Novel adaptive approach for anomaly detection in nonlinear and time-varying industrial systems

ÁLVARO MICHELENA*, *Department of Industrial Engineering, University of A Coruña, CTC, CITIC, Ferrol, A Coruña 15403, Spain.*

FRANCISCO ZAYAS-GATO**, *Department of Industrial Engineering, University of A Coruña, CTC, CITIC, Ferrol, A Coruña 15403, Spain.*

ESTEBAN JOVE†, *Department of Industrial Engineering, University of A Coruña, CTC, CITIC, Ferrol, A Coruña 15403, Spain.*

JOSÉ-LUIS CASTELEIRO-ROCA††, *Department of Industrial Engineering, University of A Coruña, CTC, CITIC, Ferrol, A Coruña 15403, Spain.*

HÉCTOR QUINTIÁN§, *Department of Industrial Engineering, University of A Coruña, CTC, CITIC, Ferrol, A Coruña 15403, Spain.*

ÓSCAR FONTENLA-ROMERO§§, *Faculty of Computer Science, University of A Coruña, LIDIA, A Coruña 15008, Spain.*

JOSÉ LUIS CALVO-ROLLE¶, *Department of Industrial Engineering, University of A Coruña, CTC, CITIC, Ferrol, A Coruña 15403, Spain.*

Abstract

The present research describes a novel adaptive anomaly detection method to optimize the performance of nonlinear and time-varying systems. The proposal integrates a centroid-based approach with the real-time identification technique Recursive Least Squares. In order to find anomalies, the approach compares the present system dynamics with the average (centroid) of the dynamics found in earlier states for a given setpoint. The system labels the dynamics difference as an anomaly if it rises over a determinate threshold. To validate the proposal, two different datasets obtained from a level control plant operation have

*E-mail: alvaro.michelena@udc.es
**E-mail: f.zayas.gato@udc.es
†E-mail: esteban.jove@udc.es
††E-mail: jose.luis.casteleiro@udc.es
§E-mail: hector.quintian@udc.es
§§E-mail: oscar.fontenla@udc.es
¶E-mail: jlcalvo@udc.es

Vol. 00, No. 0, © The Author(s) 2024. Published by Oxford University Press.

https://doi.org/10.1093/jigpal/jzae070

been used, to which anomalies have been artificially added. The results shown have determined a satisfactory performance of the method, especially in those processes with low noise.

Keywords: Anomaly detection, fault detection, online identification, centroids.

1 Introduction

The ongoing technological advancement is causing the development and implementation of complex industrial processes capable of performing very sophisticated tasks. These processes use many systems made up of a huge number of interconnected components, sensors and actuators that must work in a correct and coordinated way to ensure a good performance of the whole process. This vast number of elements involved in the processes entails working with a significant volume of data, which can sometimes be challenging. Furthermore, this massive volume of information can increase the potential risk of failures or anomalous behavior within the process, leading to a loss of production performance, safety problems or economic losses. For all these reasons, detecting anomalies or failures in this kind of system is critical in ensuring these systems' correct, efficient and safe operation.

Any process or application might have anomalies for various causes, including sensor measurement errors, actuator problems and human mistakes, among others. Therefore, identifying anomalies in an industrial process requires a sophisticated procedure that considers many characteristics and requires prior knowledge of the process's proper operation. Nowadays, failure or anomaly detection systems are widely used in many industrial processes and applications, such as the detection of failures in electric car power cells [16, 18], in the production of biocomponents [11, 12], in the medical field [17, 23] and in cybersecurity [7, 20, 24], and so on [9, 10, 13, 25].

In recent years, the intricacy of modern industrial systems and advancements in computational resources have prompted numerous researchers to focus on developing and optimizing novel anomaly detection techniques and their integration into various applications, processes and systems. Depending on the unique characteristics of the process and application, diverse methods can be employed to identify anomalous behavior [6, 26].

In the field of industrial processes, it is very frequent to implement artificial intelligence techniques based on machine learning to detect possible malfunctions. In this aspect and depending on the available information, supervised, semi-supervised or unsupervised techniques can be used [4, 22].

For example, in [21], supervised machine learning techniques are used to detect anomalies in industrial control systems. The results obtained with this proposal have been very satisfactory. However, despite the good results, supervised techniques require a high-quality labeled dataset containing samples of both normal operation and anomalous situations. The quality of the dataset and its labels is a critical factor in achieving good model performance. In many cases obtaining these datasets is a very complex task that, in most systems, is not really feasible.

Due to the complexity of obtaining quality labeled data sets in many processes, in recent years, much research has been focusing on developing and implementing semi-supervised and unsupervised techniques for anomaly detection. For example, in [15], the authors applied and compared the performance of different semi-supervised (also known as one-class) techniques to detect anomalies in industrial control loops. Modeling the system's dynamics enables anomaly identification; consequently, the model can detect anomalous measurements with high performance. On the other hand, in [14] the authors implemented unsupervised techniques to detect anomalies in industrial processes. The methods tested showed a good performance for the case study approach. However, to effectively detect the anomalous samples, a detailed analysis of the system is required to determine the cluster boundaries obtained according to whether they correspond to abnormal behaviors. Currently, a large number of research deals with the implementation of semi-supervised and unsupervised techniques in different fields [1, 5, 8].

