Fault Detection in Wastewater Treatment Plants: Application of Autoencoders with Streaming Data

Rodrigo Salles[1,2],
Jérôme Mendes[2], Rita P. Ribeiro[1,3], and João Gama[3,4]

[1] Faculty of Sciences, University of Porto, 4169-007 Porto, Portugal.
[2] Institute of Systems and Robotics, University of Coimbra, 3030-290 Coimbra, Portugal
[3] INESC TEC, 4200-465 Porto, Portugal.
[4] Faculty of Economics, University of Porto, 4200-464 Porto. Portugal

Abstract. Water is a fundamental human resource and its scarcity is reflected in social, economic and environmental problems. Water used in human activities must be treated before reusing or returning to nature. This treatment takes place in wastewater treatment plants (WWTPs), which need to perform their functions with high quality, low cost, and reduced environmental impact. This paper aims to identify failures in real-time, using streaming data to provide the necessary preventive actions to minimize damage to WWTPs, heavy fines and, ultimately, environmental hazards. Convolutional and Long short-term memory (LSTM) autoencoders (AEs) were used to identify failures in the functioning of the dissolved oxygen sensor used in WWTPs. Five faults were considered (drift, bias, precision degradation, spike and stuck) in three different scenarios with variations in the appearance order, intensity and duration of the faults. The best performance, considering different model configurations, was achieved by Convolutional-AE.

Keywords: Wastewater Treatment Plant · Fault Detection · Autoencoder · BSM2

1 Introduction

Water is a strategic and fundamental resource for human beings. Activities carried out in the industry, agriculture, and services depend directly on access to water resources. And access to water is limited. Most of the water on the planet is in the seas and oceans (97%) [1]. There is only 3% of fresh water, but more than two-thirds is frozen in glaciers and polar ice [2]. The small fraction of fresh water remaining needs to serve more and more people. It is estimated that two-thirds of the world's population, 4 billion people, face water scarcity conditions at least one month a year, and approximately 500 million people live with water shortages throughout the year [3]. Water is a strategic resource and must be managed consciously. The used water must be treated so that we can reused it. Wastewater may contain pollutants that pose risks to the environment and consequently to humans and need to be treated in appropriate places. Wastewater treatment plants (WWTPs) are structures that accelerate the treatment process in nature. The water used in human activities is sent to the WWTPs, which carry out the treatment in several stages, using chemical and physical processes. These structures are present in various parts of the world. In the United States of America, there are more than 16000 public administration WWTPs. Europe has more than 24000 treatment units. Brazil, Mexico, and China have 2820, 2540, and 1486 WWTPs, respectively [4].

WWTPs are important in dealing with water scarcity, but they must carry out their functions sustainably, with high quality and low cost. A monitoring system is needed to provide information from all stages of the treatment process so the necessary actions can be taken at the right time. With technological advances, new techniques were used to improve the functioning of WWTPs. The massive use of sensors in the monitoring of treatment plants generated a large amount of information and enabled the use of new control and optimization techniques. But the use of sensors also poses new problems. The actions taken by the control and optimization methods depend on the quality of the information provided by the sensors, and the quality of this information must be ensured. It is common for sensors to be exposed to extreme conditions at the monitoring site, as for example, temperature, vibrations, dust, chemical reagents, etc. It is of great importance that failures in these sensors are indicated as soon as possible, where undetected failures can represent damage to the structure of WWTPs, heavy fines, and environmental damage.

One of the main wastewater treatment phases occurs in the biological reactor. This reactor is composed of anoxic and oxygenated tanks, and it is the site of action of microorganisms that have the function of removing dangerous pollutants from wastewater. Oxygenated tanks need to maintain minimum oxygen levels. Lack of oxygen can result in the death of microorganisms, and excess oxygen represents a waste of energy spent on pumping. Considering that aeration is the most energy-intensive operation in wastewater treatment, amounting to 45–75% of plant energy costs [5], that of all the energy consumed in the world, approximately 3% is consumed in WWTPs [6], and that the energy spent on pumping depends on the information of the dissolved oxygen (DO) levels provided by the sensor, an efficient fault detection system is essential for the DO sensor. The main objective of this paper is to use autoencoder (AE) models to detect DO sensor failures in WWTPs. Early failure detection is important as it allows the necessary actions to be taken, benefiting higher safety, economy and quality of wastewater treatment. This work aims to assess the potential of AE in detecting failures in DO sensors, in WWTPs, in real-time, with streaming data. The simulator Benchmark Simulation Model nº2 (BSM2), which reproduces all phases of the treatment performed in a WWTP, was used to test the fault detection methodologies on the DO sensor in the biological reactor. This work analyzes the strengths and weaknesses of the Convolution-AE and LSTM-AE models, used to detect failures on DO sensors presented in WWTPs. The models are evaluated for the detection of five types of failures: bias, drift, precision degradation, spike and stuck. These faults were injected into the dataset, obtained with the help of BSM2, in three scenarios, with changes in the order of appearance, duration and intensity of the faults.

