applied sciences

MDPI

Article

Fault Type Diagnosis of the WWTP Dissolved Oxygen Sensor Based on Fisher Discriminant Analysis and Assessment of Associated Environmental and Economic Impact

Alexandra-Veronica Luca, Melinda Simon-Várhelyi, Norbert-Botond Mihály and Vasile-Mircea Cristea *

Department of Chemical Engineering, Faculty of Chemistry and Chemical Engineering, Babes-Bolyai University of Cluj-Napoca, 11 Arany János Street, 400028 Cluj-Napoca, Romania
* Correspondence: mircea.cristea@ubbcluj.ro

Featured Application: For automatically controlled WWTPs, the prompt DO sensor faults identification is essential, and the efficiency of the straightforward proposed methodology is shown. Barely revealed by the literature, the presented results also reveal the combined assessment of the impact on the environment and costs of different sensor failures. They are of great interest to researchers and practitioners seeking safe and optimal WWTP operation and provide a robust quantitative impact assessment methodology aimed at improving plant sustainability.

Abstract: Sensor failures are common events in wastewater treatment plant (WWTP) operations, resulting in ineffective monitoring and inappropriate plant management. Efficient aeration control is typically achieved by the dissolved oxygen (DO) control, and its associated sensor becomes critical to the whole WWTP's reliable and economical operation. This study presents the Fisher discriminant analysis (FDA) used for fault diagnosis of the DO sensor of a currently operating municipal WWTP. Identification of the bias, drift, wrong gain, loss of accuracy, fixed value, complete failure minimum and maximum types of DO sensor fault was investigated. The FDA-proposed methodology proved efficiency and promptitude in obtaining the diagnosis decision. The consolidated fault identification showed an accuracy of 87.5% correct identification of the seven faulty and normal considered classes. Depending on the fault type, the results of the diagnosing time varied from 2.5 h to 16.5 h during the very first day of the fault appearance and were only based on observation data not included in the training data set. The latter aspect reveals the potential of the methodology to learn from incomplete data describing the faults. The rank of the fault type detection promptitude was: bias, fixed value, complete failure minimum, complete failure maximum, drift, wrong gain and loss of accuracy. Greenhouse gases (GHGs) such as nitrous oxide (N_2O) and carbon dioxide (CO_2) emitted during wastewater treatment, electrical energy quantity in association with costs spent in the WWTP water line and clean water effluent quality were ranked and assessed for the normal operation and for each of the DO sensor faulty regimes. Both for CO_2 and N_2O, the on-site emissions showed the most significant GHG contribution, accounting for about three-quarters of the total emissions. The complete failure maximum, fixed value and loss of accuracy were the DO sensor faults with the highest detrimental impact on GHG-released emissions. The environmental and economic study reveals the incentives of the proposed DO sensor faults identification for the WWTP efficient and environmentally friendly operation.

Keywords: fault identification; Fisher discriminant analysis; dissolved oxygen sensor; energy costs assessment; GHG emissions assessment

Citation: Luca, A.-V.; Simon-Várhelyi, M.; Mihály, N.-B.; Cristea, V.-M. Fault Type Diagnosis of the WWTP Dissolved Oxygen Sensor Based on Fisher Discriminant Analysis and Assessment of Associated Environmental and Economic Impact. *Appl. Sci.* **2023**, *13*, 2554. https://doi.org/10.3390/app13042554

Academic Editors: Xiongbo Wan, Sheng Du, Wei Wang and Hao Fu

Received: 19 December 2022
Revised: 11 February 2023
Accepted: 14 February 2023
Published: 16 February 2023

1. Introduction

Wastewater treatment plants (WWTPs) are essential for eliminating pollutants from wastewater and converting it into clean effluent discharged in rivers or reused for different applications, including irrigation [1]. As treatment regulations on effluent quality are

continuously tightening, the values of primary variables, such as organic matter, ammonia, nitrates, phosphorus and suspended solids, have to strictly conform to the challenging water quality requirements. The operation of WWTPs has proven to be tough and challenging. A detailed description of wastewater treatment processes, as well as the models associated with them, are extremely complicated, exhibit nonlinear behavior and are characterized by a large number of intensely changing variables. The most difficult challenges for WWTP management and control of the operation, aside from establishing the appropriate control system structure and determining optimal values for the operated variables, are the daily, weekly or seasonal influent composition fluctuations. Comparing control systems' performance in different plants is difficult and necessitates the creation of simulation standards and benchmarks [2]. Modeling water treatment processes has become a very useful tool for the design, optimization and automatic control of the WWTP, as the importance of the treatment plants has grown today within the circular economy concept, and they are presently considered water resources recovery facilities.

The International Water Association (IWA) is well-known for its vast contributions to discovering solutions and fighting against global water problems. IWA experts have created advanced phenomenon-based models to support the construction and control of sewerage treatment processes as part of the envisioned goal of improving standards for sustainable water management [3]. Activated Sludge Model #1 (ASM1) is the most common of these. It was created by the IWA (formerly IAWQ) to standardize nomenclature and set a milestone in wastewater treatment plant modeling [4]. ASM2, ASM2d and ASM3 have been developed over time with the scope of increasing the capability of revealing the intrinsic behavior of processes that occur within the activated sludge water processing [5]; they are currently being employed for sewage treatment plant design and control and are widely used for estimation or optimization [6]. Benchmark Simulation Model #1 (BSM1) was created to enlarge the ASM1 modeling with a defined plant configuration. It has become a standard tool and method for evaluating the performance of wastewater treatment plants. BSM1 has been expanded to Benchmark Simulation Model #2 (BSM2) as a more comprehensive plant description for long-term evaluation, coupled with extensions to primary sewage treatment plant subunits and processing of sludge by digestion in anaerobic conditions [7]. Both benchmarks have become powerful tools to support the development of various control strategies and the evaluation of their performance [8].

Chemical process monitoring is important for evaluating process performance and improving process efficiency and wastewater quality in wastewater treatment plants. The advantage of using mechanical or statistical mathematical models is to estimate the behavior of main key process parameters in different operating circumstances and, in association with real-time measurements, to use them for process design or operational improvement. Calibrated models are a necessary prerequisite for model-driven control solutions due to their high usage in advanced process automation. Supervision and control of water plant process variables are attained to varying extents and complexity depending on the wastewater inflow, the specificity of the bioreactor configuration, and the actual equipment available at each facility. Successful, accurate monitoring and control of processes requires reliable information on the composition, flow rate and temperature for both influents, recycle flows or effluents. Despite the system's high level of complexity, process monitoring and simple univariate fault detection methods proved critical to ensuring that the control system has access to consistent data [9]. Consequently, efficient process control is strongly relying on the capability of detecting sensor faults before the extension of their undesired effects [10]. Failures that are detected early on can help avert further induced undesired plant performance consequences and breakdowns. In general, process monitoring addresses four different tasks: (1) fault detection, which indicates that something is wrong in the process operation; (2) fault identification (or diagnosis), which establishes the cause of the issue; (3) fault estimation, which determines the magnitude of the defect; and (4) fault reconstruction, which calculates the fault-free process variable values aimed to perform the operation in the presence of faults [11]. Traditional fault detection and

isolation methods use a mathematical model of the system to detect and isolate faults. In order to reveal discrepancies between the fault-affected and fault-free circumstances, these methods employ state estimation, parameter identification techniques, and parity relations [12–14]. However, developing precise mathematical models that characterize all physical and biochemical phenomena that occur in industrial processes is often challenging and time costly. When analytical models are not available, knowledge-based approaches such as expert systems might be considered as an alternative or supplement to analytical model-based approaches [15]. However, the development of these knowledge-based systems also necessitates a significant amount of time and work [16].

Online field-installed probes and offline laboratory analyses are used in traditional monitoring of the process variables. Real-time monitoring is difficult to be implemented, expensive and time-consuming. Field instruments require frequent and qualified maintenance, and the lack of satisfying these demands often makes the field measurements unreliable. On the other hand, time implied by laboratory measurements may range from a few minutes to several days to build a trustful assessment of the most standard wastewater monitoring metrics [17]. Both hardware and soft sensors can manifest a series of problems. While hardware sensors require a long time for maintenance and calibration, they exhibit insufficient accuracy and high noise levels. Their accuracy can quickly deteriorate in time and propose the support of soft sensors that can solve some of these issues. The soft sensors may also be subjected to faults as they rely on some sensor measurements and might depend on insufficiently accurate models [18]. As a result, detecting fault-affected measurements among the recorded data is critical for obtaining high WWTP effluent quality results. Sensor deficiencies and techniques utilized for detecting, amending and identifying faulty information or broken sensors were studied in a few works using different multivariate statistical process control (MSPC) methods, such as principal component analysis (PCA) [10,19–21], independent component analysis (ICA) [22,23], partial least squares (PLS) [24–26] or control charts [27,28]. Dynamic multiblock partial least squares (DMBPLS) was implemented to detect chemical oxygen demand (COD) bias and pH drifting sensor faults at a Chinese papermaking wastewater treatment plant [29], while a combination of t-distribution stochastic neighbor embedding with a Gaussian mixture model (t-SNE-GMM) was proposed for detecting bias, drift and complete failure sensor faults in a similar plant [30]. Because of their natural identification limit, MSPC techniques show high potential and are proficient in observing the sensor deficiencies occurring in time-varying, poorly characterized and nonlinear behavior framework of measurements in wastewater treatment plants [31–33]. MSPC methodologies are most frequently used and are superior to statistical process control (SPC) methods since they straightforwardly consider and use idle factors, prompting effective issue identification [34]. Another study combined sub-period division strategies with multiway principal component analysis for the fault diagnosis on a sequence batch reactor of the wastewater treatment process in a paper mill [35].

