca J-L, Quintián H, Méndez-Pérez JA, Calvo-Rolle JL. A fault detection system based on unsupervised techniques for industrial control loops. Expert Systems. 2019; 36:e12395.
https://doi.org/10.1111/exsy.12395

[1] University of A Coruña, CTC, Department of Industrial Engineering Avda. 19 de febrero s/n, 15405, Ferrol, A Coruña, Spain
[2] Universidad de La Laguna, Department of Computer Science and System Engineering, Avda. Astrof. Francisco Sánchez s/n, 38200, S/C de Tenerife, Spain
{esteban.jove@udc.es,jose.luis.casteleiro@udc.es,hector.quintian@udc.es, jamendez@ull.edu.es,jlcalvo@udc.es

Abstract. Power cells have presented an increasing popularity during last decades due to its importance in electric mobility, electronic devices and energy management systems. The international expansion of green policies to promote electric cars and renewable energies, has resulted in the need of ensuring their quality and reliability performance. In this context, detecting any early deviation from the correct operation must be addressed. Hence, this work is focused on the fault detection in a Lithium Iron Phosphate – LiFePO4 (LFP) cell. This is achieved by means of different one-class techniques, whose performance is assessed through artificially generated anomalies. After analysing the behaviour of each tested technique, the chosen classifier presents a successful performance.

Keywords: Power cell · Fault detection · Anomaly detection · One-class

1 Introduction

The vast majority of industrial companies perform complex and expensive processes [24, 2]. It is possible to optimize the operation of those processes leading to an increased efficiency in the use of energy and resources; and a greater quality of the final product. As a result of this, companies can become more competitive and enjoy the benefits of higher economic gains [32]. Hence, system optimization plays a key role in industrial activities.

To effectively optimize the operation of a system it is necessary that all its components function correctly, such as sensors, actuators and so on. Abnormal operation or anomalies have multiple sources [22, 37, 39, 28]; mechanical faults, changes in the system behavior, sensor error and human mistakes. Hence, anomalies are a central and frequent problem during industrial plant operation and it is mandatory to address them as soon as they occur, especially in safety-critical and high cost processes [30]. In general terms, anomaly detection is a widely

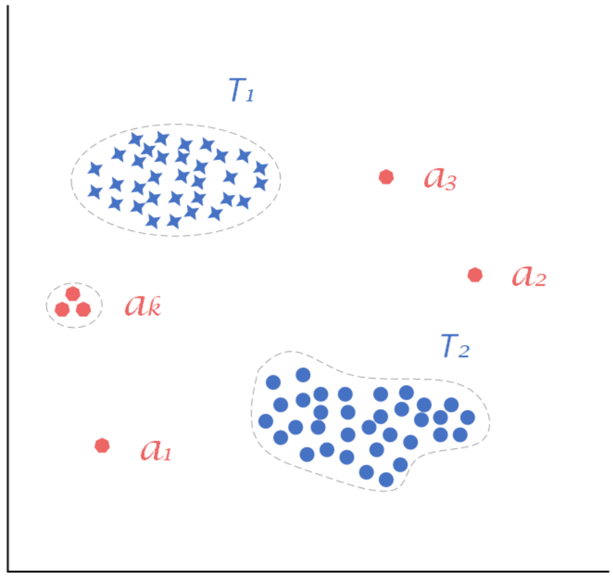


Fig. 1. Anomaly example in a two-dimensional data set

used procedure in many different applications, such as credit card fraud detection [41], fault detection in industrial processes [43, 19], intrusion detection in surveillance systems [42], medical diagnosis and so on [8, 44].

The scientific community has been focusing its attention on the anomaly detection problem for two reasons: firstly, industrial systems are becoming heavily instrumented and are therefore in need of new solutions; secondly, modern computation systems and techniques are more powerful so they can meet those needs [10]. In this scenario, factors like density distribution of the dataset or its geometric location have been taken into account in the anomaly detection process [40]. An example with different points in $\mathbb{R}^2$ is shown in Figure 1. In this case T_1 and T_2, represent expected behavior and a_1, a_2, a_3 and a_k represent clear abnormal function.

1.1 State Of The Art

Depending on the type of information within a dataset, three main cases of anomaly detection are contemplated [22]:

- Case 1: the available dataset is conformed only by normal data or by normal data with a few abnormalities. In this case, the classifier is taught with normal data and if data with different characteristics arrive, they are identified as anomalous. This kind of detection can be considered as semi-supervised. In [21], a virtual sensor for failure detection is implemented in the aircraft

folding/unfolding wings system. Anomaly detection is achieved by modeling the system dynamics and, as a result, the model is capable of detecting abnormal measurements.
- Case 2: the anomaly has to be detected without any previous knowledge of the data. The detection approach in this case is based on unsupervised clustering. Data are distributed in groups, new input data can be classified by comparing them with the data acquired during system operation. This detection assumes that normal data is well separated from the outliers. It provides successful results once the system has a large dataset with good coverage.
- Case 3: initially, normal and abnormal data are available. The data is prelabelled as correct or incorrect data before the implementation of the classifier. In [1] outlier detection has been applied to the field of medicine and has been achieved through labelling of artificially generated anomaly samples.

Depending on the prior knowledge of the data and its application, different approaches can be applied. Some techniques identify data deviations depending on the density estimation of data patterns, such as clustering techniques like DBSCAN [29]. Other techniques establish spatial boundaries of the dataset to detect an anomaly when the data is outside those boundaries, as shown in [14]. The last approach is focused on the reconstruction of data patterns using predictive models. Then, data with high reconstruction error are identified as anomalies [34].