However, despite their good result, all these anomaly detection methods suffer a loss of performance when the systems are affected by temporary variations produced by various reasons such as small component degradations, changes in working conditions (changes in temperature, humidity, etc.) and so on [19, 28]. All these variations directly affect the dynamics of the systems themselves. These changes in dynamics cause the performance of machine learning techniques to drop after a specific time since they have been trained. This loss of performance can lead to the detection of many false alarms, which can cause significant losses for the companies. Therefore, to maintain a good performance of the fault detection system, in many cases, a costly retraining process of the models generated after a specific operation time is needed.

For all these reasons and taking into account the current context of technological progress as well as the great relevance of anomaly detection systems in the industrial field, this paper presents a novel adaptive anomaly detection system for detecting faults in industrial control loops. The method presented consists of a model-based anomaly detection system. For this purpose, the proposal combines a real-time identification algorithm based on the Recursive Least Squares method and the estimation of data volumes from centroids. This new adaptive system is optimized to detect anomalies in nonlinear and time-varying industrial systems without the need to adjust the models after a specific time of operation since it can adjust to temporal variations and operating conditions.

The present paper is structured as follows: after the Introduction, Section 2 describes the case of study. Then, Section 3 presents the proposed anomaly detection method. Section 4 lists the experiments and results. Finally, Section 5 exposes the conclusions and future works.

2 Case of study

This section defines and explains in detail the industrial system used to test and prove the proposal's performance presented in this research. The system's main components are described, including its sensors and actuators, as well as the expected performance. The dataset used is also explained, commenting on the variables acquired as well as other data of relevance.

2.1 Tank control level system

The main goal of this study is to evaluate the effectiveness of the suggested fault detection system in a real scenario.

The level control plant is one of the industrial mockups in the Optimization and Control Laboratory of the Faculty of Engineering of the University of A Coruña. On a small scale, this system emulates the tank filling control system, corresponding to one of the most frequent systems in many industrial processes. Figure 1 shows the real system whereas Figure 2 represents the plant scheme.

This plant includes two tanks located at two different heights. The upper tank is where the filling level is controlled. For this purpose, the system has water stored in the lower tank, connected through a pipe to the centrifugal pump (coupled to a 0.8 HP three-phase motor) to propel the water to the upper tank. Finally, the water is discharged from the upper tank to the storage tank through two pipes, thus closing the water operation cycle.

To control the whole process, the system has different sensors and actuators. On the one hand, to control the three-phase motor coupled to the centrifugal pump, a Schneider Altivar 31 variable

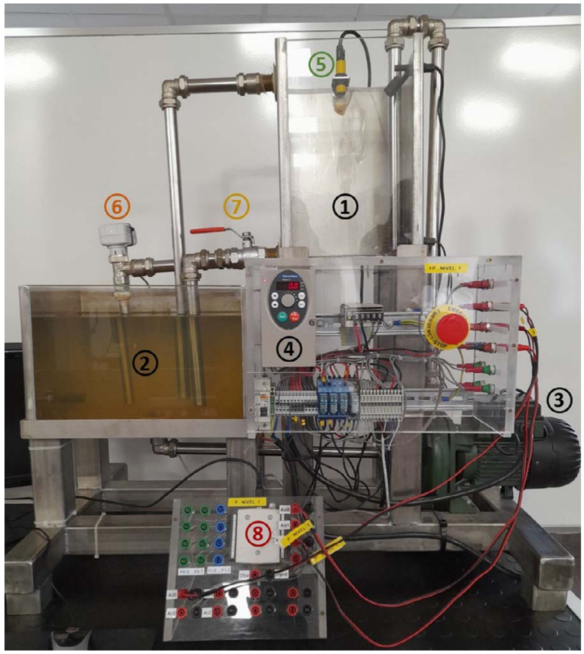

FIGURE 1. Level control plant.

frequency drive is used, which offers a 0/10 V analog input to regulate the motor rotation speed and thus control the water flow rate. On the other hand, a Banner S18UUA ultrasonic sensor is placed at the top of the tank to measure the water level in the filling tank. This sensor provides a continuous 0/10V output signal. The plant also has two valves, one manual and the other electric, controlled by a 0/10V signal to regulate the water discharge flow to the lower tank.

Finally, and because it is an educational model, the control plant has a large number of safety devices, including an emergency stop button that stops the system completely when activated, as well as safety buoys that ensure that the water does not overflow from any of the tanks and that the water pump runs out of water.