FDA is a widely used pattern classification technique [36], and its application to chemical process data analysis has continuously increased in the last two decades [37–40]. FDA application to faulty sensors was studied for air handling units [41]. The target of the FDA method is to determine the Fisher optimal discriminant vector that maximizes the Fisher criterion function. Fisher discriminant analysis uses a linear combination of features to distinguish between two or more classes in an optimal way. It is an empirical method based on observed characteristics over a large number of cases. FDA for fault identification provides the best lower dimensionality representation in terms of a discriminant between data classes, where each class corresponds to data acquired during a specific and known fault. Unlike PCA, which is looking for directions that are effective for representation, FDA is looking for directions that are effective for discrimination. From a theoretical standpoint, FDA has advantages for fault visualization and diagnostics [39].

In general, the economic efficiency and environmental friendliness of sewage treatment plants are primarily based on the process of removing nitrogen from pollutants by

biologically catalyzed oxidation using aeration control. This control is primarily sustained by measuring the content of dissolved oxygen (DO) with dedicated sensors. The effectiveness of an aeration control system can be significantly impacted by wrong information that emerges from the dissolved oxygen sensors, resulting in unexpected degradation of the system functionality and problems throughout the wastewater treatment process. Negative consequences in terms of deterioration consist of a decrease in the quality of the effluent, an increase in energy consumption, a decrease in environmentally sustainable performance or even a temporary shutdown of the plant. Few studies have addressed the problem of DO sensor failure. In typical research, the PCA approach was used to detect the set of three categories of faults [42]. Another study examined single-type DO and level sensor failures [28]. The wrong output signal from the sensor was analyzed in another work [19], and a study was devoted to detecting the clogging bias of the sensor based on PCA [20]. The variety of the detection tools was extended to element recognition, neural networks with radial transfer functions [43], binary classification instruments [44], approaches based on impulse response [45] or deep dropout neural networks (MC-DDNN) to identify incipient faults of sensors installed in wastewater treatment plants [46]. In a previous investigation, the current authors developed a study based on the PCA-based methodology that proposed detection solutions for determining the presence of inappropriate functioning of the dissolved oxygen sensor [47].

The motivation and contributions of the present work originate from several reasons. The reported fault identification studies referring to the DO sensor defects were typically focused on a single or on a limited number of fault types. Extending the set of fault types embedded in a single diagnosing tool hinders the efficiency of the fault categorization but makes it most appreciated. The DO sensor fault diagnosing in a controlled A^2O configured WWTP, where the sensor is implied in the automatic operation associated with the nitrites and nitrates concentration control loop, was not specifically reflected in the literature. The literature presenting the impact of different DO sensor fault types on the WWTP energy, effluent water quality and GHG emissions performance is still lacking, although they have become of very high interest. To the best of the authors' knowledge, no WWTP single study of fault detection has been conducted on six different sensor fault types combined with the assessment of associated environmental and economic impact.

The goal of the present paper is to use multivariate statistical methods to construct a fault diagnosis method for the wastewater treatment facility. This research was intended to explore efficient diagnostic ways for specific defects caused by faulty DO sensors and to assess the environmental and cost impact of the faults. Six different fault types were considered: bias, drift, wrong gain, loss of accuracy, fixed value or complete failure. The original contribution of the paper addresses the comprehensive and comparative disclosure of the DO sensor fault types and evaluation of the fault identification benefits for the automatically controlled treatment plant. Furthermore, for the municipal WWTP case study, the energy costs and environmental impact on the treatment plant performance were assessed for the various defects of the DO sensor.

The structure of the present work presents (i) the dynamic model of the WWTP used for simulations, (ii) the basic theory underpinning the FDA fault-diagnosis methodology for the identification of the sensors' faults and (iii) the presentation of the equations used to compute the GHGs released by the fault affected WWTP process; then, the (iv) results and discussions of the performance are obtained by the FDA approach and the environmental and economic study, and the paper ends with (v) the conclusions of the research for improving the performance of the case study WWTP.

2. Materials and Methods

2.1. Process Model

The sewage treatment plant considered in this work has an anaerobic–anoxic–aerobic (A^2O) arrangement. The sewage enters the treatment plant with a flow rate of about 115,000 m^3/day, and it undergoes a series of physical treatments: mechanical filtration,

separation of sand and grease, and the first sedimentation step. Water exits the first clarification step and enters the bioreactor tank. Here, activated sludge technology significantly reduces the concentration of carbon, nitrogen and phosphorus compounds. Three different zones characterize the biodegradation basins. The first one is anaerobic and has a capacity of about 9000 m^3; the second is anoxic and has a volume of around 12,800 m^3; the third is aerobic and comprises a total volume of approximately 33,000 m^3. Biological phosphorous removal is accomplished in the first bioreactor, and the transformation of nitrates and nitrites (NO) produced in the aerobic reactor occurs in the second bioreactor [48]. Physical separation takes place in a secondary clarifier. The clean water is discharged to the emissary river as purified water. A small fraction of 0.5% of the secondary clarifier bottom sludge flow rate is directed to the plant unit for digestion. The bulk is returned to the anaerobic degradation tank as external recycling. A second recycle flow, called internal recycling, recycles nitrate from the aerated bioreactor to the anoxic bioreactor for denitrification [49,50].

The WWTP model constructed and developed in this study was built on the foundations of BSM1 and ASM1. To comply with the A^2O configuration, size and operation characteristics of the case study targeted municipal WWTP, the appropriate modifications were made to the original BSM1. Each of the basic structural parts of the WWTP was described by a set of differential and algebraic equations: primary settler [51], anaerobic bioreactor, anoxic bioreactor, the set of 3 aerated bioreactors [51] and secondary settling unit [52]. Design and dry weather operation data were collected and reconciled from municipal WWTP measurements. The plant data were further used to update and calibrate the modified BSM1 model [53]. The model parameters were calibrated using optimization techniques. This previously calibrated model of the municipal sewage treatment plant was used in the current work to carry out the scenarios of simulations for the normal and faulty sensors. The model was implemented in MATLAB version 8.5 (MathWorks, Natick, MA, USA) and Simulink^{TM} version 8.5 software (MathWorks, Natick, MA, USA). In order to speed up the simulation of Simulink S-functions and save computational resources, the mathematical model was written in the C++ programming language and compiled as a MATLAB executable.

Automatic control is widely regarded as critical for the efficient and safe operation of the WWTP, and the prevention of pollution spread into freshwater ecosystems [54,55]. When associated with optimization, control generates a highly valuable synergy [56,57]. This study employed two typical automatic control structures: one is the dissolved oxygen control loop, and the other one is devoted to the control of nitrates and nitrites concentration. Both control loops used proportional–integral (PI) controllers. The DO concentration in the aerated bioreactor was controlled by the first loop, and the nitrate and nitrite concentrations in the anoxic bioreactor were controlled by the second loop. For the DO concentration control loop, the DO reference of 2 mg O_2/L was enforced in the aerobic bioreactor in order to prevent excessive aeration [58]. The DO controller manipulates the air control valves to supply the necessary oxygen flow rate into the aerated reactors. In the sequence of three aerobic reactors (No. 3, 4 and 5) of the modified BSM1 plant configuration, the airflow calculated by the oxygen controller was unevenly distributed. The following weighting factors for the DO controller generated airflow rate were used: 0.9 for the first aerobic (reactor no. 3), 0.5 for the second aerobic (reactor no. 4) and 0.3 for the third aerated bioreactor (reactor no. 5). The oxygen requirements for the nitrification bioreactors were used to set this distribution strategy, according to the commonly used practice emerged approach. The concentrations of nitrate and nitrite in the anoxic reactor (reactor #2) were controlled via a second control loop by adjusting the recycle flow of nitrates and nitrites. This control loop had a reference of 0.01 mg N/L. The overall configuration of the WWTP reactors, settlers and the main control loops are shown in Figure 1.

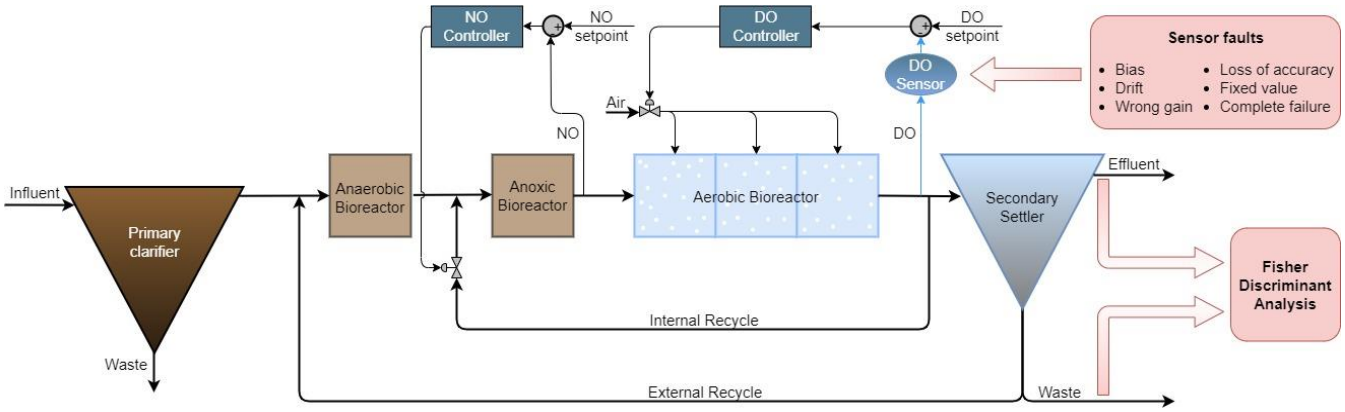

Figure 1. WWTP general structure.

The main parameters of the plant are presented in Table 1. They correspond to the typical structure and actual operating configuration of plant equipment and influent wastewater of the municipal WWTP of Cluj-Napoca, Romania.

Table 1. Municipal WWTP main equipment parameters.

Equipment	Parameter	Value	Measurement Unit
Primary settler	Area	2125	m^2
	Height	3.5	m
Anaerobic bioreactor	Volume	9015	m^3
Anoxic bioreactor	Volume	12,678	m^3
Aerated bioreactors	Volume	33,066	m^3
	Area	6012	m^2
Secondary settler	Area	11,304	m^2
	Height	3	m

Data for both the normal state and faulty state of operation were generated by simulation for each of the specifically designed cases (types) of DO sensor failure. Corresponding data to every individual type of fault and of normal operation were separated into classes and subjected to global Fisher discriminant analysis in order to obtain a distinct class representation of the high-dimensional data.