In some applications, especially the critical ones, only correct operation data from the plant is available and failure data is not statistically representative or simply remains unknown. In the cases where most of abnormal functioning situations have not occurred yet, a One Class Classification (OCC) is commonly used [27].

The well-known Support Vector Machine (SVM) algorithm employed in many different applications [38, 4, 6, 26] is frequently used to solve the OCC problem as well as Support Vector Data Description (SVDD) [40]. To solve anomalies issues in different parts of industrial plants, the use of virtual sensors or missing data imputation techniques is very common [5, 15, 25, 7, 20].

This work presents a Case 2 identification problem with a significant modification. This modification consists on using Dimensionality Reduction Techniques (DTRs) to identify the different data boundaries in a two dimensional map. Instead of identifying automatically the data groups using clustering algorithms, the limits are defined by the user.

The approach proposed was tested in a didactic real plant used to control the water level of a tank. The speed of a pump is controlled in order to maintain a constant water level while emptying through an output valve. Different abnormalities were induced during the plant operation and the model proposed is able to recognize this situation.

The outline of this paper is as follows. Section 2 describes briefly the case of study. Then, the model approach is presented. After section 3, the techniques

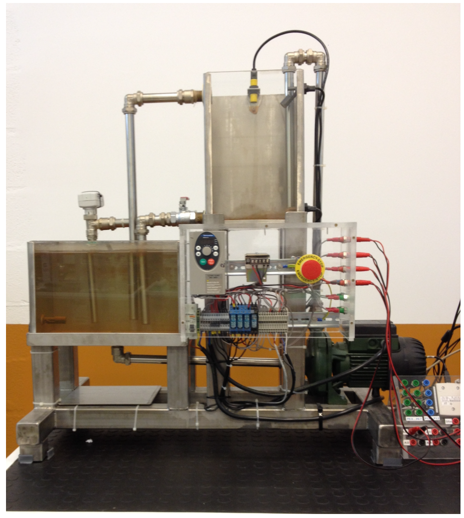


Fig. 2. Laboratory liquid level plant

applied to validate the proposed model are shown. Section 5 explains experiments and results, and finally, conclusions and future works are exposed in section 6.

2 Case of study

This section describes the plant in which the fault detection with the proposed system is tested. Moreover, the general features of the dataset are described.

2.1 Tank Level Control

The main goal of this study is to check the performance of the proposed fault detection system over a real application. The used laboratory plant was built with industrial equipment (see Figure 2).

The scheme of the industrial plant is shown in Figure 3. As stated above, the system was designed to control the level of liquid in a tank. The liquid is initially stored in a different tank placed at a lower level, and it is boosted by a three-phase pump driven by a variable frequency driver (Figure 3, V/F block). The flow rate delivered to the objective tank depends on the pump speed, driven by a three-phase motor (Figure 3, M block). The objective tank, has also two built-in output valves, one of them is a proportional electric and the other one is manual. They are used as a path for returning the fluid back to the storage tank.

The level of fluid is measured using an ultrasonic sensor. The plant structure is movable, however it has a built in mechanism to fix its wheels and avoid any kind of vibration that could distort the measurements taken by the sensor.

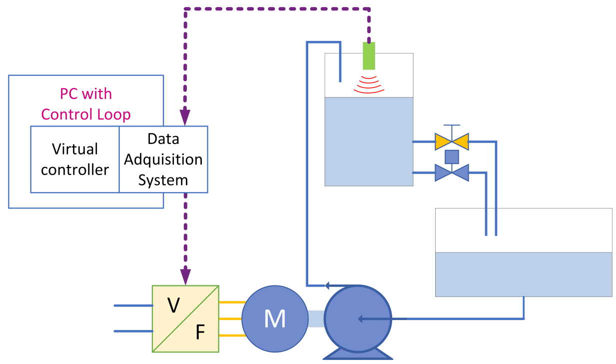


Fig. 3. Laboratory liquid level plant structure

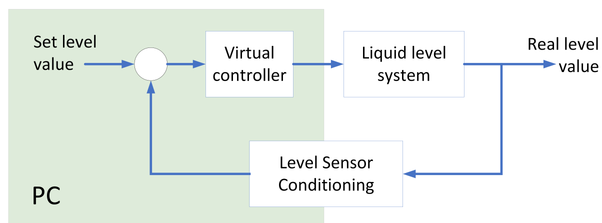


Fig. 4. Laboratory of liquid level control scheme

The scheme shown in Figure 4 represents this single input single output (SISO) system, in which the desired liquid level in a tank is achieved by controlling the pump speed.

2.2 Control System Implementation

The control system is a virtual controller that reads the current state of the plant through a data acquisition card. The set point signal represents the desired liquid level and the process value is the real level measured in the tank. The control signal value is sent by the computer and represents the speed of the centrifugal pump. The control scheme developed using Matlab is shown in Figure 4.

A National Instruments data acquisition card (model USB-6008 12-bit 10 KS /s Multifunction I/O) was used to connect the plant and the computer, and a PID (*Proportional, Integral, Derivative*) control was implemented.

3 Model approach

The aim of this research is to create a model for fault detection in real industrial plants. If the proposed method performs well it is going to detect faults in different devices used at the plant.

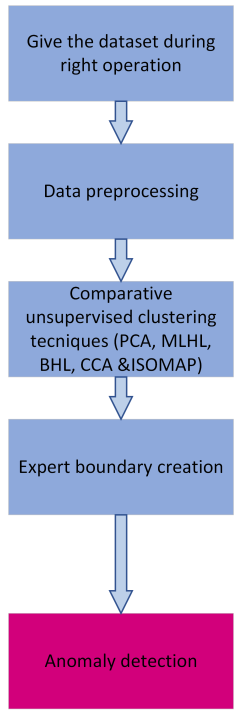


Fig. 5. Process Flowchart

Non-supervised techniques are going to be used because it is necessary to avoid the use of labeled data.