2.2 System integration and control implementation

For the control of the system, it is important to note that the plant control signals, both input and output, correspond to analog signals 0/10V. A data acquisition card, model USB-6008 12-bit of National Instruments manufacturer, is used to read and acquire these signals. This card is connected via USB to a computer where all the control logic is implemented. It is important to note that the data acquisition board used has analog inputs in the 0-10 V range. However, it does not have 0-10 V analog outputs since they are 0-5 V, so an auxiliary amplifier circuit is used to convert the output signal to the 0-10 V range so that it can be sent to the variable frequency drive.

Figure 3 shows the control loop that is implemented for its control. A virtual PID controller implemented in Matlab has been used for this research.

Following the scheme shown in Figure 6, the desired fill value is initially set in the computer, and the controller generates a control signal from its programmed logic and the error signal (difference between the expected and real fill value). This control signal is sent to the data acquisition card,

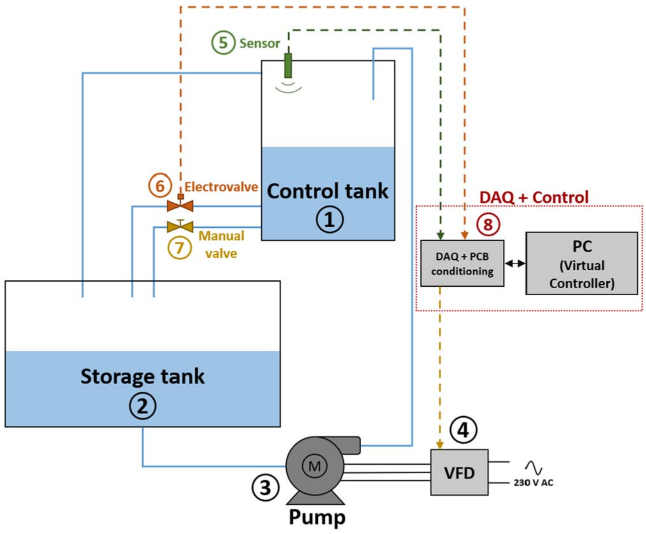

FIGURE 2. Level control plant scheme.

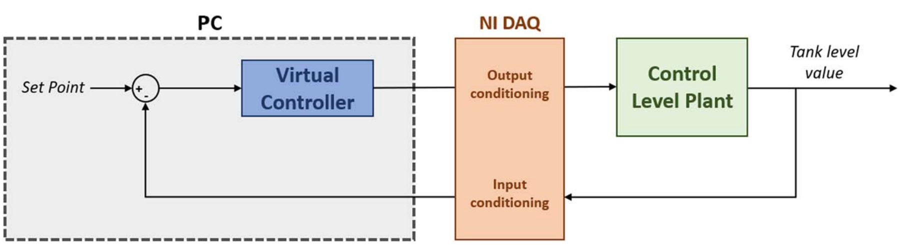

FIGURE 3. Scheme of the control loop implemented.

which is converted to a 0/10V voltage signal that can be interpreted by the variable frequency drive to control the centrifugal pump. The level measured by the sensor is also acquired and conditioned by the data acquisition card to be processed by the executed control algorithm.

2.3 Dataset

The control signal, set point and process value were recorded during 35 minutes of normal plant operation to obtain the dataset used. An adaptive PID controller based on the Dahlin PID [3] was implemented for data collection. For this purpose, the manual and electric piloted valves were fully open. The sampling time was 0.5 seconds, so 4200 samples were recorded for different operating points from 25% to 85% in steps of 10%. The range limitation is necessary since the plant does not perform well for below or above percentage values due to its design. The 10% increment is implemented since the change in system dynamics is not very appreciable for smaller increments.

The real data registered correspond to a correct operation, and this work aims to detect anomalies. Therefore, a total of 30 anomalies have been generated by modifying the process value signal by deviating a random percentage between 4 and 10% of the total filling. Due to the control loop used to control the plant, the system output signal affects the control signal, so this signal is also modified

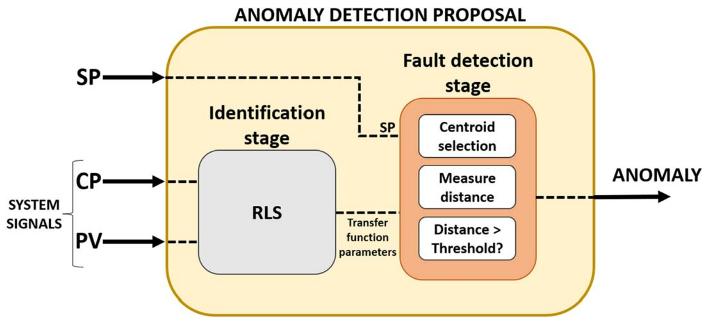

FIGURE 4. Identification system scheme.

in the same samples in an inverse way to the output signal deviation. Since a variation of the control signal affects the system output to a greater extent, this signal is modified the half percentage varied in the process value signal. These anomalies simulate small interferences in the signals or problems of valve obstruction, leaks, etc. Likewise, the sensor output signal has also been modified for 0-100% values to emulate ultrasonic sensor measurement failures.