The rest of the paper is organized as follows: Section 2 brings a review of literature related to fault detection in WWTPs. Section 3 describes the simulator used and the case studies. Section 4 presents the structure of AE-based methodologies used to identify faults on DO sensors. Section 5 presents the experimental results, and compares the performance of Convolution-AE and LSTM-AE models in identifying the failures. Finally, in Section 6, the conclusion of the work is elaborated.

2 Related Work

Many works have already been proposed with the objective of detecting failures in sensors in WWTPs. The works can be divided into two large groups: failure detection using statistical methods, and failure detection using machine learning techniques. In [7], the use of artificial neural networks (ANNs) is proposed to identify six types of faults, one of which is the DO concentration sensor. The results proved a good ability of the ANN to recognize the faults, identifying 97% of case study failures. In [8], the authors use a Long short-term memory (LSTM) networks to identify collective failures in the sensors. The results obtained by the LSTM were compared to the results of the autoregressive integrated moving average (ARIMA), principal component analysis (PCA) and support-vector machines (SVM) models, and achieved the best performance, with a fault detection rate of over 92%. In [9], a radial basis function (RBF) neural network is used to identify faults in DO sensors by calculating the error limits. The proposed method obtained 0% false alarm, and a delay of 0.22 days in detection. In fault detection, unsupervised ANN can be trained to model a process by estimating the values of inputs and comparing the estimation to the actual values, also known as autoencoder. In [10], the authors used an AE, and the proposed model was used for detection of abrupt changes and drift in the sensor signal. The results showed that AE is capable of detecting sensor faults with good accuracy under different scenarios. In [11], a variational AE is used for fault detection. The proposed model takes into account the temporal evolution of the treatment process. The slow feature variational AE (SFAVAE) model is used to monitor processes and tries to identify faults such as sludge expansion fault and small magnitude variable step. Among the statistical methods the most used is the PCA. The PCA has many applications in WWTP, from direct fault detection [12] to data reconstruction [13]. In [14], the Incremental Principal Component Analysis (IPCA) method was used to identify several types of failures in WWTPs, one of them being failures in the DO sensor. The failures were injected into the dataset, and IPCA proved to be able to detect the failures and isolate the variable that originated the failure, with false alarm rate and missed detection rate of 0.07% and 18.53% respectively. A probabilistic PCA approach in process monitoring and fault diagnosis with application in WWTP is proposed in [15]. The probabilistic PCA is compared to PCA, PPCA (probabilistic interpretation of the PCA), GPLVM (version of the PPCA for nonlinear situations) and Bayesian GPLVM (uses the Bayesian theory for training). The GPLVM and GPLVM models showed better performance in detecting failures in relation to the other models analyzed in the paper, in relation to the considered metrics. The major drawback of PCA for WWTP, is the assumption that process variables are linearly related to each other [16]. In the present work, models based on AE will be used. The case studies in which the models will be evaluated will be explained in the following section.

3 Case Studies

The water resulting from human activities, which carry pollutants, cannot be returned to the environment without undergoing treatment. This treatment occurs in WWTPs and is done in several stages: primary treatment (removes floatable and settleable solids), secondary treatment (secondary decantation and activated sludge), tertiary treatment (reuse of treated water), and sludge treatment (mechanical and biological treatments) [17]. Before implementing and evaluating new techniques in real treatment plants, simulators are commonly used. A widely used simulator of WWTPs is BSM2 [14,18]. Section 3.1 provides a brief description of the BSM2, and Section 3.2 presents the case studies with the description of the injected failures.

3.1 Benchmark Simulation Model No 2 - BSM2

The BSM2 is a simulation environment in which the plant layout, the simulation model, influence loads, test procedures and evaluation criteria are defined. For each of these items, compromises were pursued to combine plainness with realism and accepted standards [19]. BSM2 allows the implementation of several techniques and the manipulation of many parameters related to WWTPs [20, 21]. The influent dynamics are defined for 609 days, which takes into account rainfall effect, and temperature [19], with the data sampled every 15 minutes. The structure of BSM2 can be seen in the Fig. 1. The BSM2 was used to test the failure detection methods used in this work.

3.2 Faults in Dissolved Oxygen Sensor

BSM2 simulates the various stages of WWTPs. The biological reactor or activated sludge reactor can be seen in Fig. 1. Fig. 2 shows the biological reactor in more detail. It consists of five tanks, the first two being anoxic and the next three oxygenated. In oxygenated tanks, the DO level must be kept at 2 [mg/L], by a proportional-integral (PI) controller, which receives the DO values measured in Tank 4, compares it with the reference value, and drives the pumps responsible for maintaining the DO at the correct levels.

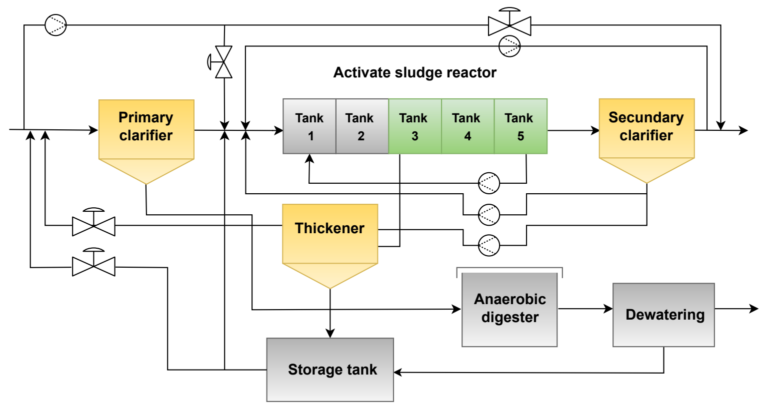


Fig. 1: Layout of Benchmark Simulation Model No 2 - BSM2.