2.2. Sensor Faults

Lack of sensor functioning can be caused by a variety of factors, including a signal of doubtful quality [59], a bad electrical connection or sensor failure [60]. The diagnosis of different DO sensor malfunction types is of critical importance for the overall WWTP operation. The sensor defects investigated in the present work are:

- Bias—caused by the miscalibrated sensor, is a change in the sensor-generated signal, such as the supplied value being altered typically by a constant value in addition to the actual sensor signal [59].
- Drift—characterized by the continuous time-varying divergence of the sensor signal from the genuine one [59].
- Wrong gain—happens when the sensor slope is affected by an inappropriate gain factor, frequently determined incorrectly during the calibration process [61].
- Loss of accuracy—occurs when the signal of the sensor changes randomly, and its value is imprecise when it is compared to the genuine value [59].
- Fixed value—the sensor displays the same value all the time [61].
- Complete failure (with two cases, minimum and maximum)—characterized by a measured value that is equal or close to the sensor minimum calibration limit or is equal or close to the maximum calibration value [61].

The majority of the simulated fault types originate from the WWTP operating practice, and they were chosen to follow the typical behavioral pattern. For the loss of accuracy type of fault, the random scenario used for the faulty sensor signal was chosen to roughly reflect an irregular change in the true signal. For other magnitudes of the bias, drift, wrong gain and loss of accuracy types of faults, where fault samples are not directly available from the actual WWTP measurements, the use of the calibrated dynamic model can be used for generating data to be embedded in the FDA fault identification methodology. As new faults and magnitudes show up, the database of faults behavior can be completed by the plant measurements, and the diagnosing methodology becomes adaptive.

2.3. Simulation Methodology of the Faults

Specific software modules were created to simulate the behavior of the 6 DO concentration sensor errors. These are integrated into the dynamic WWTP simulator. For each error type, the error signal produced by the sensor is given to the proportional–integral oxygen controller as the measured oxygen process variable. The DO control loop adjusts the airflow accordingly, returning back the controlled DO to the desired reference of 2 mg O_2/L without offset. For each failure, the scenario was run as follows. The simulation was carried out without any fault for 139 days. Then, each fault was implemented starting from the 140th day of the simulations and lasted for a period of 28 days. The process variables taken into consideration for the fault identification were considered from the first day of faulty operation, i.e., day 140 [47]. The same type period of measurements for 28 days was considered for collecting data that describe the normal operation.

The DO sensor uses a galvanic measurement technique. The considered scenarios for the faults were generated, such as to imitate the real DO sensor's steady state and dynamic behavior.

For the bias type of fault, the DO sensor signal was considered to have a bias of +1.5 mg O_2/L added to the actual DO process value. It simulates a constant difference between the true DO value and the DO signal produced by the defective air blast assembly of the sensor.

To mimic the drift, a constant value of 0.05 mg O_2/L was integrated in time, and it supplemented the DO genuine value. This fault simulates a defect in the electronic circuits of the transmitter, produced by a time-increasing parasite electric capacitance.

The sensor's wrong gain type of fault was considered by an incorrect gain factor of 1.4, which consists of an incorrect relationship between the actual DO process value and the sensor output. This fault type is potentially generated by a wrong calibration of the sensor or by calibration sudden change due to internal membrane deterioration. The first order filter with a time constant of 0.3 days was used to smoothly introduce the faulty gain over time for passing from normal to fault-affected operation.

The defective measured process variable for the case of loss of accuracy type of fault was generated by adding to the DO true value a random signal value from the interval $-2.5, 2.5$. Such sensor fault can originate from the partial dislocation of the anode of the cathode electrodes of the sensor. Each random sample value had a duration of 0.1 days.

For the fixed value type of fault, the constant of 2.2 mg O_2/L was used. This sensor defect can be produced due to the leak of the sensor filling solution.

For the complete failure minimum and maximum types of faults, the two very low and very high values of 0.1 mg O_2/L and 6 mg O_2/L were considered. They correspond to the practical circumstances when the electrical supply or parts of the transmitter electronic circuits are malfunctioning or due to the defective self-cleaning assembly, which leads to sensor clogging.

2.4. Fisher Discriminant Analysis

Fisher discriminant analysis is a pattern categorization approach that features a very efficient classification potential. FDA's main goal is to determine the Fisher optimal discriminant vector that maximizes the Fisher criterion function. The higher-dimensional feature

space of process measurements can be projected onto the obtained optimal discriminant vector space for constructing a lower-dimensional feature space. Let $X \in R^{n \times m}$ be the matrix that contains the training data for all classes. The total set of n observations for the m measured variables that build the X matrix contains the submatrix X_i as the subset of measurements consisting of n_i rows and corresponding to the class i of samples. $\overline{x_i}$ denotes the m-dimensional sample mean vector for the class i and is given by:

$$\overline{x_i} = \frac{1}{n_i} \sum_{x_j \in X_i} x_j \tag{1}$$

with x_j the set of vectors that belong to class j, then the within-class scatter matrix is defined by:

$$S_w = \sum_{i=1}^{c} S_i \tag{2}$$

where c is the number of classes, and

$$S_i = \sum_{x_j \in X_i} (x_j - \overline{x_i})(x_j - \overline{x_i})^T \tag{3}$$

is the within-scatter matrix for class i.

The between-class scatter matrix is then defined by:

$$S_b = \sum_{i=1}^{c} n_i \, (\overline{x_i} - \overline{x})(\overline{x_i} - \overline{x})^T \tag{4}$$

where $\overline{x}$ is the total mean vector of all means of the columns of X.

The optimal discriminant direction is found by maximizing the Fisher criterion:

$$J(\varphi) = \frac{\varphi^T S_b \varphi}{\varphi^T S_w \varphi} \tag{5}$$

where the maximizer φ is the Fisher optimal discriminant direction that maximizes the ratio of the between-class scatter to the within-class scatter. It may be shown that a vector φ that maximizes $J(\cdot)$ must satisfy the equation:

$$S_b \varphi = \lambda S_w \varphi \tag{6}$$

for some constants, λ indicates the separability between classes. If S_w is nonsingular, it is obtained as a conventional eigenvalue problem, described by the following expression:

$$S_w^{-1} S_b \varphi = \lambda \varphi \tag{7}$$

The total-scatter matrix is given by the sum of S_b and S_w:

$$S_t = S_b + S_w \tag{8}$$

If data vector x_j from the m-dimensional space is reduced to the a-dimensional space of the FDA vectors, then its linear transformation is given by:

$$z_i = W_a^T x_j \tag{9}$$

where W_a^T has the a FDA vectors as columns and $z_i \in R^a$.

To diagnose the faults, FDA examines observed data collected under various faults and uses a discriminant function that assesses the similarity between the current data and

the data belonging to each class. When the maximum discriminant function value, g_i, satisfies the following conditions, the observation is allocated to the class i:

$$g_i(x) > g_j(x),\ \forall\, j \neq i \tag{10}$$

$g_i(x)$ is the discriminant function given by a measured vector x for class i, and $g_j(x)$ is the discriminant function given by the measured vector x for class j. The discriminant function can be calculated, for each class i, with the following equations:

$$g_i(x) = -\frac{1}{2}(x-\overline{x_i})^T W_a\left(\frac{1}{n_i-1}W_a^T S_i W_a\right)^{-1} W_a^T(x-\overline{x_i}) + lnP_i - \frac{1}{2}\,ln\left[det\left(\frac{1}{n_i-1}W_a^T S_i W_a\right)\right] \tag{11}$$

where P_i is the a posteriori probability of x to belong to class i, and it is computed by the ratio between the number of observations from a class and the total number of observations for all classes. FDA was implemented using the Statistics and Machine Learning Toolbox version 10.0 of Matlab.

2.5. Assessment of the GHGs Impact

Conventional wastewater treatment facilities purify wastewater and reduce water pollution, but they also discharge GHGs into the air through direct emissions, and as they require a considerable amount of energy to process the influent, they also indirectly contribute to the release of GHGs in the atmosphere. Biological wastewater treatment using activated sludge technology removes organic matter and N and P nutrients from wastewater in an effective manner. Comprehensive environmental implications of the WWTP's contribution to GHG emissions have been highlighted in different studies. Greenhouse gas emissions have been identified as a key negative impact of the WWTP operation and have been studied in several works [62–64]. Other studies offer a complex image of the total environmental impact of a wastewater treatment plant through life cycle assessments [65–67]. However, no studies have looked at the environmental impact of the wastewater treatment plant when anomalies occur in the process due to sensor fault inappropriate operation. This study emphasizes the differences, with respect to both the environmental impact and the energy costs, between the plant performance when it is operated normally and when its operation is affected by faulty operation of the dissolved oxygen sensor.

Overall impact assessment of the WWTP disturbed operation on plant performance was performed using a cumulative performance index composed of three different performance indices: aeration energy (AE), pumping energy (PE) and effluent quality (EQ) [50]. The oxygen mass transfer coefficient of the aerobic bioreactors (K_La), which is directly related to the airflow rate, is used to compute the aeration energy index, as stated in Equation (12).

$$AE = \frac{SO_{sat}}{T\cdot 1.8\cdot 1000}\cdot \int_0^T \sum_{aerated\ reactors} V_{bioreactor}\cdot K_L a_i(t)dt \tag{12}$$

where SO_{sat} is the oxygen saturation concentration (mg O_2/L), T is the time in days of faulty operation, $V_{bioreactor}$ represents the volume of the bioreactor (m^3) and K_La_i is the mass transfer coefficient in the aerated bioreactor i.