The selected unsupervised techniques are used to allow the visualization of the operation point in a two dimension graph regardless of the number of variables. As the industrial plants usually work in one or a few working point, all the visualized data in an operation point should be near to each other. Hence, data from different working points displayed in the two dimension graph must be clearly separated.

Therefore, the data is projected into two dimension graph and the user could define a contour in the data. With the defined contour, the algorithm used detects if the working point is out of the working area defined.

Automatic selection of a contour is possible, however there is a possibility that the created contour is not going to be capable of detecting failures with the same precision as expert operator would with the manual contour definition. Figure 5 outlines the steps followed.

3.1 Datasets

The datasets used in this research are obtained by registering data from at least 10 minutes of normal operation of the plant and from fault situations corresponding to water leaks through an electrovalve. The sample rate time is one second.

The different experiments are based on an adaptive PID algorithm for water level control. For the implementation of the PID, the RLS (*Recursive Least Squared*) algorithm was used to identify the parameters of the plant transfer function. The transfer function weights allow to calculate the controller parameters according to the actual plant operating point. The adaptive PID algorithm helps to control non-linear systems, adjusting the plant transfer function each time.

Due to the strong non-linearity of the plant under control, an adaptive PID control is implemented [3]. The first step consisted in identifying the laboratory plant on-line as a second order transfer function using the Recursive Least Squares algorithm, according to the Equation 1:

$$G_{plant}(z^{-1}) = \frac{b_0 \cdot z^{-k}}{1 - a_1 \cdot z^{-1} - a_2 \cdot z^{-2}} \tag{1}$$

where:

- b_0 - Open loop gain
- k - System delay
- a_1 - First order coefficient
- a_2 - Second order coefficient

Then, from the transfer function obtained during the identification process, an adaptive PID is self-tuned following the equation 2.

$$G_{controller}(z^{-1}) = \frac{p_0 + p_1 \cdot z^{-1} + p_2 \cdot z^{-2}}{1 - z^{-1}} \tag{2}$$

and:

- $p_0 = \frac{1}{b_0 \cdot T_c^2 \cdot (2 \cdot K + 1)}$
- $p_1 = -a_1 \cdot p_0$
- $p_2 = -a_2 \cdot p_0$

where:

- T_c - Critical period
- K - Critical gain

In this research, two different datasets were created to test the performance of the proposed model with datasets whose complexity, quantity (number of samples) and quality (number of features measured) differ. One dataset had three parameters and the other had five. In both datasets, failures consist in water leaks from the main water tank to the lower tank through a valve.

4 Techniques applied to validate the proposed model

To provide a visual representation of the abnormal behavior of the tested system, this work proposes the application of several Dimensionality Reduction Techniques to detect anomalies by means of visual inspection [36]. The problem of identifying patterns of anomalies that exist across dimensional boundaries in high-dimensional datasets, can be solved by using projection methods. These methods project high dimensional data points onto a lower dimensional space in order to identify "interesting" directions in terms of any specific index or projection.

In this work, Principal Component Analysis (PCA), MLHL (Maximum Likelihood Hebbian Learning), Beta Hebbian Learning Algorithm (BHL), Curvilinear Component Analysis and ISOMAP DRT techniques have been applied to a real control liquid level system to validate our approach.

DRT [16] has been used for the purpose of identifying structure in high-dimensional data. This challenging problem was tackled by projecting the data onto a low dimensional subspace in which we searched for structures by visual inspection using raw human vision. Therefore, the visual presentation is the standard measure widely accepted by the DRT community.

4.1 Principal Component Analysis

Principal Component Analysis (PCA) is a well-known statistical model, introduced in [17], that describes the variation in a set of multivariate data in terms of a set of uncorrelated variables each of which is a linear combination of the original variables. From a geometrical point of view, this method mainly consists in the rotation of the axes of the original coordinate system to a new set of orthogonal axes that are ordered in terms of the amount of variance of the original data they account for.

PCA can be performed by means of neural models such as those described in [31] or [18]. It should be noted that even if it is possible to characterize the data with a few variables, it will not ensure a logical interpretation of these variables.

4.2 Maximum Likelihood Hebbian Learning

Maximum Likelihood Hebbian Learning (MLHL) [12] which is based on Exploration Projection Pursuit (EPP).

The statistical method of EPP [23] was designed to solve the complex problem of identifying structure in high-dimensional data by projecting it onto a lower dimensional subspace in which its structure is searched visually. Therefore, the visual presentation is the standard measure widely accepted by the EPP community.

To that end, an "index" must be defined to measure the varying degrees of interest associated with each projection. Subsequently, the data is transformed by maximizing the index and the associated interest. From a statistical point of view, the most interesting directions are those that are as non-Gaussian as possible.

4.3 Beta Hebbian Learning Algorithm

Beta Hebbian Learning algorithm (BHL) [33], is an EPP technique that belongs to a novel family of learning rules derived from the Probability Density Function (PDF) of the residual based on Beta distribution.