The aeration system depends on the value measured by the DO sensor. A failure in this signal leads the system to malfunction. Through BSM2, we injected several faults. The objective is to identify the faults as early as possible. Also, in Fig. 2, it can be seen that the DO sensor performs its measurement in Tank 4 and it is used by the PI controller, and that the AE, represented in yellow, was positioned between the DO sensor and the PI controller, and has the function of detecting anomalous behavior in the measurement signal coming from the DO sensor.

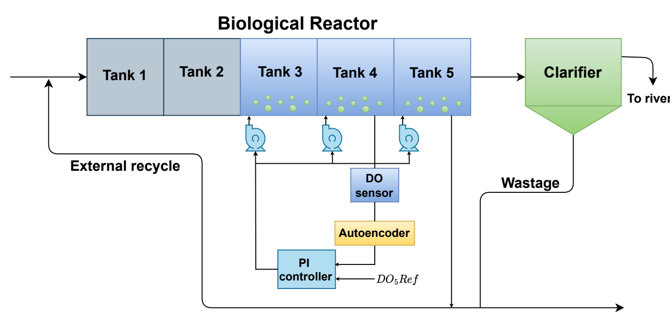


Fig. 2: Biological reactor details and case studies framework.

To evaluate the performance of the fault detection system, some faults were injected into the signal from the sensor. Deviations from expected behavior in the sensor output are considered faults, and they are classified according to the deviation from normal behavior. Let $s(t) = h(t) + \eta$ be the expected output of a sensor without the presence of faults, where $h(t)$ is the output of the sensor at time t, and $\eta \sim N(0, \delta_n^2)$ is noise, and δ_n^2 is the noise variance [22]. The failures considered in the present work are presented below [23,24]. They are common failures in sensors caused by corrosion, calibration errors, presence of noise and physical damage presented on the WWTPs.

Drift fault When the sensor output increases at a constant rate, this type of fault is called as drift fault. A drift fault can be defined as

$$\begin{aligned} s(t) &= h(t) + \eta + b(t), \\ b(t) &= b(t-1) + v, \end{aligned} \tag{1}$$

where $b(t)$ is the bias added to the signal at time t, and v is a constant. The $s(t)$ value increases linearly from the normal value over time.

Bias fault In a bias fault, a constant value v is added to the sensor output and, as a consequence, a shift from the normal value is observed:

$$s(t) = h(t) + \eta + v. \tag{2}$$

Precision degradation (PD) fault This type of fault adds noise with a zero mean and high variance (δ_v^2) to the output of a sensor:

$$s(t) = h(t) + \eta + v,\ v \sim N(0, \delta_v^2),\ \delta_v^2 \gg \delta_n^2. \tag{3}$$

Spike fault In spike faults, large amplitude peaks are observed at constant time intervals at the sensor output:

$$\begin{aligned} s(t) &= h(t) + \eta + v(t), \\ \forall t &= v \times \tau,\ h(t) + \eta, \\ \text{otherwise},\ v &= \{1, 2, ...\},\ \tau \geq 2, \end{aligned} \tag{4}$$

where τ is the interval in which the spikes occur in the sensor output.

Stuck fault It is a complete failure, with the sensor output being locked at a fixed value v:

$$s(t) = v. \tag{5}$$

4 Fault Detection using Autoencoders

WWTPs exhibit marked nonlinear characteristics due to biochemical reactions and nitrification processes [25]. Thus, traditional statistical methods present difficulties in correctly identifying changes in the variables involved in the wastewater treatment process. In order to have a good performance in identifying the changes that may occur in WWTPs, the method used must be able to deal with non-linearities. The AE present, among other characteristics, the ability to deal with non-linear processes.

AE is an unsupervised machine learning algorithm that aims to reconstruct its input signal. The generic AE model consists of three parts: encoder (responsible for reducing the dimensionality of the data), code (reduced representation of the encoder input data), and decoder (responsible for expanding the dimensionality represented in the code, and reconstructing the input signal). A representation of the AE can be seen in the Fig. 3.

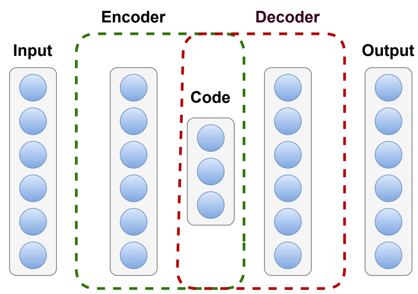


Fig. 3: Illustration of a generic AE model.