In order to evaluate the anomaly detection system in processes with signals with a considerable noise ratio, a second dataset has been generated from the previously mentioned. This second dataset differs in that a random variation between $\pm 1\%$ of the level has been added to the system output signal, thus emulating a higher system noise.

3 Methodological approach

This research aims to develop a new adaptive system for anomaly detection in nonlinear and time-varying systems based on on-line identification algorithms. In general, the system uses the RLS on-line identification method to identify the system dynamics at a particular time, and this value is then contrasted with the centroid obtained from the last identification processes determined for that setpoint. An anomaly is detected if the measured distance between the centroid and the new data exceeds a specified threshold. Therefore, this method consists of two main stages: the RLS identification algorithm, to identify the system dynamics in real-time, and the fault detection process based on data clusters and centroids.

The proposal scheme is shown in Figure 4 where SP corresponds to the setpoint value; CP is the control process signal and PV is the process value.

3.1 Online identification stage

For the correct performance of the proposed approach, it is essential to accurately detect changes in the dynamics of the level control plant. Therefore, this identification has to be performed in real-time (online).

Due to its efficiency and simplicity, the Recursive Least Square (RLS) approach is one of the most used methods in the field of online identification. This technique aims to determine the value of transfer function parameters, as described in the θ vector that reduces prediction error and best correlate system input and output signals [2]. Additionally, since it is a simple computational method, a wide range of devices with minimal processing power, such as low-cost boards (Arduino, BeagleBone...) or microcontrollers, can execute this method. As a result, the RLS approach may be incorporated into many other systems and applications.

In general terms, the RLS method works as follows:

1. Firstly, the gain matrix, K defined by Equation 1, is calculated.

$$K_k = \frac{P_{k-1}x_{k-1}}{\lambda + x_{k-1}^T P_{k-1}x_{k-1}} \tag{1}$$

   In Equation 1, x is the regressor vector containing input and output system signals, P corresponds to the covariance matrix defined in Equation 4, and λ is the forgetting factor that will be explained below.
2. Secondly, the prediction error, ε_k, is obtained by means of Equation 2.

$$\varepsilon_k = y_k - x_{k-1}^T \theta_{k-1} \tag{2}$$

3. Then, transfer function parameter vector, θ, is updated following Equation 3.

$$\theta_k = \theta_{k-1} + \varepsilon_k K_k \tag{3}$$

4. Finally, covariance matrix, P is obtained by using Equation 4.

$$P_k = \frac{1}{\lambda}\left(P_{k-1} - K_k x_{k-1}^T P_{k-1}\right) \tag{4}$$

The vector, θ, can be initialized with random numbers. In addition, the regressor vector must be initialized with a close to zero value, whereas covariance matrix initial value is defined by $P = \alpha I$, with α being high integer and I the identity matrix.

Considering Equation 4, it is essential to highlight that RLS is a recursive method, so an exponential forgetting factor, defined as λ, frequently $\lambda \in [0.8, 1]$ [27]), is added to the algorithm. To ensure optimal algorithm performance is crucial to tune this parameter properly by considering its impact on the identification process. A lower value of λ leads to higher sensitivity of the identification algorithm and reduced memory, potentially resulting in errors due to system noise. Conversely, higher values, closer to 1, result in lower sensitivity and more significant memory, which makes the algorithm more robust to system noise, but much slower in detecting changes in process dynamics.

Since the transfer function best fits most systems, the tank fill level plant is defined in this study as a second-order transfer function with time delay, Equation 5:

$$G_p\left(z^{-1}\right) = \frac{b_0 z^{-k}}{1 - a_0 z^{-1} - a_1 z^{-2}} \tag{5}$$

where:

- k is the time delay (for this research k=1)
- b_0 corresponds to the system gain
- a_0 defined the first order coefficient
- a_1 is the second order transfer function coefficient

Therefore, the RLS method is used to obtain the parameters b_0, a_0 and a_1 that best identify the performance of the tank filling plant for a given level.

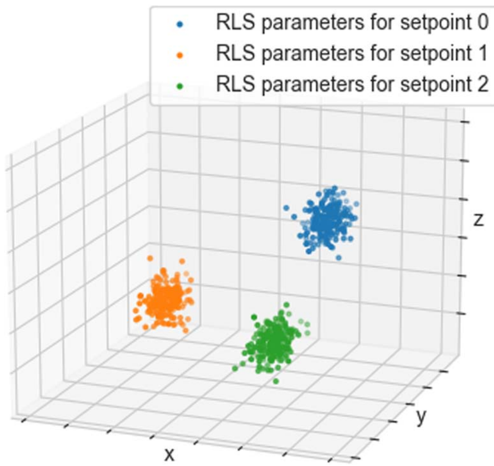

FIGURE 5. Example of identified parameter subsets in nonlinear systems.