In general, the minimization of the cost function associated with this network, may be thought to make the probability of the residuals more dependent on the PDF of the residuals. Thus, if the probability density function of the residuals is known, this knowledge could be used to determine the optimal cost function. So, the residual (e) is draw from the Beta distribution, $B(\alpha, \beta)$, with the following probability density function (equation 3):

$$p(e) = e^{\alpha-1}(1-e)^{\beta-1} = (x - Wy)^{\alpha-1}(1 - x + Wy)^{\beta-1} \tag{3}$$

Where α and β are the parameters that determine the shape of the PDF curve of the Beta distribution, x is the input of the network, W is the weight vector associated with network neurons and y is the output of the network.

Then, to maximize the likelihood of the data with respect to the weights, the gradient descent is performed using equation 4:

$$\begin{aligned}\frac{\partial p}{\partial W} = (e_j^{\alpha-2}(1-e_j)^{\beta-2}(-(\alpha-1)(1-e_j) + e_j(\beta-1))) = \\ (e_j^{\alpha-2}(1-e_j)^{\beta-2}(1-\alpha+e_j(\alpha+\beta-2)))\end{aligned} \tag{4}$$

In the case of the BHL, by maximizing the likelihood of the residual with respect to the actual distribution, the learning rule is matched to the PDF of the residual. The BHL may also be linked to the standard statistical method of Exploratory Projection Pursuit, as the nature and quantification of the interestingness is in terms of the likelihood of the residuals being under a particular model of the residuals PDF. Therefore, the new neural architecture for BHL is defined as follows:

$$Feedforward : y_i = \sum_{j=1}^{N} W_{ij} x_j, \forall i \tag{5}$$

$$Feedback : e_j = x_j - \sum_{i=1}^{M} W_{ij} y_i \tag{6}$$

$$Weightsupdate : \varDelta W_{ij} = \eta(e_j^{\alpha-2}(1-e_j)^{\beta-2}(1-\alpha+e_j(\alpha+\beta-2)))y_i \tag{7}$$

Where α and β are the parameters that determine the shape of the PDF curve of the Beta distribution, x is the input of the network, W is the weight vector associated with the network neurons, e is the residual and y is the output of the network.

4.4 Curvilinear Component Analysis

Curvilinear Component Analysis (CCA) [13, 11] is a non-linear projection method that preserves distance relationships in both input and output spaces. CCA is a useful method for redundant and non-linear data structure representations and can be used in dimensionality reduction. CCA is useful with highly non-linear data, where PCA or any other linear method fails to give suitable information. CCA improves other methods like Sammon's Mapping [35], although when unfolding a nonlinear structure, Sammon's Mapping cannot reproduce all distances. One way to get round this problem consists in favoring local topology: CCA tries to reproduce short distances first, viewing long distances as secondary. Formally, this reasoning led to the following error function (without normalization) defined in equation 8.

$$E_{CCA} = \sum_{i,j=1}^{N} \left(d^{n}_{i,j} - d^{p}_{i,j}\right)^2 F_\lambda \left(d^{p}_{i,j}\right) \tag{8}$$

In comparison with E_{Sammon}, E_{CCA} has an additional weighting function F depending on d^p_{ij} and on parameter λ. The F factor is a decreasing function of its argument, so it is used to favor local topology preservation.

4.5 ISOMAP Algorithm

ISOMAP nonlinear Dimensionally Reduction Technique [9] attempts to preserve pairwise geodesic (or curvilinear) distance between data points. Geodesic distance is the distance between two points measured over the manifold. ISOMAP defines the geodesic distance as the sum of edge weights along the shortest path between two nodes. The doubly-centered geodesic distance matrix K in ISOMAP is of the form given by equation 9.

$$K = -{}^1\!/_2 H D^2 H \tag{9}$$

Where $D^2 = D^2_{ij}$ means the element wise square of the geodesic distance matrix $D = [D_{ij}]$, and H is the centring matrix, given by equation 10.

$$H = I_n - {}^1\!/_N e_N e_N^T \tag{10}$$

In which $e_N = [1...1]^T \in R^N$ The top N eigenvectors of the geodesic distance matrix represent the coordinates in the new n-dimensional Euclidean space.

5 Experiments and Results

This section describes the experiments performed to validate the proposed system and the obtained results. Both linear (PCA) and nonlinear models (MLHL, BHL, CCA and ISOMAP) have been applied to the previously described datasets (section 3.1), in order to identify the system malfunction states.

5.1 Dataset 1

The first dataset consists of 4871 samples and 3 variables (the 3 parameters of the PID controller) with a total of 21 failure samples.

For the purpose of making a comparative, five different projection models have been applied, whose results are shown in Figure 6, and the total number of mixed samples (fault and normal) are presented in Table 1.