AE can have a simple structure, with only the code as the hidden layer, or can have several hidden layers. The representation of the input data, made by the code layer, can be classified as undercomplete or overcomplete. In undercomplete representation, the dimension of the representation of the input data by the code is smaller than the dimension of the input data, which forces the model to learn the most important characteristics of the input data. If the dimension of the code's input data representation is equal to the input dimension, the overcomplete representation, the model will just copy the input signal to the output without learning the most important characteristics of the input signal. The AE can be implemented as fully connected, convolution based or recurrent based units [26]. In this work, two models of AE will be used: LSTM-AE and Convolutional-AE. These models will be described in the next subsections.

4.1 LSTM Autoencoder

LSTM is a recurrent neural network that takes into account the historical context of events to make its predictions, with the help of memory cells. In an LSTM cell there are input, forget, memory and output gates.

In [27], the LSTM-AE is described as an extension to RNN based AE for learning the representation of time series sequential data. In this model, encoder and decoder are built using LSTM. The encoder LSTM receives a sequence of vectors that represents the signal from the DO sensor, and the decoder has the function of recreating the target sequence of input vectors in the reverse order. This is the model used in the present work. The generic structure of a LSTM-AE can be seen in Fig. 4

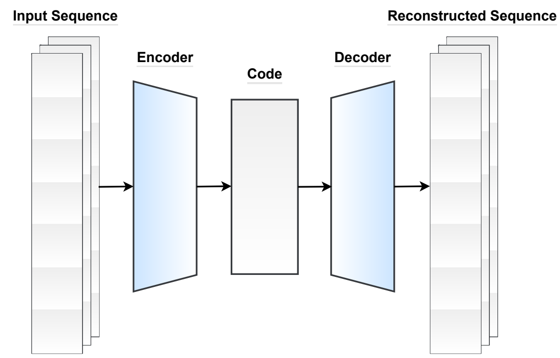


Fig. 4: Generic structure of the LSTM-AE.

4.2 Convolutional Autoencoder

Fully connected AEs ignore the spatial structure of the input signal, and this spatial structure can represent important information for the final reconstruction. To solve this problem, in [28] is proposed a model known as Convolutional-AE. Instead of using fully connected layers, Convolutional-AE use convolutional operators.

The Convolutional-AE is trained to reproduce the input signal from the DO sensor to the output layer. The signal passes through the encoder, composed of a convolution layer, which reduces the dimension of the representation of the input signal. In the decoder, composed of deconvolution layers, the compressed signal is reconstructed to obtain the original input signal, the DO sensor signal. The generic structure of a Convolutional-AE can be seen in Fig. 5.

The LSTM-AE and Convolutional-AE will be used to identify failures in these case studies. The experimental results will be described in Section 5.

5 Experimental Results

This section presents the results obtained by Convolutional-AE and LSTM-AE presented in Section 4, in detecting the faults described in Section 3.2. This section also describes the characteristics of injected faults and the evaluation metrics.

The dataset from the DO sensor, was separated into two equal parts, each part being equivalent to 100 days (9600 samples), keeping the temporal order. The first part was used to inject the faults described in Table 1 and represented

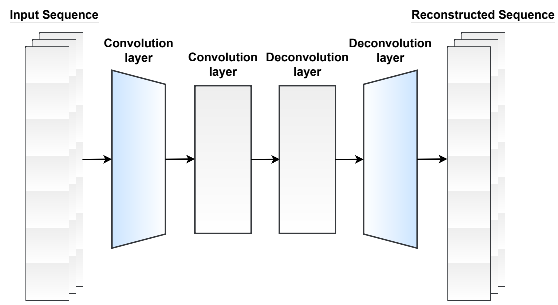


Fig. 5: Generic structure a Convolutional-AE.

in the graphs of Figure 6. From the second part, without failures, 70% of the data was used for training the Convolutional-AE and LSTM-AE, and the remaining 30% was used for evaluating the models, according to the input signal reconstruction error. In order to obtain the best model for each algorithm, a grid search were performed, with the evaluation of the combination of hyperparameters. All the work was developed in Python programming language, version 3.7, with the help of the Keras neural network package, version 2.8.0. The following combinations of hyperparameters were analyzed:

– Convolutional-AE: Epochs $= [10, 20, 30, 40, 50]$; Batch size $= [32, 64, 128]$; AE layout : $[16, 32, 64, 128]$.
– LSTM-AE: Epochs $= [10, 20, 30, 40, 50]$; Batch size $= [32, 64, 128]$; LSTM cells (AE layout) $= [16, 32, 64, 128]$.

The best model was the one with the lowest mean absolute error (MAE):

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |s_t - \hat{s}_t|, \tag{6}$$

where s_t and $\hat{s}_t$ are the real and estimated DO values at the instant of time t, and n is the number of samples used to validation.

The best models found, according to the MAE, were:

– Convolutional-AE: Epochs = 20; Batch size = 128; AE layout = [32, 16, 16, 32]
– LSTM-AE: Epochs = 20; Batch size = 64; LSTM cells (AE layout) = [128, 64, 32, 32, 64, 128]

Table 1 presents the types of faults injected into the dataset, as well as their duration. To better evaluate the performance of the AE three scenarios were considered with variations in the order of appearance, duration and intensity of failures. Figure 6 depicts the signal from the DO sensor, after the faults described in Table 1. The objective of Convolutional-AE and LSTM-AE is to identify the faults that can be seen in Figure 6. The process of training AEs and choosing the best models will be described below.