3.2 Fault detection stage

Once the RLS algorithm has identified the system dynamics for a certain sample, the transfer function parameters and the desired filling level value are sent to the anomaly detection stage. This stage is divided into three different blocks:

- Data cloud storage and centroid calculation block
- Distance measurement module
- Decision module

3.2.1 Data cloud storage and centroid calculation block This stage stores the last N samples of the transfer function parameters identified for normal operation and computes the centroid (mean value). Since the proposed solution is designed to work with nonlinear systems, it is essential to split the transfer function parameters according to the specified set point, generating small subsets of data that are related to an operating point or a specific range. This is necessary because, in nonlinear systems, the change of operating point causes a change in the dynamics of the process. Figure 5 shows an example of the different data clouds that can be formed depending on the setpoint. This figure represents the different samples identified for a nonlinear system in a three-dimensional space.

For this module, it is essential to determine the number of samples, N, to be stored to calculate the associated centroid. A very large value of stored samples may cause the system to have too much memory, which may not reflect the dynamics of the last samples well. Conversely, a small value of stored samples may cause the system to become very sensitive to noise.

3.2.2 Distance measurement module This module calculates the distance measurement between the calculated centroid associated with the selected working point (selected fill level) and the new sample obtained in the identification process, defined by the transfer function coefficients (b_0, a_0 and a_1). The Euclidean distance, d, defined by Equation 6, is used as the distance measure in the

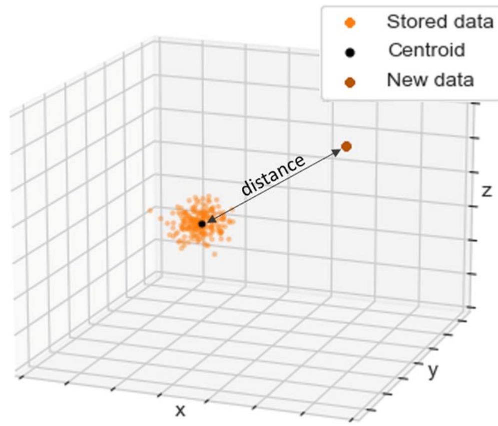

FIGURE 6. Example of Euclidean distance measurement between two three-dimensional points.

proposed system.

$$\|d(a,b))\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2} \tag{6}$$

Finally, it is important to note that the value of the distance between the centroid and the new samples corresponds to a dimensionless value since this distance depends directly on the value of the transfer function coefficients, which are also dimensionless values. Figure 6 shows an example in three-dimensional space of the functioning of this block.

3.2.3 Decision module Once the distance between the centroid and the identified sample has been calculated, the decision module determines whether the new data corresponds to normal operating dynamics or to an anomaly. To this end, this block uses a distance threshold, which is used to determine the volume that is taken up by the system dynamics for a given filling value. Therefore, this system can detect anomalies if the measured distance exceeds the threshold, Equation 7. On the other hand, if the data is considered normal, its value is added to the cluster subset linked with the chosen operating point.

$$\|d(centroid, new \quad data))\| > threshold \rightarrow Anomaly \quad detected \tag{7}$$

This creates an adaptive system that employs iterative centroids to adjust to the slow changes in the dynamics of time-varying systems. Figure 7 shows an example of how the subsets of data are shifted according to the temporal variations of the systems.

Considering this block’s functioning, it is essential to correctly adjust the decision threshold to achieve a good system performance. The value of this threshold will largely depend on the system on which the anomaly detection method is applied and the amount of noise contained in the identified samples. Therefore, in order to select its value, an initial study of the system is required to determine the maximum distances measured between the centroids obtained and the furthest points of their corresponding normal operating samples.

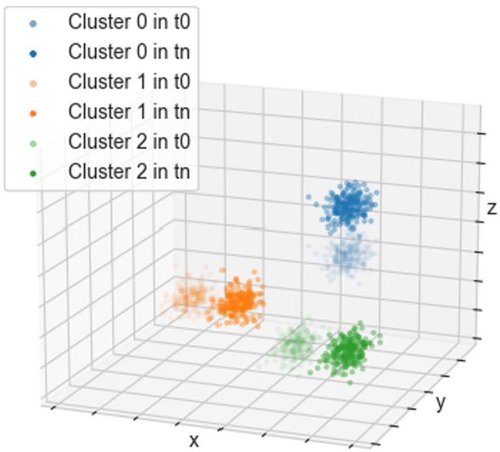

FIGURE 7. Example of system dynamics changes over time.

3.3 Anomaly detection system workflow

For a better understanding of the operation of the developed fault detection system, this section explains its operation flow.

1. Initially, the system waits until the steady state is reached. This is essential because, during the transitory regime, the RLS identification process usually does not provide very accurate results.
2. Once steady state is reached, the fault detection system is activated.
3. It checks if data is stored for the selected operating range. If there is stored data, the centroid of that subset is calculated. If not, the data is stored, and a new iteration is started.
4. The distance between the data cluster's centroid and the RLS's new sample is measured.
5. The distance is compared with the determined threshold. If the distance is greater, an anomaly is detected.
6. If the data is anomalous, the system discards the new sample, and the RLS method working variables are updated to the previous values so that this anomaly does not affect the identification process of future samples.
7. On the other hand, if the data is identified as normal operation, it is stored in the associated data cloud. If the number of data stored in the subset equals the maximum stored data, defined by N, the oldest stored sample is deleted, and the new data is added. This ensures that only the latest samples that best identify the system's dynamics are stored.