Table 1. Number of mixed samples for each algorithm

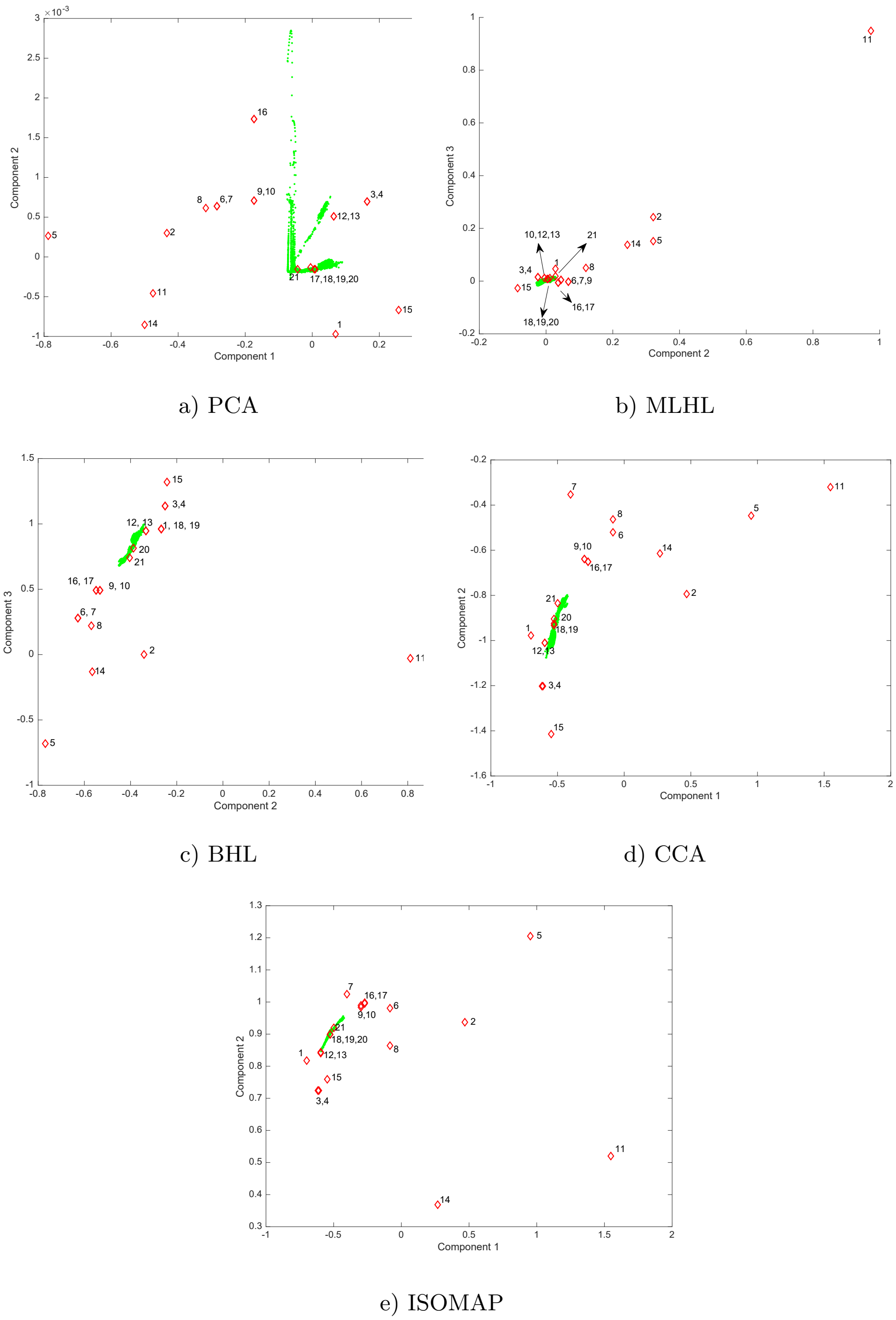
Algorithm	Parameters
PCA	5 samples
MLHL	14 samples
BHL	4 samples
CCA	6 samples
ISOMAP	6 samples

Figure 6 presents the best projection of each algorithm (PCA, MLHL, BHL, CCA and ISOMAP) and the parameters that were used in each projection are presented in Table 2, where *lrate* is the learning rate used during training process, *iters* the number of iterations, *lambda* the initial radius of influence in CCA algorithm, and α and β are the parameters that determine the shape of the PDF curve of the Beta distribution.
These parameters were chosen in an experimental process of trial and error. As parameter selection is a task that is very dependent on the type of dataset used, several initial experiments were conducted with a range of combinations of these parameters. In each figure, anomalies are displayed using red diamond shapes ($\diamond$) and normal samples with green dots ($\cdot$).

Table 2. PCA, MLHL, BHL, CCA and ISOMAP parameters for dataset 1

Algorithm	Parameters
PCA	-
MLHL	iters=1000, lrate=0.01, p=0.5
BHL	iters=5000, lrate=0.01, α=3, β=4
CCA	100 epochs, alpha=0.5 and lambda=1.5152.
ISOMAP	number of neighbours: 5

Results obtained by PCA show normal samples (green dots - Figure 6) are very sparse in the graph so it has been difficult to establish the boundaries for the anomalous data.
Although, MLHL presents the normal samples in a very compact group, several anomalies are over or very near to this group. In the case of CCA and ISOMAP, their projections are very similar and some anomalies are also over the normal



a) PCA

b) MLHL

c) BHL

d) CCA

e) ISOMAP

Fig. 6. PCA, MLHL, BHL, CCA and ISOMAP projections for dataset 1

sample, so it has been difficult to clearly separate them from the cluster generated by normal samples.