Table 1: Three scenarios of faults injected in the signal obtained by the DO sensor: drift (Eq. 1), bias (Eq. 2), PD (Eq. 3), spike (Eq. 4) and stuck (Eq. 5). The faults in the scenarios have different order of appearance, duration and intensity.

Fault	**Start [day]**	**Duration [hours]**
Drift	10	120
Bias	30	120
PD	50	120
Spike	70, 72, 74, 76, 78	0.25
Stuck	90	120

(a) Scenario I

Fault	**Start [day]**	**Duration [hours]**
Drift	40	72
Bias	92	48
PD	18	96
Spike	60, 62, 64, 66, 68	0.25
Stuck	30	72

(b) Scenario II

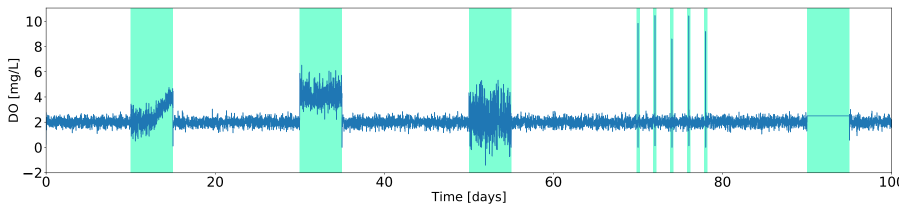
Fault	**Start [day]**	**Duration [hours]**
Drift	58	96
Bias	45	72
PD	90	72
Spike	26, 30, 31, 33, 38	0.25
Stuck	15	48

(c) Scenario III

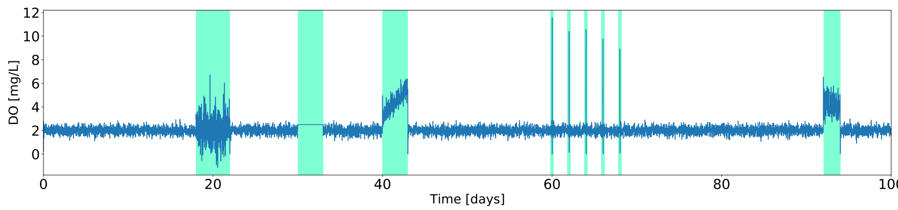
The purpose of the AE is to reconstruct the input signal. During its training, with the DO sensor data, the maximum value for the reconstruction error is adopted as a threshold. In tests, a failure is identified if the difference between the real and estimated DO values is greater than the determined threshold.

The fault identification methods were evaluated as follows:

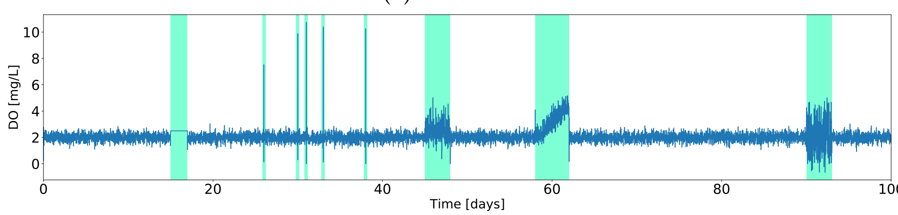
- if a sample is identified as faulty, within the fault duration period, it is classified as true positive (TP);
- if a sample is identified as a failure, outside the fault duration period, is classified as a false positive (FP);
- if a sample, within the failure duration time, is classified as normal, we have a false negative (FN);
- if a sample outside the fault duration period is identified as normal, it is classified as true negative (TN).



(a) Scenario I



(b) Scenario II



(c) Scenario III

Fig. 6: Faults implemented in the signal obtained from the DO sensor. Variations in the order of appearance, duration and intensity of faults.

The evaluation metrics used was for this study are TP rate (TP_r), FP rate (FP_r) and FN rate (FN_r) - cf. Eq.7-9. The use of these metrics makes it possible to assess the reliability of the implemented error detection system.