The pumping energy index is calculated using the flow rates of nitrate recirculation, return-activated sludge recycling and waste, as shown in Equation (13). The energy used for aeration and pumping is computed in kWh per day. The effluent quality index is determined by a weighted sum of total suspended solids (TSS), chemical oxygen demand, biochemical oxygen demand (BOD), total Kjeldahl nitrogen (TKN), and nitrates and nitrites

concentrations in the effluent flow stream, as indicated by Equation (14). The effluent quality is expressed in kilograms of pollutant units per day.

$$PE = \frac{1}{T} \cdot \int_0^T [0.004 \cdot Q_{NR}(t) + 0.08 \cdot Q_{RAS}(t) + 0.05 \cdot Q_{waste}(t)]dt \quad (13)$$

where Q_{NR} is the flow rate of the nitrate recirculation (m^3/day), Q_{RAS} is the flow rate of the return-activated sludge (m^3/day) and Q_{waste} is the flow rate of waste from the secondary settler (m^3/day).

$$EQ = \frac{1}{T \cdot 1000} \cdot \int_0^T [PU_{TSS}(t) + PU_{COD}(t) + PU_{BOD}(t) + PU_{TKN}(t) + PU_{NO}(t)] \cdot Q_{effluent}(t)dt \quad (14)$$

where PU_{TSS} denotes total suspended solids, PU_{COD} refers to the chemical oxygen demand and PU_{BOD} to the biochemical oxygen demand, PU_{TKN} considers the Total Kjeldahl Nitrogen, PU_{NO} accounts for the nitrates and nitrites, and $Q_{effluent}$ is the effluent flow rate.

Total GHG emissions generated by the water line of the municipal wastewater treatment plant are composed of both on-site and off-site emissions and consist of CO_2 and N_2O gases. Besides CO_2, N_2O is also considered an important contributor to the GHGs, as it has a global warming potential (GWP) of about 265–298 times higher than CO_2, with an average residence time of 100 years [68].

Off-site CO_2 emissions (kg CO_2/day) include indirect CO_2 emissions from the electrical power generation plant that are associated with the electrical energy consumed at the WWTP. They are described by:

$$P_{CO_2,\ off-site} = k_{PG} \cdot e_D \quad (15)$$

where k_{PG} is the site-specific emission factor per unit of energy generated, considered with a value of 0.19 kg CO_{2e}/kWh, and e_D is the total energy demand, calculated as the sum of the aeration energy and pumping energy [64,69].

Off-site N_2O emissions include N_2O that results from biological degradation in the effluent (downstream) of the wastewater treatment plant [64,70]:

$$P_{N_2O,\ off-site} = N_{effluent} \cdot EF_{effluent} \quad (16)$$

where $N_{effluent}$ is the nitrogen load in the effluent discharged into aquatic environments and $EF_{effluent}$ = 0.005 · 44/28 kg N_2O/kg N [71] is the emission factor for N_2O emissions from the discharged wastewater.

The on-site CO_2 emissions emerging from the water line of the aerobic biological processes are computed by the following expression:

$$P_{CO_2,on-site} = Q_{influent} \cdot 0.99 \cdot (1 - Y_H) \cdot \eta_{ASP} \cdot bCOD + Q_{influent} \cdot 1.03 \cdot Y_H \cdot \eta_{ASP} \cdot bCOD \cdot \frac{k_{d,H} \cdot MCRT}{1 + k_{d,H} \cdot MCRT} \quad (17)$$

where $Q_{influent}$ is the plant influent flow rate (m^3/day); 0.99 kg CO_{2e}/kg COD is the emission factor related to organic compounds; Y_H is the heterotrophic biomass yield (massVSS/massCOD) [2]; η_{ASP} is the biodegradable COD (bCOD) removal in the activated sludge reactors; 1.03 kg CO_{2e}/kg COD is the emission factor related to activated sludge biomass; $k_{d,H}$ is the decay rate of heterotrophic biomass and has a value of 0.3 day^{-1} [2]; and MCRT is the mean cell retention time, which is 15 days for this case [64,72].

The on-site N_2O emissions from the water line can be estimated using the following relationship:

$$P_{N_2O,on-site} = Q_{influent} \cdot (TN_{in} - TN_{out}) \cdot r_{N_2O} \quad (18)$$

where TN_{in} represents the total nitrogen from the influent (kg N/m^3), TN_{out} is the total nitrogen in the effluent (kg N/m^3) [73] and r_{N_2O} is the emission rate of N_2O (kg N_2O/kg N) [74].

3. Results and Discussion

3.1. Normal and Abnormal Operation Data Sets

The FDA methodology relied on 17 WWTP process variables. They were the bottom effluent sludge concentrations (10 variables), secondary settler clean effluent concentrations (six variables) and temperature. The set of these variables consisted of total nitrogen (N_{total}); total Kjeldahl nitrogen; chemical oxygen demand; free and saline ammonia (S_{NH}); nitrate and nitrite nitrogen (S_{NO}); total suspended solids (TSS); slowly biodegradable substrate (X_S); heterotrophic biomass ($X_{B,H}$); autotrophic biomass ($X_{B,A}$); inert particulate products (X_P); particulate biodegradable organic nitrogen (X_{ND}); soluble, biodegradable organic nitrogen (S_{ND}); readily biodegradable substrate (S_S); alkalinity (S_{alk}); and temperature (T). The first six of them characterized the secondary settler clean water, while the last eleven variables described the bottom effluent.

Eight separate scenarios were created, and simulations were performed to generate the data sets for the different fault classes, one for normal functioning and six for malfunctioning of the DO sensor. The DO and NO-controlled WWTP were simulated for 168 days of operation. The starting set of 139 days of nominal (faults lacking) functioning was considered to bring the plant to a quasi-steady state. Sensor faults were applied from day 140 in the simulation scenario. Operational data of the simulated wastewater treatment plant were collected with a sampling time of 15 min. Data generated in the time period from the 141st to the 145th day of DO sensor normal and faulty operation were used for training the FDA model. Each set of 480 observations is considered to be a class. They formed the training matrix of observations (3840 lines and 17 columns). The fault diagnosis performance of the trained FDA model was tested using data corresponding to the 140th day, i.e., the first day of faulty sensor operation. The testing data set contained 96 measurements for each fault. This testing approach was designed in order to investigate the capability of the FDA diagnosis method to identify the type of fault in the very first hours following the fault appearance.

The chosen scenarios and the emerged dimensions of the data sets used for training and validation of the fault detection methodology were considered to make a fair trade-off between diagnosis accuracy and required computational resources.

3.2. Fault Diagnosis

The values obtained for the discriminant function of each class, g_i, were compared in order to diagnose the flawed sensor operation. The discriminant function with the highest value indicated the class of the faulty sensor, and consequently, it diagnosed the fault. The values of the discriminant functions g_i were computed for each of the 15 min time-sampled measurements of the testing day no. 140, affected by the different types of faults (seven classes) and for measurements corresponding to normal operation (one class). They are presented in Figures 2–9.

Each of these figures presents the values $g_i(x)$, i = 1 to 8, i.e., the values of the FDA discriminant functions associated with the eight classes of the normal and of the seven fault types. For every sampling time moment of the 140th day (considered as multiples of the 15 min sample time), the eight discriminant function values were computed and presented with a particular line and marker. The vector of observations x, as the independent variable of the discriminant functions $g_i(x)$, consists of the set of values of the features at a particular sampling time moment. Separately, each of the Figures 2–9 shows the values of the discriminant functions at all considered validation sampling time moments and the corresponding vectors of observations obtained for only one of the normal or of the seven investigated fault types cases. In each figure, the discriminant function with the permanent highest value identifies the type of fault indicated by its associated class.

In order to reveal the moment of time within the first day after the fault appearance, when the fault is firmly and permanently diagnosed by the maximum discriminant function, a magnified representation of the time interval in the vicinity of this moment is presented in the special inserted detailed graphical representation.

As may be observed in Figure 2, all of the g_i values confirmed the normal operation (faulty-free values) diagnosis after 16.5 h of the fault appearance.

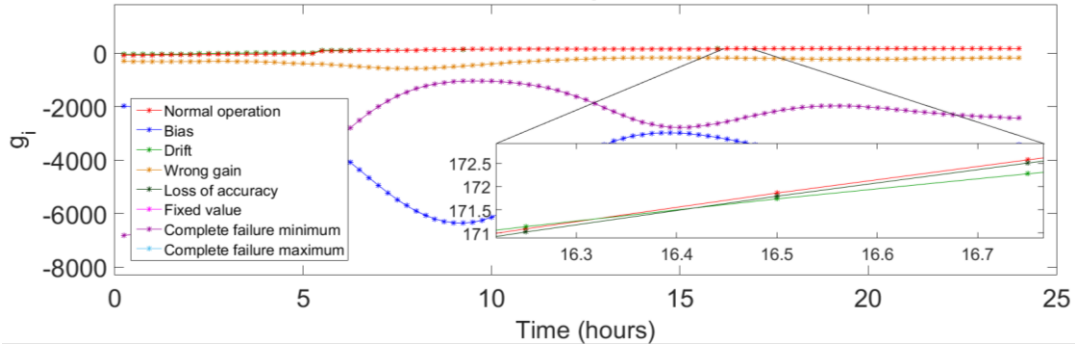

Figure 2. Normal operation diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm normal operation identification.

It is worth noticing that for the first part of the testing day (no. 140) when sensor faults are not yet fully developed, the discrimination between normal operation and faults is less evident.

In Figure 3, the bias fault type diagnosis is performed after 2.5 h, for a time period of 5.75 h. After that period, the automatic controlled WWTP operation is driven progressively (due to the integral component of the PI controller) to low aeration and to values of the process variables that are becoming similar to those corresponding to the maximum values of the DO sensor and, as a result, the FDA discrimination is revealing the complete failure maximum class for the last part of the day.