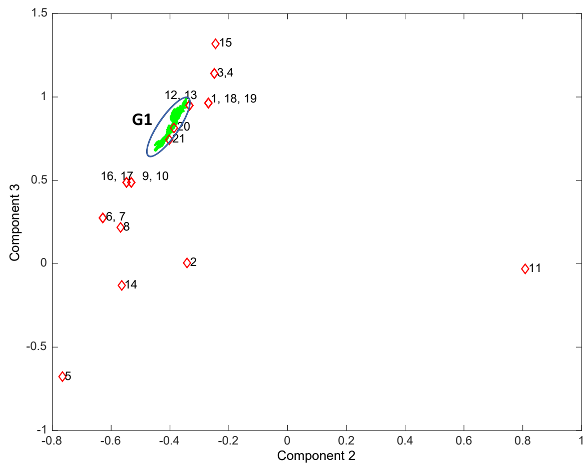


Fig. 7. BHL projection of the dataset 1

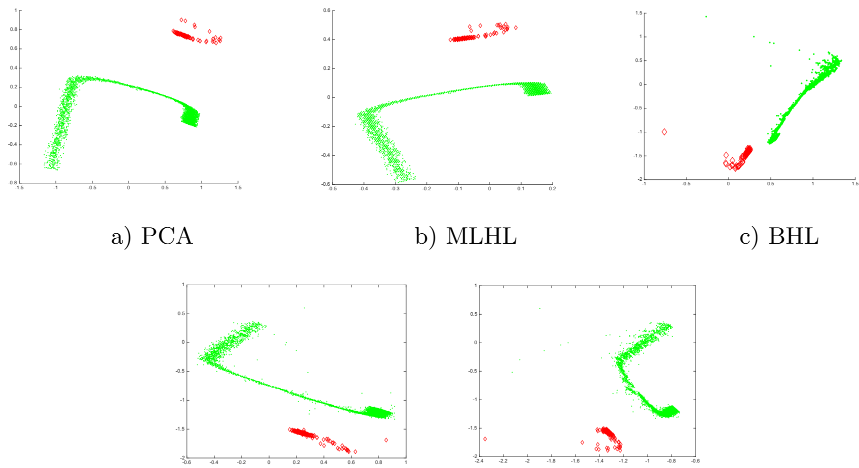
In spite of the fact that none of the 5 methods have been able to separate the anomalies from normal clusters in 100%, it is evident that the projections of the BHL are superior to those of the PCA and the MLHL. BHL has provided a clear visualization of samples which represent anomalies and has been able to present compactly grouped (G1, see Figure 7) samples belonging to correct system operation and has separated the anomalies from this compact group (G1).

5.2 Dataset 2

Dataset 2 consists of 7000 samples and 5 variables (identification parameters of the plant -a_0, a_1, a_2-, water level, process value signal). This dataset provides more and better information than the previous one, as the number of samples is higher and 5 variables are measured instead of 3. In this dataset, the number of anomalous samples is higher; 200 anomalous samples out of the total of 7000.

Figure 8 presents the best projections for each of the 5 tested algorithms, based on the parameters in Table 3.

In this case, the results obtained by all algorithms are similar. All of them are able to clearly distinguish between a normal plant process and an anomalous one. The complexity of this second dataset is smaller than that of dataset 1, as it provides a more informative description of its internal structure, and therefore there are no significant differences between the results obtained by the 5 algorithms.



a) PCA b) MLHL c) BHL

d) CCA e) ISOMAP

Fig. 8. PCA, MLHL, BHL, CCA and ISOMAP projections (component 1 and 2 for all cases) for dataset 2

Table 3. PCA, MLHL, BHL, CCA and ISOMAP parameters for dataset 2

Algorithm	Parameters
PCA	-
MLHL	iters=10000, lrate=0.01, p=0.8
BHL	iters=10000, lrate=0.01, α=3, β=3
CCA	100 epochs, alpha=0.5 and lambda=1.4.
ISOMAP	number of neighbours: 7

6 Conclusions and Future Works

This study has proposed a method for accomplishing fault, anomaly or malfunction detection with unsupervised and projectionist learning techniques. The new approach has been successfully validated in a real laboratory plant where a level control loop was implemented. The correct operation of the plant meant that the output electrovalve had to be completely closed. The feasibility of the proposed method was checked by opening the electrovalve and simulating a water leak.

The obtained results have been very satisfactory in general terms, although their accuracy varied depending on the employed technique and especially on the complexity of the dataset; the higher the complexity the less accurate the results.

The results indicate that the BHL is capable of generating projections in which the dataset is clearly structured, with the lowest number of mixed samples (normal and fault) among the 5 tested algorithms and when complexity of the dataset was high (dataset 1). However, when complexity is low, there are no significant differences in the performance of the 5 algorithms, in all cases they provided projections in which the samples were not mixed.

Finally, by means of the knowledge of an expert user, the boundaries of the normal behavior were established in order to automatically detect future faults in the system.

It is very important to remark that the present contribution approach is particularly suitable for cases where human expertise must be taken into account. The described methodology is going to work well when complemented with expert knowledge in some aspects that cannot be performed automatically (i.e. conditional and predictive maintenance, when for instance a failure in a gear is caused by floor vibration.

After analyzing the results of performed experiments we have identified as an advantage of this approach the fact that it allows a skilled operator to define the contour detection limit for the application of unsupervised techniques. This feature is not available in typical fault detection techniques. Thanks to this contribution, it is possible to include expert knowledge in the fault detection process and, consequently, achieve better performance. Moreover, the approach allows to visually monitor the status of the industrial process.

Thus, the developed tool is going to contribute to maintenance, product quality, efficiency, energy saving or system optimization.

On the basis of the obtained results we can conclude that unsupervised techniques are powerful tools in the detection of anomalies. They allow to monitor and supervise the correct operation of industrial processes, especially when the complexity of the system is high.

In a future work we are going to contemplate the ability to define new contours in real time as new data arrive to the system. Furthermore, we are going to study the possibility of validating additional fault situations using a bigger dataset from the industrial control level plant. Also, it would be interesting to analyze other fault detection techniques and compare their performance with the results obtained in the present work. In addition, different kinds of unsupervised approaches could be used to this end.