4 Experiments and results

Several tests were performed to test and validate the proposed anomaly detection method. First, the optimal value for the centroid data window, forgetting factor and distance threshold had to be determined. The original data set was evaluated without adding anomalies or noise to calculate the threshold value. The system was run to acquire the greatest measured distance between the identified data point and its associated centroid. Using this method, it was found that the greatest distance ranged from 0.9 to 1.1, depending on the operational point chosen.

The forgetting factor and distance threshold were adjusted empirically using the datasets with anomalies to find the best values. Different parameter values were tested by running the system with dataset with and without noise. As it was an anomaly detection task, the metric used to compare results was the f1-score, which calculates the harmonic mean of precision and recall. A score of 1 in f1 indicates that all anomalies were detected and no normal data points were incorrectly labeled as anomalies. Also, precision and recall are measured.

In Table 1, the results obtained for the dataset without noise are shown, whereas Table 2 shows the results for the dataset with noise.

Analyzing the tables, it can be seen that in the case of the dataset without noise, Table 1, a high system performance is obtained for different configurations of its hyperparameters. In this case, all anomalies are detected without any false positive, with a threshold value of 1.1, a forgetting factor of 0.94 and time windows of 10, 20 and 50 samples. On the other hand, the proposal's performance does not reflect good efficiency when the forgetting factor of the RLS is 0.96, since the identification method becomes too robust and the changes in dynamics caused by the anomalies are not detected.

On the other hand, the results obtained with the dataset with noise were worse than those obtained with the data without noise. In this case, a maximum F1 score value of 0.967 is reached for a threshold of 1, storing the last 10 values and a forgetting factor of the RLS method of 0.94. It can also be seen that the use of high values of the forgetting factor, 0.96, significantly worsens the results obtained.

5 Conclusions and future works

This research presents a new adaptive method based on online identification for anomaly detection in nonlinear and time-varying systems. The method has been validated using two different datasets obtained from a real-level control plant, to which anomalies and noise have been artificially added.

For the dataset without noise, the results have been excellent, being able to detect all the anomalies generated and presenting an F1 score value of 1, which indicates that no normal operating data has been classified as anomalous. Using the dataset with noise, the system's performance has been found to be worse, however, with a fine-tuning of the system operating hyperparameters, a good performance can also be obtained, with F1 score values higher than 0.9 in some cases. Despite the good results, tuning the system operating parameters, i.e. the forgetting factor of the RLS identification method, the number of data stored for each operating point and the threshold, can be challenging, as a prior performance analysis is required. The correct setting of these parameters is key to achieving good performance of the proposal.

On the other hand, one of the great advantages of the proposed system is that it is very light and does not require a large computational capacity so it can be implemented and executed in a large number of low-cost devices.

In future work, and for a correct generalization of the proposed method, we will analyze its performance using a more complete dataset that includes operating samples over a longer period of time. In addition, we will try to optimize the system to facilitate the configuration of the operating hyperparameters, such as the forgetting factor, the number of stored samples and the threshold. We will also consider the possibility of developing a method capable of automatically self-adjusting these values as a function of the process. On the other hand, we will analyze the performance of the proposal in other types of industrial processes, such as temperature control loops and systems with

TABLE 1. Results obtained for the dataset without noise

Threshold	Centroid data window (N)	Forgetting factor	Precision	Recall	F1-score
0.9	10	0.92	0.857	1.000	0.923
		0.94	0.938	1.000	0.968
		0.96	0.053	0.933	0.1
	20	0.92	0.769	1.000	0.87
		0.94	0.938	1.000	0.968
		0.96	0.067	0.833	0.124
	50	0.92	0.055	1.000	0.105
		0.94	0.833	1.000	0.909
		0.96	0.053	0.933	0.1
	100	0.92	0.055	1.000	0.104
		0.94	0.057	1.000	0.109
		0.96	0.055	0.967	0.103
1	10	0.92	0.909	1.000	0.952
		0.94	1.000	1.000	1.000
		0.96	0.029	0.933	0.056
	20	0.92	0.882	1.000	0.938
		0.94	0.968	1.000	0.984
		0.96	0.107	0.867	0.19
	50	0.92	0.056	1.000	0.107
		0.94	0.938	1.000	0.968
		0.96	0.107	0.867	0.19
	100	0.92	0.056	1.000	0.106
		0.94	0.857	1.000	0.923
		0.96	0.107	0.867	0.19
1.1	10	0.92	0.968	1.000	0.984
		0.94	1.000	1.000	1.000
		0.96	0.028	0.9	0.054
	20	0.92	0.938	1.000	0.968
		0.94	1.000	1.000	1.000
		0.96	0.107	0.867	0.19
	50	0.92	0.811	1.000	0.896
		0.94	1.000	1.000	1.000
		0.96	0.055	0.967	0.103
	100	0.92	0.732	1.000	0.845
		0.94	0.882	1.000	0.938
		0.96	0.053	0.933	0.1