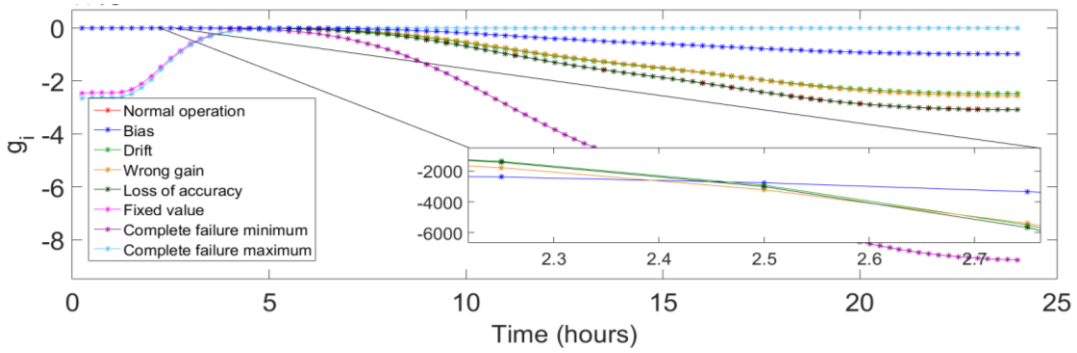

Figure 3. Bias diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm fault type identification.

Drift diagnosis is firmly confirmed after 13.75 h, as it is shown in Figure 4. Drift fault effects are growing in time due to the intrinsic nature of this fault.

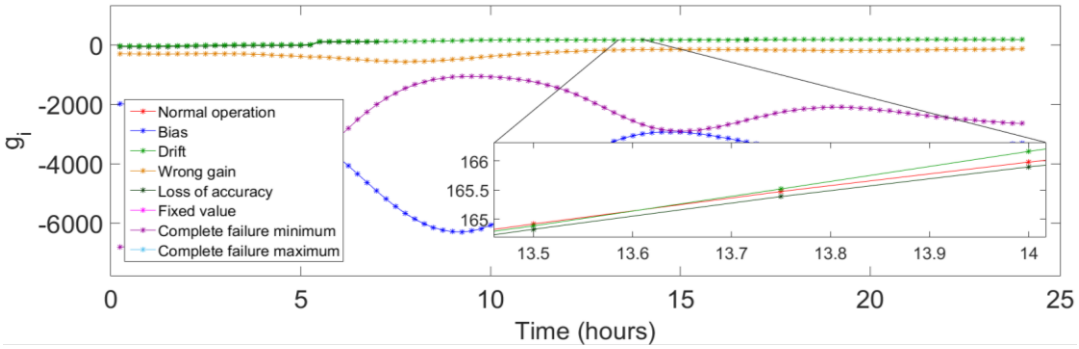

Figure 4. Drift diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm fault type identification.

The diagnosis of wrong gain fault was confirmed after 14 h of the fault incidence moment. The graphical representation is shown in Figure 5. As the wrong gain fault was introduced by a time lag constant, the identification was relatively promptly achieved.

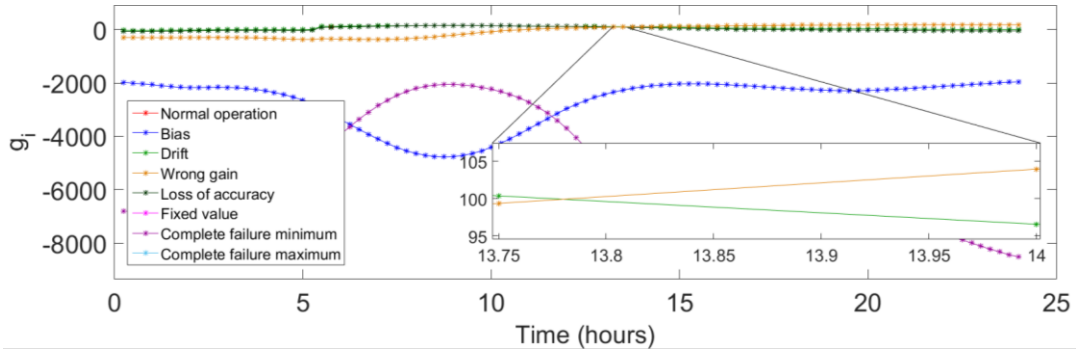

Figure 5. Wrong gain diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm fault type identification.

The loss of accuracy fault was identified after 16.5 h of the fault start moment of action, as it is presented in Figure 6. Despite the irregular character of this fault, determined by the random component of the simulated faulty signal, the permanent discrimination type of fault was well achieved.

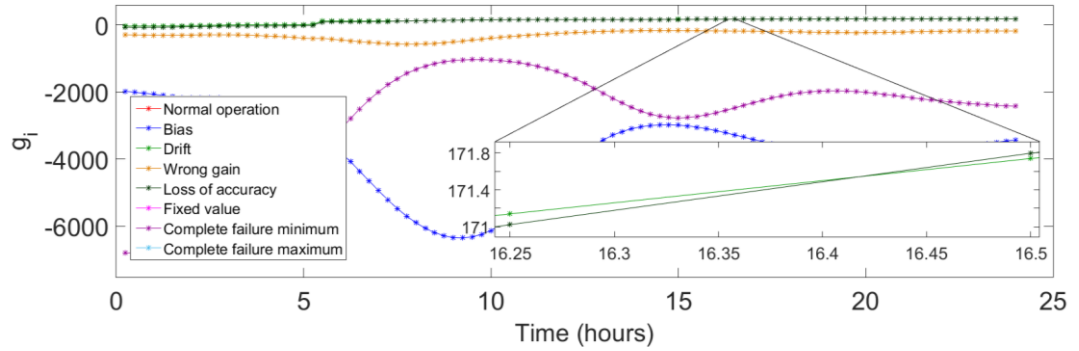

Figure 6. Loss of accuracy diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm fault type identification.

Figure 7 shows the graphical representation of the fixed value fault, which was correctly and promptly identified after 6 h of the fault appearance.

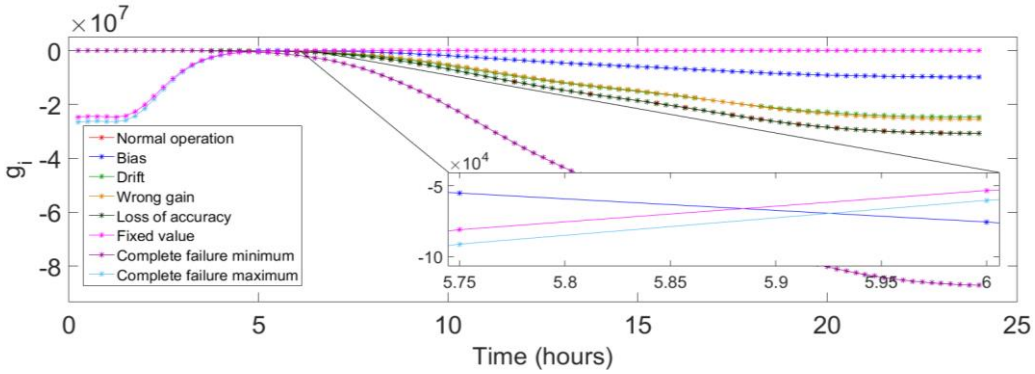

Figure 7. Fixed value diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm fault type identification.

The complete failure minimum is correctly diagnosed after 6.75 h of its intervention, as it is revealed in Figure 8.

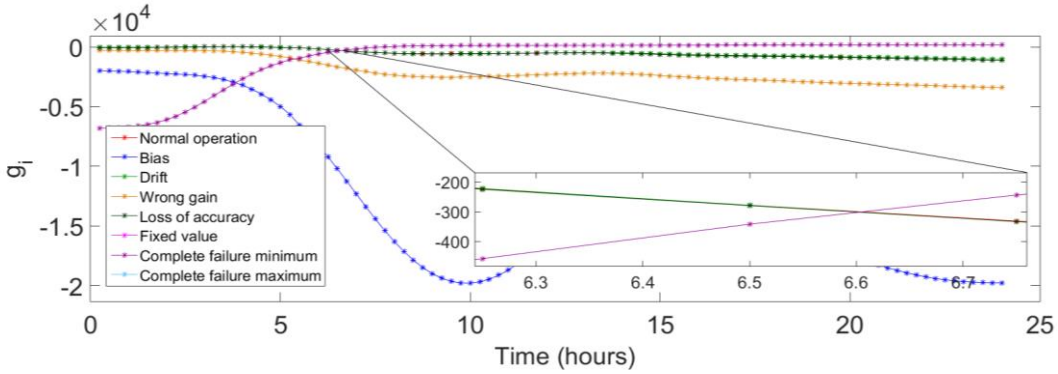

Figure 8. Complete failure minimum diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm fault type identification.

The complete failure maximum diagnosis is diagnosed successfully after 9.5 h, as is shown in Figure 9.

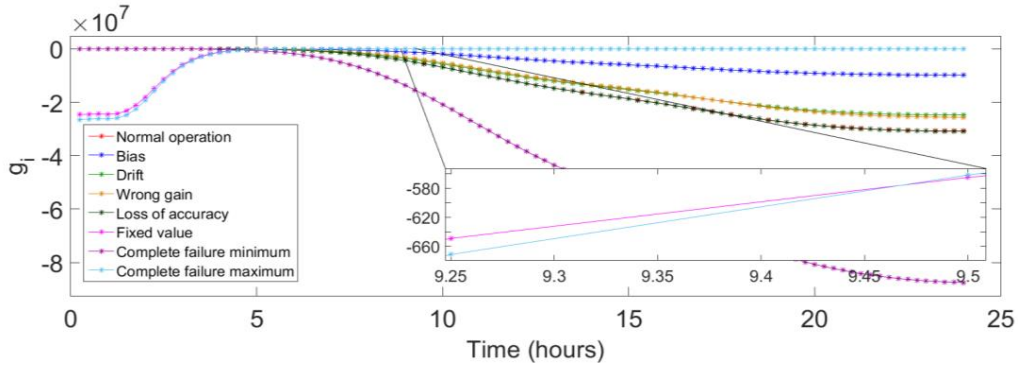

Figure 9. Complete failure maximum diagnosis: graphs of the FDA discriminant functions $g_i(x)$ for each of the normal and seven fault classes, along the 24 h of the 140th day and the detailed representation revealing the class of the observations and the time moment of the firm fault type identification.