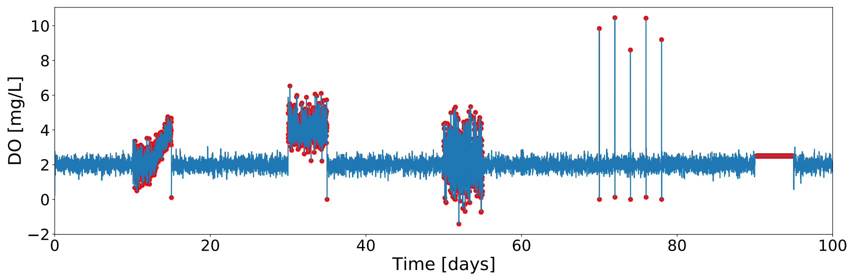
$$TP_r = TP/(TP + FN) \tag{7}$$

$$FP_r = FP/(FP + TN) \tag{8}$$

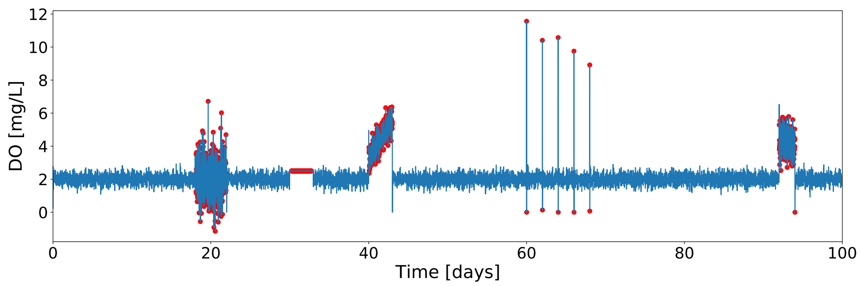
$$FN_r = FN/(FN + TP) \tag{9}$$

The results obtained by the models are graphically represented in Figures 7–8. Table 2 presents the performance of LSTM-AE and Convolutional-AE in identifying the present faults, where the values represent the arithmetic average of the algorithms' performance, for each fault implemented, in the three scenarios.

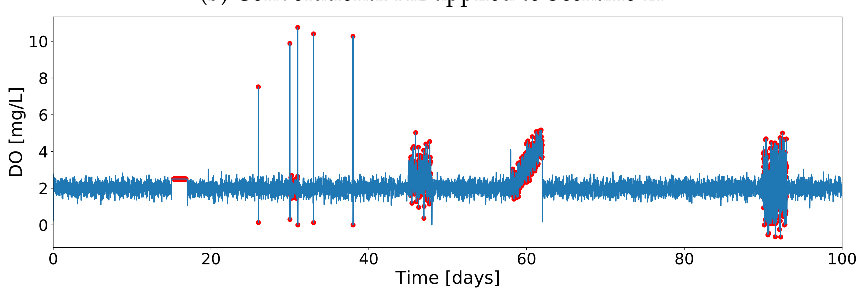
It is noticed that the two models are efficient in identifying the failures according to the evaluation metrics. The bias, drift and PD faults were correctly identified by Convolutional-AE, with TP_r of 95.6%, 95.15% and 97.8%, respectively. The LSTM-AE model presented a little lower performance with TP_r of



(a) Convolutional-AE applied to Scenario I.



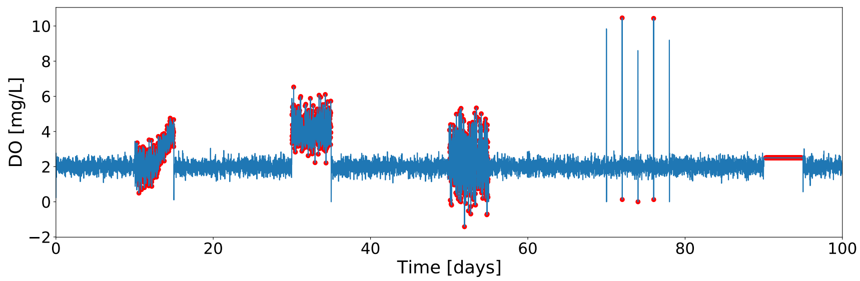
(b) Convolutional-AE applied to Scenario II.



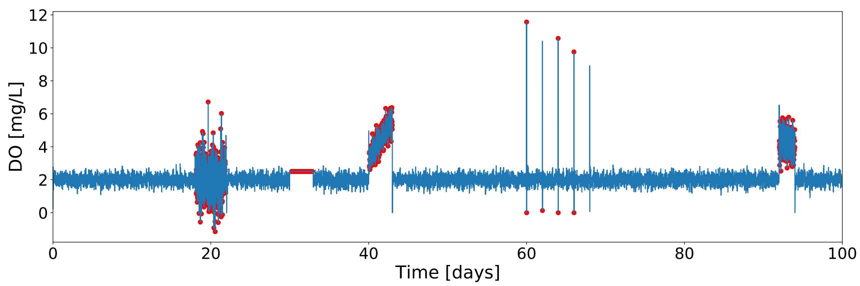
(c) Convolutional-AE applied to Scenario III.

Fig. 7: Faults identified in the three scenarios by the Convolutional-AE.

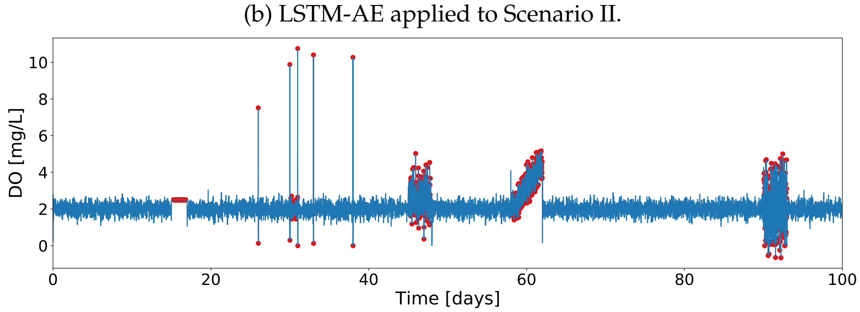
95%, 94% and 94.1% for the same failures, respectively. Both algorithms presented greater difficulties with the spike failure. The Convolutional-AE and LSTM-AE showed correct identification of faults in 88.23% and 84.61%, and considerable FN_r, with values of 11.76% and 15.38%, respectively. Both models performed well on the Stuck fault. The Convolutional-AE presented 88%, and the LSTM-AE 86.39% for TP_r. But these two models showed considerable value for FN_r, with 12% for the Convolutional-AE, and 13.61% for the LSTM-AE model. There was an identification problem near the spike faults, in the third scenario considered. Outside the evaluation area, between the second and third spikes, the models identified 99 normal samples as faults.