Conflict of interest

Authors declare no conflict of interest

References

1. Abe, N., Zadrozny, B., Langford, J.: Outlier detection by active learning. In: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining. pp. 504–509. ACM (2006)
2. Alaiz Moretón, H., Calvo Rolle, J., García, I., Alonso Alvarez, A.: Formalization and practical implementation of a conceptual model for pid controller tuning. Asian Journal of Control **13**(6), 773–784 (2011)
3. Bishop, G., Welch, G., et al.: An introduction to the kalman filter. Proc of SIGGRAPH, Course **8**(27599-3175), 59 (2001)

4. Calvo-Rolle, J.L., Quintian-Pardo, H., Corchado, E., del Carmen Meizoso-López, M., García, R.F.: Simplified method based on an intelligent model to obtain the extinction angle of the current for a single-phase half wave controlled rectifier with resistive and inductive load. Journal of Applied Logic **13**(1), 37–47 (2015)
5. Casteleiro-Roca, J.L., Jove, E., Sánchez-Lasheras, F., Méndez-Pérez, J.A., Calvo-Rolle, J.L., de Cos Juez, F.J.: Power cell soc modelling for intelligent virtual sensor implementation. Journal of Sensors **2017** (2017)
6. Casteleiro-Roca, J.L., Pérez, J.A.M., Piñón-Pazos, A.J., Calvo-Rolle, J.L., Corchado, E.: Modeling the electromyogram (emg) of patients undergoing anesthesia during surgery. In: 10th International Conference on Soft Computing Models in Industrial and Environmental Applications. pp. 273–283. Springer (2015)
7. Casteleiro-Roca, J.L., Quintián, H., Calvo-Rolle, J.L., Corchado, E., del Carmen Meizoso-López, M., Piñón-Pazos, A.: An intelligent fault detection system for a heat pump installation based on a geothermal heat exchanger. Journal of Applied Logic **17**, 36–47 (2016)
8. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM computing surveys (CSUR) **41**(3), 15 (2009)
9. Chang, H., Yeung, D.Y., Xiong, Y.: Super-resolution through neighbor embedding. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. vol. 1, pp. I–I (June 2004). https://doi.org/10.1109/CVPR.2004.1315043
10. Chiang, L.H., Russell, E.L., Braatz, R.D.: Fault detection and diagnosis in industrial systems. Springer Science & Business Media (2000)
11. Cirrincione, G., Hérault, J., Randazzo, V.: The on-line curvilinear component analysis (oncca) for real-time data reduction. In: 2015 International Joint Conference on Neural Networks (IJCNN). pp. 1–8 (July 2015). https://doi.org/10.1109/IJCNN.2015.7280318
12. Corchado, E., MacDonald, D., Fyfe, C.: Maximum and minimum likelihood hebbian learning for exploratory projection pursuit. Data Mining and Knowledge Discovery **8**(3), 203–225 (May 2004). https://doi.org/10.1023/B:DAMI.0000023673.23078.a3, https://doi.org/10.1023/B:DAMI.0000023673.23078.a3
13. Demartines, P., Herault, J.: Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets. IEEE Transactions on Neural Networks **8**(1), 148–154 (Jan 1997). https://doi.org/10.1109/72.554199
14. Fernández-Francos, D., Fontenla-Romero, O., Alonso-Betanzos, A.: One-class convex hull-based algorithm for classification in distributed environments. IEEE Transactions on Systems, Man, and Cybernetics: Systems pp. 1–11 (2018). https://doi.org/10.1109/TSMC.2017.2771341
15. Fernández-Serantes, L.A., Vázquez, R.E., Casteleiro-Roca, J.L., Calvo-Rolle, J.L., Corchado, E.: Hybrid intelligent model to predict the soc of a lfp power cell type. In: International Conference on Hybrid Artificial Intelligence Systems. pp. 561–572. Springer (2014)
16. Friedman, J.H., Tukey, J.W.: A projection pursuit algorithm for exploratory data analysis. IEEE Transactions on computers **100**(9), 881–890 (1974)
17. F.R.S., K.P.: Liii. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **2**(11), 559–572 (1901). https://doi.org/10.1080/14786440109462720, https://doi.org/10.1080/14786440109462720