The FDA-based identification methodology for the type of DO sensor fault proved to be successful for all of the investigated faults except the bias one (which has very similar effects with complete failure maximum failure). This means an accuracy of 87.5% correct identification of the eight considered cases. The time needed for obtaining the consolidated diagnosis decision varied from 2.5 h to 16.5 h.

From the rare previous studies that addressed fault diagnosis in WWTPs sensors, one showed good detection of bias, drift and precision degradation (loss of accuracy) of sensors used in a WWTP with classic and dynamic PCA-based methods but showed some limitations regarding the fault identification with variable reconstruction-based methods [42]. These limitations included the fact that a reconstruction-based method cannot identify the faulty sensor which causes process transition, so the approach is inappropriate for identifying the fault type connected to a control loop. Additionally, it did not study the fault identification time in detail. The present study investigated several fault types, and the identification time was determined for each of them. Another work proposed a complex-valued slow independent component analysis (CSICA) based method for fault detection and diagnosis with applications to wastewater treatment processes. Despite the fact that the method had good performance in detecting and diagnosing incipient faults, it was inadequate for scenarios of multi-fault and large-scale nonlinear systems [23]. The sub-period division strategies combined with multiway principal component analysis for only two faults diagnosis showed that the used methods could manifest false identification results during normal operation periods and the inability to detect the fault during some time intervals [35]. The present investigations considered six different fault types that implied increased difficulty in the promptitude and accuracy of the fault diagnosis.

The efficiency of the proposed FDA identification is substantiated by the fact that discrimination was achieved by the FDA discrimination model that was trained with faulty data from the five days subsequent to the first one, when the fault actually started to act, and testing was performed for this very first day.

3.3. Performance Indices

AE, PE and EQ were calculated for both normal and faulty operation cases. They were determined as a mean value over 28 days, i.e., for the period of the days ranging from day 140 to day 168. Table 2 shows the values of the AE, PE and EQ for the normal operation and for the six different faults.

Table 2. Values of the performance indices for the normal and faulty operation cases.

Operating Regime	AE (kWh/day)	PE (kWh/day)	Total Energy Demand (kWh/day)	EQ (kg PU/day)
Normal operation	16,992	1329	18,321	16,852
Bias fault	14,206	2415	16,621	21,461
Drift fault	15,569	1746	17,315	17,134
Wrong gain fault	15,866	1593	17,459	16,706
Loss of accuracy fault	9150	2403	11,553	219,189
Fixed value fault	1968	2415	4383	338,737
Complete failure minimum fault	23,537	1039	24,576	19,804
Complete failure maximum fault	1968	2415	4383	338,750

It can be observed that for the bias, loss of accuracy, fixed value and complete failure maximum types of faults, the PE and EQ indices values are much higher than the normal operation values, but AE is lower. The AE values can be explained by the fact that the four fault types are characterized by high values of the DO sensor signal, sent as a wrong feedback signal to the DO controller. As a result, the DO controller diminishes the aeration with the aim of reducing the value of the faulty DO signal, and the aeration energy drops.

This is followed by a decrease in nitrification, which induces a lower concentration of nitrates and nitrites in the aerated bioreactor and, subsequently, in the anoxic bioreactor. Consequently, the NO control loop reacts to this change by increasing the internal recycle flow rate, which determines higher pumping energy. Additionally, the quality of the effluent drastically deteriorated due to the inefficiency of the treatment process, affected by the air-lacking circumstances, and higher amounts of pollutants are discharged with the clean effluent. The high AE value in the case of the complete failure minimum type of fault is due to the increase in the DO controller manipulated airflow rate as a response to the low but faulty value of the DO sensor signal.

It is worth mentioning that in cases of the faulty operation of the DO sensor that is characterized by reduced values of the DO signal, compared to the true one, the DO controller will drive the aeration to increased airflow rates with straightforward consequences on the rise of the AE index values.

3.4. Energy Costs Assessment for Operation Affected by Faults

Operating a sewage treatment plant is highly energy intensive, with implicitly high economic costs. The price of energy is determined by a variety of supply and demand factors, such as the geopolitical situation, the national particular energy mix, distribution network expenses, environmental protection taxes, severe weather conditions or excise and taxation levels. The cost of energy also depends on the energy source. Energy prices vary from 4.8 to 12.10 eurocents depending on the energy source, as can be seen in Table 3. These values were extracted from an International Energy Agency (IEA) report on the mean costs of generating energy. This report provided data for a total of 243 plants in 24 countries from Europe, Africa and Asia [75].

Table 3. Costs of different energy sources.

Source/Technology	Eurocents/kWh
Lignite	9.12
Coal	8.80
Gas (CCGT [1])	6.81
Nuclear	6.62
Wind onshore	4.80
Wind offshore	8.45
Solar PV [2] commercial	5.38
Solar PV residential	12.10
Solar thermal (CSP [3])	11.62
Hydro reservoir	6.91
Hydro run of river	6.53
Geothermal	9.50
Biomass	11.33

[1] CCGT—combined cycle gas turbines, [2] PV—photovoltaic, [3] CSP—concentrating solar power.

The daily costs of normal operation and each type of DO sensor malfunction were determined for the sewage treatment plant considered in this study. The daily costs were individually evaluated for different potential energy sources or energy production technologies. The data in Table 4 were obtained by multiplying the costs of different energy sources [75] with the total energy demand values computed for the normal and faulty operation cases. Table 4 summarizes the findings.

Table 4. Source depending sum of AE and PE energy costs computed for normal and fault affected operation.

Source/Technology	Daily Operation Costs (EUR)							
	Normal Operation	Bias	Drift	Wrong Gain	Loss of Accuracy	Fixed Value	Complete Failure Minimum	Complete Failure Maximum
Lignite	1671	1516	1579	1592	1054	400	2241	400
Coal	1612	1463	1524	1536	1017	386	2163	386
CCGT	1249	1133	1180	1190	787	299	1675	299
Nuclear	1214	1101	1147	1156	765	290	1628	290
Onshore wind	879	798	831	838	555	210	1180	210
Offshore wind	1548	1404	1463	1475	976	370	2076	370
Solar PV commercial	985	894	931	938	621	236	1321	236
Solar PV residential	2216	2011	2094	2112	1397	530	2973	530
Solar thermal (CSP)	2128	1931	2011	2028	1342	509	2855	509
Hydro reservoir	1266	1149	1197	1207	799	303	1699	303
Hydro run of river	1196	1085	1130	1140	754	286	1604	286
Geothermal	1741	1580	1646	1659	1098	417	2336	417
Biomass	2075	1883	1961	1978	1309	497	2784	497

By analyzing the costs of electrical energy spent for the different cases of the faults, it may be observed that complete failure minimum fault of the DO sensor implies the largest energy costs, as the reduced faulty value of the DO sensor determines the control system to considerably increase the airflow rate. On the contrary, the fixed value and complete failure maximum faults imply the smallest and comparable energy costs due to the large values of the DO sensor signal that make the DO controller reduce the airflow rate. However, for these latter cases, the quality of the effluent dramatically deteriorates by a factor higher than 20 and ranks these faults as having the worst overall effect.

Values of the daily energy costs presented in Table 4 also reveal that onshore wind, solar PV commercial and hydro run of rivers are the most favorable sources of energy to be used from the economic costs point of view, while the solar PV residential, solar thermal (CSP) and biomass are the most expensive ones. The lignite (CSS) and coal (CSS) energy sources may be as well considered as having high costs.

Depending on the technical solution availability, switching to the alternative cheaper energy sources and implementing a customized program for the use of a mixed energy source in case of faulty sensor operation might reduce the implied energy costs before the replacement of the defective sensor.

3.5. Environmental Assessment of CO_2 and N_2O Emissions

The on-site and off-site emissions of CO_2 and N_2O, as main contributors to the Green House Gases released by the water line of the WWTP, were estimated for each type of faulty operation and also for the normal one. Their daily mean values are presented in Table 5.

The data in Table 5 revealed that on-site emissions are the most significant, accounting in the case of normal operation for about 75% of the total emissions, both for $P_{CO_2,total}$ and $P_{N_2O,total}$. Computed total CO_2 emissions, i.e., the sum of the on-site and off-site values, showed increased values for all cases of the DO sensor faults, with the exception of the complete failure minimum fault type. Very large total CO_2 emissions (more than one order of magnitude higher than those of the normal operation) were produced in the cases of the following faults: fixed value, complete failure maximum and loss of accuracy fault. Complete failure minimum fault type showed the least emission values, while drift and wrong gain had emissions that were slightly higher than those of the normal operation case.

Table 5. GHG emissions due to DO sensor defect.

Emissions Type	Source/Technology	Emitted Gas	Daily GHG Emissions							
			Normal	Bias	Drift	Wrong Gain	Loss of Accuracy	Fixed Value	Complete Failure Minimum	Complete Failure Maximum
Off-site emissions	Power generation	CO_2, $P_{CO_2,off-site}$, kg CO_2/day	3481	3158	3290	3317	2195	833	4669	833
	Biological degradation in the WWT effluent	N_2O, $P_{N_2O,off-site}$, kg N_2O/day	3.61	2.47	2.89	2.97	14.39	21.49	6.50	21.49
On-site emissions	Water-line aerobic biological processes	CO_2, $P_{CO_2,on-site}$, kg CO_2/day	13,689	30,459	17,851	16,178	461,439	921,028	10,604	921,168
		N_2O, $P_{N_2O,on-site}$, kg N_2O/day	10.35	10.81	10.64	10.05	6.07	3.27	9.20	3.27
Total emissions		CO_2, $P_{CO_2,total}$, kg CO_2/day	17,170	33,617	21,141	19,495	463,634	921,861	15,274	922,001
		N_2O, $P_{N_2O,\ total}$, kg N_2O/day	13.96	13.28	13.53	13.02	20.46	24.76	15.70	24.76
		CO_{2e}, $P_{CO_{2e},overall}$, kg CO_{2e}/day	21,330	37,574	25,173	23,375	469,731	929,239	19,953	929,379

Assessment of the total N_2O emissions for the fixed value, complete failure maximum and loss of accuracy faults also reveal increased values for the N_2O emissions. They are characterized by a factor ranging from 1.4 to 1.7 when compared to the normal operation case. It may be noticed that, contrary to the total CO_2 emissions observed trend, complete failure minimum fault type led to increased values of the N_2O emissions. N_2O emissions produced due to bias, drift and wrong gain faults were marginally smaller when compared to the normal reference case, especially due to the decreased values of the off-site N_2O emissions.