TABLE 2. Results obtained for the dataset with noise

Threshold	Centroid data window (N)	Forgetting factor	Precision	Recall	F1-score
0.9	10	0.92	0.833	1.000	0.909
		0.94	0.031	0.933	0.059
		0.96	0.107	0.867	0.19
	20	0.92	0.667	1.000	0.8
		0.94	0.055	0.967	0.104
		0.96	0.121	0.933	0.214
	50	0.92	0.099	0.867	0.178
		0.94	0.714	1.000	0.833
		0.96	0.031	0.833	0.06
	100	0.92	0.096	0.867	0.173
		0.94	0.05	0.933	0.096
		0.96	0.062	0.833	0.115
1	10	0.92	0.833	1.000	0.909
		0.94	0.967	0.967	0.967
		0.96	0.069	0.8	0.127
	20	0.92	0.75	1.000	0.857
		0.94	0.444	0.933	0.602
		0.96	0.2	0.9	0.327
	50	0.92	0.714	1.000	0.833
		0.94	0.106	0.867	0.188
		0.96	0.057	0.733	0.105
	100	0.92	0.089	0.933	0.163
		0.94	0.106	0.867	0.188
		0.96	0.068	0.967	0.127
1.1	10	0.92	0.882	1.000	0.938
		0.94	0.068	0.967	0.128
		0.96	0.024	0.733	0.046
	20	0.92	0.109	0.867	0.194
		0.94	0.059	0.9	0.111
		0.96	0.194	0.867	0.317
	50	0.92	0.625	1.000	0.769
		0.94	0.236	0.867	0.371
		0.96	0.045	0.867	0.085
	100	0.92	0.102	0.867	0.182
		0.94	0.789	1.000	0.882
		0.96	0.036	0.833	0.07

multiple inputs and outputs. Finally, we will study the possibility of obtaining graphs that explain the transitory when an anomaly appears.

Acknowledgements

Álvaro Michelena's research was supported by the Spanish Ministry of Universities (https://www.universidades.gob.es/), under the 'Formación de Profesorado Universitario' grant with reference FPU21/00932.

CITIC, as a Research Center of the University System of Galicia, is funded by Consellería de Educación, Universidade e Formación Profesional of the Xunta de Galicia through the European Regional Development Fund (ERDF) and the Secretaría Xeral de Universidades (Ref. ED431G 2019/01).

It should be noted that this initiative is being carried out within the framework of the Recovery, Transformation and Resilience Plan funds, financed by the European Union (Next Generation).

Xunta de Galicia. Grants for the consolidation and structuring of competitive research units, GPC (ED431B 2023/49).

Funding for open access charge: Universidade da Coruña/CISUG.

References

[1] S. Ahmad, A. Lavin, S. Purdy and Z. Agha. Unsupervised real-time anomaly detection for streaming data. *Neurocomputing*, **262**, 134–147, 2017.

[2] K. J. Åström and B. Wittenmark. *Adaptive Control*. Courier Corporation, 2013.

[3] V. Bobál, J. Böhm, J. Fessl and J. Macháček. *Self-Tuning PID Controllers*. Springer, 2005.

[4] J. L. Calvo-Rolle, H. Quintian-Pardo, E. Corchado, M. D. C. Meizoso-López and R. Ferreiro García. 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**, 37–47, 2015.

[5] R. Chalapathy, A. K. Menon and S. Chawla. Anomaly detection using one-class neural networks. arXiv preprint arXiv:1802.06360, 2018.

[6] V. Dutta, M. Pawlicki, R. Kozik and M. Choraś. Unsupervised network traffic anomaly detection with deep autoencoders. *Logic Journal of the IGPL*, **30**, 912–925, 2022.

[7] G.-M. Go, B. Seok-Jun and S.-B. Cho. Insider attack detection in database with deep metric neural network with Monte Carlo sampling. *Logic Journal of the IGPL*, **30**, 979–992, 2022.

[8] M. Goldstein and S. Uchida. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLoS One*, **11**, e0152173, 2016.

[9] J. M. Gonzalez-Cava, R. Arnay, J. A. Mendez-Perez, A. León, M. Martín, J. A. Reboso, E. Jove-Perez and J. L. Calvo-Rolle. Machine learning techniques for computer-based decision systems in the operating theatre: application to analgesia delivery. *Logic Journal of the IGPL*, **29**, 236–250, 2021.

[10] C. Guevara and M. Santos. Intelligent models for movement detection and physical evolution of patients with hip surgery. *Logic Journal of the IGPL*, **29**, 874–888, 2021.

[11] E. Jove, J. Casteleiro-Roca, H. Quintián, J. A. Méndez-Pérez and J. L. Calvo-Rolle. Anomaly detection based on intelligent techniques over a bicomponent production plant used on wind generator blades manufacturing. *Revista Iberoamericana de Automática e Informática industrial*, **17**, 84–93, 2020.