18. Fyfe, C.: A neural network for pca and beyond. Neural Processing Letters **6**(1), 33–41 (Aug 1997). https://doi.org/10.1023/A:1009606706736, https://doi.org/10.1023/A:1009606706736
19. Garoudja, E., Harrou, F., Sun, Y., Kara, K., Chouder, A., Silvestre, S.: A statistical-based approach for fault detection and diagnosis in a photovoltaic system. In: 2017 6th International Conference on Systems and Control (ICSC). pp. 75–80 (May 2017). https://doi.org/10.1109/ICoSC.2017.7958710
20. Gonzalez-Cava, J.M., Reboso, J.A., Casteleiro-Roca, J.L., Calvo-Rolle, J.L., Méndez Pérez, J.A.: A novel fuzzy algorithm to introduce new variables in the drug supply decision-making process in medicine. Complexity **2018** (2018)
21. Heredia, G., Ollero, A.: Virtual sensor for failure detection, identification and recovery in the transition phase of a morphing aircraft. Sensors **10**(3), 2188–2201 (2010)
22. Hodge, V., Austin, J.: A survey of outlier detection methodologies. Artificial intelligence review **22**(2), 85–126 (2004)
23. Hou, S., Wentzell, P.D.: Re-centered kurtosis as a projection pursuit index for multivariate data analysis. Journal of Chemometrics **28**(5), 370–384 (2014). https://doi.org/10.1002/cem.2568, http://dx.doi.org/10.1002/cem.2568, cEM-13-0108.R1
24. Jove, E., Alaiz-Moretón, H., García-Rodríguez, I., Benavides-Cuellar, C., Casteleiro-Roca, J.L., Calvo-Rolle, J.L.: Pid-its: An intelligent tutoring system for pid tuning learning process. In: International Joint Conference SOCO'17-CISIS'17-ICEUTE'17 León, Spain, September 6–8, 2017, Proceeding. pp. 726–735. Springer (2017)
25. Jove, E., Blanco-Rodríguez, P., Casteleiro-Roca, J.L., Moreno-Arboleda, J., López-Vázquez, J.A., de Cos Juez, F.J., Calvo-Rolle, J.L.: Attempts prediction by missing data imputation in engineering degree. In: International Joint Conference SOCO'17-CISIS'17-ICEUTE'17 León, Spain, September 6–8, 2017, Proceeding. pp. 167–176. Springer (2017)
26. Jove, E., Gonzalez-Cava, J.M., Casteleiro-Roca, J.L., Pérez, J.A.M., Calvo-Rolle, J.L., de Cos Juez, F.J.: An intelligent model to predict ani in patients undergoing general anesthesia. In: International Joint Conference SOCO'17-CISIS'17-ICEUTE'17 León, Spain, September 6–8, 2017, Proceeding. pp. 492–501. Springer (2017)
27. Khan, S.S., Madden, M.G.: A survey of recent trends in one class classification. In: Irish Conference on Artificial Intelligence and Cognitive Science. pp. 188–197. Springer (2009)
28. Kozik, R., Choraś, M.: Protecting the application layer in the public domain with machine learning methods. Logic Journal of the IGPL p. jzy029 (2018). https://doi.org/10.1093/jigpal/jzy029, http://dx.doi.org/10.1093/jigpal/jzy029
29. Lauro, F., Moretti, F., Capozzoli, A., Khan, I., Pizzuti, S., Macas, M., Panzieri, S.: Building fan coil electric consumption analysis with fuzzy approaches for fault detection and diagnosis. Energy Procedia **62**, 411 – 420 (2014). https://doi.org/https://doi.org/10.1016/j.egypro.2014.12.403, http://www.sciencedirect.com/science/article/pii/S1876610214034341, 6th International Conference on Sustainability in Energy and Buildings, SEB-14
30. Miljković, D.: Fault detection methods: A literature survey. In: MIPRO, 2011 proceedings of the 34th international convention. pp. 750–755. IEEE (2011)
31. Oja, E.: Principal components, minor components, and linear neural networks. Neural Networks **5**(6), 927 – 935 (1992).

https://doi.org/https://doi.org/10.1016/S0893-6080(05)80089-9, http://www.sciencedirect.com/science/article/pii/S0893608005800899
32. Parmee, I., Hajela, P.: Optimization in industry. Springer Science & Business Media (2012)
33. Quintián, H., Corchado, E.: Beta hebbian learning as a new method for exploratory projection pursuit. Int. J. Neural Syst. **27**(6), 1–16 (2017). https://doi.org/10.1142/S0129065717500241, https://doi.org/10.1142/S0129065717500241
34. Sakurada, M., Yairi, T.: Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis. p. 4. ACM (2014)
35. Sammon, J.W.: A nonlinear mapping for data structure analysis. IEEE Transactions on Computers **C-18**(5), 401–409 (May 1969). https://doi.org/10.1109/T-C.1969.222678
36. Sánchez, R., Herrero, Á., Corchado, E.: Clustering extension of MOVICAB-IDS to distinguish intrusions in flow-based data. Logic Journal of the IGPL **25**(1), 83–102 (2017). https://doi.org/10.1093/jigpal/jzw047, https://doi.org/10.1093/jigpal/jzw047
37. Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: Niethammer, M., Styner, M., Aylward, S., Zhu, H., Oguz, I., Yap, P.T., Shen, D. (eds.) Information Processing in Medical Imaging. pp. 146–157. Springer International Publishing, Cham (2017)
38. Schlkopf, B., Smola, A.J., Bach, F.: Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. The MIT Press (2018)
39. Stojanovic, N., Dinic, M., Stojanovic, L.: A data-driven approach for multivariate contextualized anomaly detection: Industry use case. In: 2017 IEEE International Conference on Big Data (Big Data). pp. 1560–1569 (Dec 2017). https://doi.org/10.1109/BigData.2017.8258090
40. Tax, D.M.J.: One-class classification: concept-learning in the absence of counterexamples [ph. d. thesis]. Delft University of Technology (2001)
41. Tran, P.H., Tran, K.P., Huong, T.T., Heuchenne, C., HienTran, P., Le, T.M.H.: Real time data-driven approaches for credit card fraud detection. In: Proceedings of the 2018 International Conference on E-Business and Applications. pp. 6–9. ICEBA 2018, ACM, New York, NY, USA (2018). https://doi.org/10.1145/3194188.3194196, http://doi.acm.org/10.1145/3194188.3194196
42. Vega Vega, R., Quintián, H., Calvo-Rolle, J.L., Herrero, A., Corchado, E.: Gaining deep knowledge of android malware families through dimensionality reduction techniques. Logic Journal of the IGPL p. jzy030 (2018). https://doi.org/10.1093/jigpal/jzy030, http://dx.doi.org/10.1093/jigpal/jzy030
43. Yin, S., Gao, H., Qiu, J., Kaynak, O.: Fault detection for nonlinear process with deterministic disturbances: A just-in-time learning based data driven method. IEEE Transactions on Cybernetics **47**(11), 3649–3657 (Nov 2017). https://doi.org/10.1109/TCYB.2016.2574754
44. Zhang, H., Liu, J., Kato, N.: Threshold tuning-based wearable sensor fault detection for reliable medical monitoring using bayesian network model. IEEE Systems Journal **12**(2), 1886–1896 (June 2018). https://doi.org/10.1109/JSYST.2016.2600582