The overall CO_{2e} emission values were computed by cumulating the total CO_2 and total N_2O (as CO_2 equivalent) emissions. According to the results of the overall CO_{2e} emission, all faults show higher CO_2 emission levels than the normal operation case. The only exception is the complete failure minimum type of fault, but in this particular case, the effluent quality deteriorates by more than 15%, and the sum of the aeration and pumping energy has the highest values of all investigated cases. Assessment of the CO_2 and N_2O emissions, issued due to different DO sensor faults, provides valuable quantitative information on the extent and ranking of the most unfavorable sensor fault types that may affect the sustainable operation of the WWTP.

The results validate the logical assumption that the presence of all faults is detrimental to the WWTP operation, and the severity assessment of their consequences has to be considered in an integrated approach of energy, costs, water effluent quality and GHG issues. These evaluations constitute the foundation for the control and safety systems design aimed at achieving the plant sustainability objective.

4. Conclusions

The performance of the FDA-based sensor faults identification approach was proposed and assessed for the key DO sensor belonging to the automatically controlled A^2O WWTP with the typical dissolved oxygen and the nitrates and nitrites concentration control loops. The DO feedback control system performance is responsible for the WWTP nitrification process, directly determining the spent energy, effluent quality and GHG emissions of the entire plant.

The FDA fault identification was investigated for six different types of faults. They were: bias, drift, wrong gain, loss of accuracy, fixed value and complete failure minimum and complete failure maximum faults. The data sets of variables for both the normal

and fault-affected operation of the DO sensor originated from simulations of the calibrated plant model in which the scenario of WWTP influent variations emerged from plant measurements.

The prompt and reliable FDA fault identification methodology was successful. The time needed for obtaining the consolidated diagnosis decision varied from 2.5 h to 16.5 h following the moment of the fault appearance. It is noteworthy that data used for building the training data matrix for all faulty classes, i.e., data from days 2 to 6, did not contain the data of the first day. The latter was only used for testing the promptitude and efficacy of the identification methodology. The complexity of the diagnosis increases when the number of fault types subject to investigation expands or the sensor faults produces comparable effects on the process variables.

The results showed that the WWTP effluent quality performance index depreciated during all of the DO sensor faults. From the EQ index perspective, the most detrimental cases were the complete failure maximum, fixed value and loss of accuracy faults, while the less affected case was the wrong gain fault type.

The WWTP normal and faulty operation modes were further investigated to assess the environmental effect of the on-site and off-site emissions of CO_2 and N_2O GHG and the economic impact of spent aeration and pumping energy. On-site emissions have the most significant GHG contribution, being responsible for about three-quarters of the total emissions, both for CO_2 and N_2O. The complete failure maximum, fixed value and loss of accuracy were the DO fault-affected operation modes that had the most undesired impact on the amount of GHG released emissions. The identification of simultaneous acting fault types needs specific investigations, and this is the subject of future research work.

The comprehensive evaluation of effluent quality, energy costs and GHG emissions is a useful quantitative assessment basis for the control and safety systems design aimed at satisfying multiple objective targets and the overall plant sustainability goal. Instrumentation maintenance tasks, metrological calibration or verification services and designers of intelligent DO sensors may benefit from the presented results for promoting the safe, efficient and environmentally friendly operation of the WWTP.

Author Contributions: Conceptualization, A.-V.L. and V.-M.C.; Methodology, A.-V.L. and V.-M.C.; Software, A.-V.L., M.S.-V. and N.-B.M.; Formal analysis, A.-V.L. and V.-M.C., Investigation, A.-V.L. and V.-M.C.; Resources, M.S.-V. and N.-B.M.; Data curation, A.-V.L., M.S.-V. and V.-M.C.; Writing—original draft preparation, A.-V.L.; Writing—Reviewing and Editing, V.-M.C.; Supervision, A.-V.L. and V.-M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No data availability.

Acknowledgments: The present work has received support from the project: Entrepreneurship for innovation through doctoral and postdoctoral research, POCU/380/6/13/123886, co-financed by the European Social Fund through the Operational Program for Human Capital 2014–2020.

Conflicts of Interest: "The author Melinda Simon-Varhelyi is an employee of MDPI, however she does not work for the journal Applied Sciences at the time of submission and publication." The authors declare no conflict of interest.

References

1. Baklouti, I.; Mansouri, M.; Ben Hamida, A.; Nounou, H.; Nounou, M. Monitoring of wastewater treatment plants using improved univariate statistical technique. *Process. Saf. Environ. Prot.* **2018**, *116*, 287–300. [CrossRef]
2. Alex, J.; Benedetti, L.; Copp, J.; Gernaey, K.V.; Jeppsson, U.; Nopens, I.; Pons, M.N.; Rieger, L.; Rosen, C.; Steyer, J.P.; et al. *Benchmark Simulation Model No. 1 (BSM1)*; Lund University: Lund, Sweden, 2008; pp. 2–29.
3. Mamandipoor, B.; Majd, M.; Sheikhalishahi, S.; Modena, C.; Osmani, V. Monitoring and detecting faults in wastewater treatment plants using deep learning. *Environ. Monit. Assess.* **2020**, *192*, 148. [CrossRef] [PubMed]

4. Rieger, L.; Gillot, S.; Langergraber, G.; Ohtsuki, T.; Shaw, A.; Takacs, I.; Winkler, S. *Guidelines for Using Activated Sludge Models*; IWA Publishing: London, UK, 2012; pp. 11–147, ISBN 9781780401164.
5. Henze, M.; Gujer, W.; Mino, T.; van Loosedrecht, M. *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3*; IWA Publishing: London, UK, 2002; pp. 1–120, ISBN 9781780402369.
6. Hauduc, H.; Rieger, L.; Ohtsuki, T.; Shaw, A.; Takacs, I.; Winkler, S.; Heduit, A.; Vanrolleghem, P.A.; Gillot, S. Activated sludge modelling: Development and potential use of a practical applications database. *Water Sci. Technol.* **2011**, *63*, 2164–2182. [CrossRef] [PubMed]
7. Jeppsson, U.; Rosen, C.; Alex, J.; Copp, J.; Gernaey, K.V.; Pons, M.N.; Vanrolleghem, P.A. Towards a benchmark simulation model for plant-wide control strategy performance evaluation of WWTPs. *Water Sci. Technol.* **2006**, *53*, 287–295. [CrossRef]
8. Jeppsson, U. The benchmark simulation modelling platform—Areas of recent development and extension. *Lect. Notes Civ. Eng.* **2017**, *4*, 81–91. [CrossRef]
9. Marais, H.L.; Nordlander, E.; Thorin, E.; Wallin, C.; Dahlquist, E.; Odlare, M. Outlining process monitoring and fault detection in a wastewater treatment and reuse system. In Proceedings of the 2020 European Control Conference (ECC), St. Petersburg, Russia, 12–15 May 2020; pp. 558–563. [CrossRef]
10. Tao, E.P.; Shen, W.H.; Liu, T.L.; Chen, X.Q. Fault diagnosis based on PCA for sensors of laboratorial wastewater treatment process. *Chemom. Intell. Lab. Syst.* **2013**, *128*, 49–55. [CrossRef]
11. Qin, S.J. Statistical process monitoring: Basics and beyond. *J. Chemom.* **2003**, *17*, 480. [CrossRef]
12. Gertler, J. Survey of model-based failure detection and isolation in complex plants. *IEEE Contr. Syst. Mag.* **1988**, *8*, 3–11. [CrossRef]
13. Benveniste, A.; Basseville, M.; Moustakides, G. The asymptotic local approach to change detection and model validation. *IEEE Trans. Automat. Contr.* **1987**, *32*, 583–592. [CrossRef]
14. Isermann, R. Process fault detection based on modeling and estimation methods—A survey. *Automatica* **1984**, *20*, 387–404. [CrossRef]
15. Frank, P. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy—A survey and some new results. *Automatica* **1990**, *26*, 459–474. [CrossRef]
16. Yoon, S.; MacGregor, J. Fault diagnosis with multivariate statistical models, part i: Using steady state fault signatures. *J. Process Contr.* **2001**, *11*, 387–400. [CrossRef]
17. Ching, P.M.L.; So, R.H.Y.; Morck, T. Advances in soft sensors for wastewater treatment plants: A systematic review. *J. Water Process Eng.* **2021**, *44*, 102367. [CrossRef]
18. Kano, M.; Nakagawa, Y. Data-based process monitoring, process control, and quality improvement: Recent developments and applications in steel industry. *Comput. Chem. Eng.* **2008**, *32*, 12–24. [CrossRef]
19. Sanchez-Fernandez, A.; Fuente, M.J.; Sainz-Palermo, G.I. Fault detection in wastewater treatment plants using distributed PCA methods. In Proceedings of the 20th Conference on Emerging Technologies & Factory Automation (ETFA), Luxembourg, 8–11 September 2015; pp. 1–7. [CrossRef]
20. Andersson, S.; Hallgren, F. Sensor fault detection methods applied on dissolved oxygen sensors at a full scale WWTP. In Proceedings of the 9th IWA Symposium on Systems Analysis and Integrated Assessment (Watermatex 2015), Gold Coast, Australia, 14–17 June 2015.
21. Tarcsay, B.L.; Bárkányi, Á.; Chován, T.; Németh, S. A Dynamic Principal Component Analysis and Fréchet-Distance-Based Algorithm for Fault Detection and Isolation in Industrial Processes. *Processes* **2022**, *10*, 2409. [CrossRef]
22. Villegas, T.; Fuente, M.J.; Sainz-Palmero, G.I. Fault diagnosis in a wastewater treatment plant using dynamic independent component analysis. In Proceedings of the 18th Mediterranean Conference on Control and Automation (MED'10), Marrakech, Morocco, 23–25 June 2010; pp. 874–879. [CrossRef]
23. Xu, C.; Huang, D.; Cai, B.; Chen, H.; Liu, H. A complex-valued slow independent component analysis based incipient fault detection and diagnosis method with applications to wastewater treatment processes. *ISA Trans.* **2022**, *128*, 1–20. [CrossRef] [PubMed]
24. Yang, C.; Zhang, Y.; Huang, M.; Liu, H. Adaptive dynamic prediction of effluent quality in wastewater treatment processes using partial least squares embedded with relevance vector machine. *J. Clean. Prod.* **2021**, *314*, 128076. [CrossRef]
25. Chen, A.; Zhou, H.; An, Y.; Sun, W. PCA and PLS monitoring approaches for fault detection of wastewater treatment process. In Proceedings of the 25th International Symposium on Industrial Electronics (ISIE), Santa Clara, CA, USA, 8–10 June 2016; pp. 1022–1027. [CrossRef]
26. Liu, H.; Yang, J.; Zhang, Y.; Yang, C. Monitoring of wastewater treatment processes using dynamic concurrent kernel partial least squares. *Process Saf. Environ. Prot.* **2021**, *147*, 274–282. [CrossRef]
27. Kazemi, P.; Giralt, J.; Bengoa, C.; Steyer, J.P. Data-driven fault detection methods for detecting small-magnitude faults in anaerobic digestion process. *Water Sci. Technol.* **2020**, *81*, 1740–1748. [CrossRef]
28. Schraa, O.; Tole, B.; Copp, J.B. Fault detection for control of wastewater treatment plants. *Water Sci. Technol.* **2006**, *53*, 375–382. [CrossRef]
29. Yang, J.; Zhang, Y.; Zhou, L.; Zhang, F.; Jing, Y.; Huang, M.; Liu, H. Quality-related monitoring of papermaking wastewater treatment processes using dynamic multiblock partial least squares. *J. Bioresour. Bioprod.* **2022**, *7*, 73–82. [CrossRef]
30. Ma, X.; Zhang, Y.; Zhang, F.; Liu, H. Monitoring of papermaking wastewater treatment processes using t-distributed stochastic neighbor embedding. *J. Environ. Chem. Eng.* **2021**, *9*, 106559. [CrossRef]