[12] E. Jove, J.-L. Casteleiro-Roca, R. Casado-Vara, H. Quintián, J. A. M. Pérez, M. S. Mohamad and J. Luis Calvo-Rolle. Comparative study of one-class based anomaly detection techniques for a bicomponent mixing machine monitoring. *Cybernetics and Systems*, **51**, 649–667, 2020.

[13] E. Jove, J.-L. Casteleiro-Roca, H. Quintián, J.-A. Méndez-Pérez and J. L. Calvo-Rolle. A new method for anomaly detection based on non-convex boundaries with random two-dimensional projections. *Information Fusion*, **65**, 50–57, 2021.

[14] E. Jove, J.-L. Casteleiro-Roca, H. Quintián, J. A. Méndez-Pérez and J. L. Calvo-Rolle. A fault detection system based on unsupervised techniques for industrial control loops. *Expert Systems*, **36**, e12395, 2019.

[15] E. Jove, J.-L. Casteleiro-Roca, H. Quintián, D. Simić, J.-A. Méndez-Pérez and J. L. Calvo-Rolle. Anomaly detection based on one-class intelligent techniques over a control level plant. *Logic Journal of the IGPL*, **28**, 502–518, 2020.

[16] E. Jove, J.-L. Casteleiro-Roca, H. Quintián, F. Zayas-Gato, G. Vercelli and J. L. Calvo-Rolle. A one-class classifier based on a hybrid topology to detect faults in power cells. *Logic Journal of the IGPL*, **30**, 679–694, 2021.

[17] A. Leira, E. Jove, J. M. Gonzalez-Cava, J.-L. Casteleiro-Roca, H. Quintián, F. Zayas-Gato, S. T. Alvarez, S. Simic, J.-A. Méndez-Pérez, J. Luis. One-class-based intelligent classifier for detecting anomalous situations during the anesthetic process. *Logic Journal of the IGPL*, **11**, **30**, 326–341, 2020.

[18] J. L. Casteleiro-Roca, H. Quintián, J. L. Calvo-Rolle, J.-A. Méndez-Pérez, F. J. Perez-Castelo and E. Corchado. Lithium iron phosphate power cell fault detection system based on hybrid intelligent system. *Logic Journal of the IGPL*, **28**, 71–82, 2020.

[19] I. Machón-González, H. López-Garcáa and J. L. Calvo-Rolle. A hybrid batch som-ng algorithm. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–5, 2010.

[20] Á. Michelena, J. Aveleira-Mata, E. Jove, M. Bayón-Gutiérrez, P. Novais, O. F. Romero, J. L. Calvo-Rolle and H. Aláiz-Moretón. A novel intelligent approach for man-in-the-middle attacks detection over internet of things environments based on message queuing telemetry transport. *Expert Systems*, **41**, 2023.

[21] S. Mokhtari, A. Abbaspour, K. K. Yen and A. Sargolzaei. A machine learning approach for anomaly detection in industrial control systems based on measurement data. *Electronics*, **10**, 407, 2021.

[22] S. Omar, A. Ngadi and H. H. Jebur. Machine learning techniques for anomaly detection: an overview. *International Journal of Computer Applications*, **79**, 33–41, 2013.

[23] G. Pachauri and S. Sharma. Anomaly detection in medical wireless sensor networks using machine learning algorithms. *Procedia Computer Science*, **70**, 325–333, 2015.

[24] H. Quintián, E. Jove, J.-L. Casteleiro-Roca, D. Urda, Á. Arroyo, J. L. Calvo-Rolle, Á. Herrero and E. Corchado. Advanced visualization of intrusions in flows by means of beta-hebbian learning. *Logic Journal of the IGPL*, **30**, 1056–1073, 2022.

[25] S. Simić, E. Corchado, D. Simić, J. Dordević and S. D. Simić. A novel fuzzy metaheuristic approach in nurse rerostering problem. *Logic Journal of the IGPL*, **28**, 583–595, 2020.

[26] F. Zayas-Gato, Á. Michelena, H. Quintián, E. Jove, J.-L. Casteleiro-Roca, P. Leitão and J. L. Calvo-Rolle. A novel method for anomaly detection using beta hebbian learning and principal component analysis. *Logic Journal of the IGPL*, **31**, 390–399, 2022.

[27] H. Zhang, S.-J. Gong and Z.-Z. Dong. On-line parameter identification of induction motor based on rls algorithm. In *2013 International Conference on Electrical Machines and Systems (ICEMS)*, pp. 2132–2137. IEEE, 2013.

[28] H. Zhou, Z. Lei, E. Zio, G. Wen, Y. S. Zimin Liu and X. Chen. Conditional feature disentanglement learning for anomaly detection in machines operating under time-varying conditions. *Mechanical Systems and Signal Processing*, **191**, 110139, 2023.

Received 20 May 2022