(a) LSTM-AE applied to Scenario I.



(b) LSTM-AE applied to Scenario II.



(c) LSTM-AE applied to Scenario III.

Fig. 8: Faults identified in the three scenarios by the LSTM-AE.

The treatment process carried out by WWTPs presents a characteristic of slow changes. In the BSM2 simulator, the water retention time in the activated sludge tank (where the DO sensor takes its measurements) is 14 hours. Taking water retention time into account, both models had satisfactory results. The delay for fault detection detection, for each fault was calculated, as in the previous cases, by the arithmetic average of the delay detection times in the three proposed scenarios. The time that each algorithm took to identify the failures can be seen in Table 2. Related to the delay, it is observed that the Convolutional-AE obtained better results than the LSTM-AE, with the lowest average delay for fault detection. Only when detecting the bias fault, LSTM-AE obtain a better result, with 13.6% less delay in fault identification, in relation to Convolutional-

Table 2: Performance of LSTM-AE and Convolutional-AE. The values are the arithmetic average of the models' performance in the three scenarios. In bold the best results.

Autoencoder	Fault	Assessment metrics			
		$\mathbf{TP_r}$[%]	$\mathbf{FP_r}$[%]	$\mathbf{FN_r}$[%]	**Delay** [h]
Convolutional-AE	Bias	**95.6**	0	**4.4**	**2.5**
	Drift	**95.15**	0	**4.85**	**4.41**
	PD	**97.8**	0	**2.2**	**1.08**
	Spike	**88.23**	12.5	**11.76**	0
	Stuck	**88**	0	**12**	**5.25**
LSTM-AE	Bias	95	0	5	2.16
	Drift	94	0	6	5.33
	PD	94.1	0	4	1.83
	Spike	84.61	12.5	15.38	0
	Stuck	86.39	0	13.61	6.03

AE. Both models readily identify the first peak of the skipe fault. For the other faults, the Convolutional-AE performed better. It took 20.86%, 69.4% and 14.8% less time to identify drift, PD and stuck faults, respectively, when compared with LSTM-AE.

6 Conclusions

WWTPs play a key role in dealing with the problem of water scarcity and thus alleviating the resulting economic and social problems. The work proposed in this paper helps to make WWTPs more secure and reliable. This paper proposed the application of AE for fault detection in DO sensors in biological reactors of WWTPs. Convolutional-AE and LSTM-AE were used to detect five types of faults: bias, drift, PD, spike and stuck. The models had their hyperparameters chosen with the help of the grid search process, using the MAE metric to evaluate the input signal reconstruction error. Three scenarios were considered, with variations in the order of appearance, duration and intensity of faults injected into the dataset. The best performance was obtained by the Convolutional-AE, with better detection values, according to the considered metrics, and less delay time when identifying faults. The analysis of other combinations for hyperparameters or the use in conjunction with other methods that allow less delay in fault detection would make the Convolutional-AE a good option to detect faults such as bias and drift in real WWTPs, representing an important contribution to its safety and sustainability.

Acknowledgments

This research was supported by the ERDF and national funds through the project SYNAPPS (CENTRO-01-0247-FEDER-046978). We also acknowledge the support of the EC project CHIST-ERA-19-XAI-012, and project CHIST-ERA/0004/2019 funded by FCT.

References

1. W. D. Haglund and M. H. Sorg, “Human remains in water environments,” *Advances in forensic taphonomy: method, theory, and archaeological perspectives*, pp. 201–218, 2002.
2. C. Cassardo and J. A. A. Jones, “Managing water in a changing world,” *Water*, vol. 3, no. 2, pp. 618–628, 2011.
3. M. M. Mekonnen and A. Y. Hoekstra, “Four billion people facing severe water scarcity,” *Science advances*, vol. 2, no. 2, p. e1500323, 2016.
4. H. E. Macedo., B. Lehner, J. Nicell, G. Grill, J. Li, A. Limtong, and R. Shakya, “Distribution and characteristics of wastewater treatment plants within the global river network,” *Earth System Science Data*, vol. 14, no. 2, pp. 559–577, 2022.
5. D. Rosso, L. L. E, and M. K. Stenstrom, “Aeration of large-scale municipal wastewater treatment plants: state of the art,” *Water Science and Technology*, vol. 57, no. 7, pp. 973–978, 2008.
6. S. Nakkasunchi, N. J. Hewitt, C. Zoppi, and C. Brandoni, “A review of energy optimization modelling tools for the decarbonisation of wastewater treatment plants,” *Journal of Cleaner Production*, vol. 279, p. 123811, 2021.
7. M. Miron, L. Frangu, S. Caraman, and L. Luca, “Artificial neural network approach for fault recognition in a wastewater treatment process,” in *2018 22nd International Conference on System Theory, Control and Computing (ICSTCC)*, 2018, pp. 634–639.
8. B. Mamandipoor, M. Majd, S. Sheikhalishahi, C. Modena, and V. Osmani, “Monitoring and detecting faults in wastewater treatment plants using deep learning,” *Environmental monitoring and assessment*, vol. 192, no. 2, pp. 1–12, 2020.
9. X. Li, W. Chai, T. Liu, and J. Qiao, “Fault detection of dissolved oxygen sensor in wastewater treatment plants,” in *IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society*, 2020, pp. 225–230.
10. H. Xiao, D. Huang, Y. Pan, Y. Liu, and K. Song, “Fault diagnosis and prognosis of wastewater processes with incomplete data by the auto-associative neural networks and arma model,” *Chemometrics and Intelligent Laboratory Systems*, vol. 161, pp. 96–107, 2017.
11. K. Wang, P. Chang, and F. Meng, “Monitoring of wastewater treatment process based on slow feature analysis variational autoencoder,” in *2021 IEEE 10th Data Driven Control and Learning Systems Conference (DDCLS)*. IEEE, 2021, pp. 495–502.
12. K. L. King, Z. Wang, and D. J. Kroll, “Classification of deviations in a process,” Feb. 14 2006, uS Patent 6,999,898.
13. O. Schraa, B. Tole, and J. B. Copp, “Fault detection for control of wastewater treatment plants,” *Water science and technology*, vol. 53, no. 4-5, pp. 375–382, 2006.
14. P. Kazemi, J. Giralt, C. Bengoa, A. Masoumian, and J.-P. Steyer, “Fault detection and diagnosis in water resource recovery facilities using incremental PCA,” *Water Science and Technology*, vol. 82, no. 12, pp. 2711–2724, 08 2020.

15. B. Wang, Z. Li, Z. Dai, N. Lawrence, and X. Yan, "A probabilistic principal component analysis-based approach in process monitoring and fault diagnosis with application in wastewater treatment plant," *Applied Soft Computing*, vol. 82, p. 105527, 2019.
16. K. B. Newhart, R. W. Holloway, A. S. Hering, and T. Y. Cath, "Data-driven performance analyses of wastewater treatment plants: A review," *Water research*, vol. 157, pp. 498–513, 2019.
17. F. R. Spellman, *Handbook of water and wastewater treatment plant operations*. CRC press, 2003.
18. J. Mendes, R. Araújo, T. Matias, R. Seco, and C. Belchior, "Automatic extraction of the fuzzy control system by a hierarchical genetic algorithm," *Engineering Applications of Artificial Intelligence*, vol. 29, pp. 70–78, March 2014.
19. U. Jeppsson, M.-N. Pons, I. Nopens, J. Alex, J. Copp, K. Gernaey, C. Rosén, J.-P. Steyer, and P. Vanrolleghem, "Benchmark simulation model no 2: general protocol and exploratory case studies," *Water Science and Technology*, vol. 56, no. 8, pp. 67–78, 2007.
20. J. Mendes, N. Sousa, and R. Araújo, "Adaptive predictive control with recurrent fuzzy neural network for industrial processes," in *Proc. 16th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2011)*. Toulouse, France: IEEE, September 5-9 2011, pp. 1–8.
21. J. Mendes, R. Araújo, T. Matias, R. Seco, and C. Belchior, "Evolutionary learning of a fuzzy controller for industrial processes," in *Proc. of The 40th Annual Conference of the IEEE Industrial Electronics Society (IECON 2014)*. Dallas, TX, USA: IEEE, October 29 - November 1 2014, pp. 139–145.
22. S. U. Jan, Y. D. Lee, and I. S. Koo, "A distributed sensor-fault detection and diagnosis framework using machine learning," *Information Sciences*, vol. 547, pp. 777–796, 2021.
23. S. U. Jan and I. Koo, "A novel feature selection scheme and a diversified-input svm-based classifier for sensor fault classification," *Journal of Sensors*, vol. 2018, 2018.
24. S. Zidi, T. Moulahi, and B. Alaya, "Fault detection in wireless sensor networks through svm classifier," *IEEE Sensors Journal*, vol. 18, no. 1, pp. 340–347, 2017.
25. F. Li, Z. Su, and G. Wang, "An effective dynamic immune optimization control for the wastewater treatment process," *Environmental Science and Pollution Research*, pp. 1–16, 2021.
26. W. Pedrycz and S.-M. Chen, *Deep learning: Concepts and architectures*. Springer, 2020.
27. N. Srivastava, E. Mansimov, and R. Salakhudinov, "Unsupervised learning of video representations using lstm," in *International conference on machine learning*. PMLR, 2015, pp. 843–852.
28. J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *International conference on artificial neural networks*. Springer, 2011, pp. 52–59.