31. Li, Z.; Yan, X. Ensemble model of wastewater treatment plant based on rich diversity of principal component determining by genetic algorithm for status monitoring. *Control Eng. Pract.* **2019**, *88*, 38–51. [CrossRef]
32. Dong, Y.; Qin, Y. A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. *J. Process Control* **2018**, *67*, 1–11. [CrossRef]
33. Corominas, L.; Villez, K.; Aguado, D.; Rieger, L.; Rosén, C.; Vanrolleghem, P.A. Performance evaluation of fault detection methods for wastewater treatment processes. *Biotechnol. Bioeng.* **2010**, *108*, 333–344. [CrossRef] [PubMed]
34. Garcia-Alvarez, D. Fault detection using principal component analysis (PCA) in a wastewater treatment plant (WWTP). In Proceedings of the International Student's Scientific Conference, Online Conference, 15 January 2009; pp. 55–60.
35. Zhou, J.; Huang, F.; Shen, W.; Liu, Z.; Corriou, J.-P.; Seferlis, P. Sub-period division strategies combined with multiway principal component analysis for fault diagnosis on sequence batch reactor of wastewater treatment process in paper mill. *Process Saf. Environ. Prot.* **2021**, *146*, 9–19. [CrossRef]
36. Duda, R.O.; Hart, P.E.; Stork, D.G. *Pattern Classification*; Wiley: New York, NY, USA, 2000; ISBN 978-0-471-05669-0.
37. Fuente, M.; Garcia, G.; Sainz, G. Fault diagnosis in a plant using fisher discriminant analysis. In Proceedings of the 16th Mediterranean Conference on Control and Automation, Ajaccio, France, 25–27 June 2008; pp. 53–58. [CrossRef]
38. He, Q.; Qin, S.; Wang, J. A new fault diagnosis method using fault directions in fisher discriminant analysis. *AIChE J.* **2005**, *51*, 555–571. [CrossRef]
39. Chiang, L.H.; Russell, E.L.; Braatz, R.D. Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares, and principal component analysis. *Chemom. Intell. Lab. Syst.* **2000**, *50*, 243–252. [CrossRef]
40. Chiang, L.H.; Russell, E.L.; Braatz, R.D. *Fault Detection and Diagnosis in Industrial Systems*; Springer Science & Business Media: London, UK, 2001; ISBN 1-85233-327-8.
41. Du, Z.; Jin, X. Multiple faults diagnosis for sensors in air handling unit using Fisher discriminant analysis. *Energy Convers. Manag.* **2008**, *49*, 3654–3665. [CrossRef]
42. Lee, C.; Choi, S.W.; Lee, I.B. Sensor fault diagnosis in a wastewater treatment process. *Water Sci. Technol.* **2006**, *53*, 251–257. [CrossRef]
43. Li, X.; Chai, W.; Liu, T.; Qiao, J. Fault detection of dissolved oxygen sensor in wastewater treatment plants. In Proceedings of the 46th Annual Conference of the IEEE Industrial Electronics Society, Singapore, 18–21 October 2020; pp. 225–230. [CrossRef]
44. Chistiakova, T.; Zambrano, J.; Samuelsson, O.; Carlsson, B. Binary classifiers applied to detect DO sensors faults during washing events. In Proceedings of the 2nd New Developments in IT & Water, Rotterdam, The Netherlands, 8–10 February 2015. [CrossRef]
45. Samuelsson, O.; Bjork, A.; Zambrano, J.; Carlsson, B. Fault signatures and bias progression in dissolved oxygen sensors. *Water Sci. Technol.* **2018**, *78*, 1034–1044. [CrossRef] [PubMed]
46. Mali, B.; Laskar, S.H. Incipient fault detection of sensors used in wastewater treatment plants based on deep dropout neural network. *SN Appl. Sci.* **2020**, *2*, 2121. [CrossRef]
47. Luca, A.-V.; Simon-Várhelyi, M.; Mihály, N.-B.; Cristea, V.-M. Data driven detection of different dissolved oxygen sensor faults for improving operation of the WWTP control system. *Processes* **2021**, *9*, 1633. [CrossRef]
48. Focht, D.D.; Chang, A.C. Nitrification and denitrification processes related to waste water treatment. *Adv. Appl. Microbiol.* **1975**, *19*, 153–186. [CrossRef] [PubMed]
49. Simon-Varhelyi, M.; Cristea, V.-M.; Luca, A.-V. Reducing energy costs of the wastewater treatment plant by improved scheduling of the periodic influent load. *J. Environ. Manag.* **2020**, *262*, 110294. [CrossRef] [PubMed]
50. Nair, A.; Cristea, V.-M.; Agachi, P.S.; Brehar, M. Model calibration and feed-forward control of the wastewater treatment plant—Case study for Cluj-Napoca WWTP. *Water Environ. J.* **2018**, *32*, 164–172. [CrossRef]
51. Otterpohl, R.; Freund, M. Dynamic models for clarifiers of activated sludge plants with dry and wet weather flows. *Water Sci. Technol.* **1992**, *26*, 1391–1400. [CrossRef]
52. Takacs, I.; Patry, G.G.; Nolasco, D. A dynamic model of the clarification-thickening process. *Water Res.* **1991**, *25*, 1263–1271. [CrossRef]
53. Varhelyi, M.; Cristea, V.-M.; Brehar, M.; Nemes, E.D.; Nair, A. WWTP model calibration based on different optimization approaches. *Environ. Eng. Manag. J.* **2019**, *18*, 1657–1670. [CrossRef]
54. Cristea, V.-M. Counteracting the accidental pollutant propagation in a section of the river Somes by automatic control. *J. Environ. Manag.* **2013**, *128*, 828–836. [CrossRef]
55. Viktoryova, N.; Szarka, A. Hrouzkova, Recent developments and emerging trends in paint industry wastewater treatment methods. *Appl. Sci.* **2022**, *12*, 10678. [CrossRef]
56. Revollar, S.; Vilanova, R.; Vega, P.; Francisco, M.; Meneses, M. Wastewater treatment plant operation: Simple control schemes with a holistic perspective. *Sustainability* **2020**, *12*, 768. [CrossRef]
57. Viveros, P.; Miqueles, L.; Mena, R.; Kristjanpoller, F. Opportunistic strategy for maintenance interventions planning: A case study in a wastewater treatment plant. *Appl. Sci.* **2021**, *11*, 10853. [CrossRef]
58. Ostace, G.S.; Cristea, V.-M.; Agachi, P.S. Extension of activated sludge model no. 1 with two-step nitrification and denitrification processes for operation improvement. *Environ. Eng. Manag. J.* **2011**, *10*, 1529–1544. [CrossRef]
59. Schneider, M.Y.; Carbajal, J.P.; Furrer, V.; Sterkele, B.; Maurer, M.; Villez, K. Beyond signal quality: The value of unmaintained pH, dissolved oxygen, and oxidation-reduction potential sensors for remote performance monitoring of on-site sequencing batch reactors. *Water Res.* **2019**, *161*, 639–651. [CrossRef]

60. Teh, H.Y.; Kempa-Liehr, A.W.; Wang, K.I. Sensor data quality: A systematic review. *J. Big Data* **2020**, *7*, 11. [CrossRef]
61. Rosen, C.; Rieger, L.; Jeppsson, U.; Vanrolleghem, P.A. Adding realism to simulated sensors and actuators. *Water Sci. Technol.* **2008**, *57*, 337–344. [CrossRef]
62. Lv, Z.; Shan, X.; Xiao, X.; Cai, R.; Zhang, Y.; Jiao, N. Excessive greenhouse gas emissions from wastewater treatment plants by using the chemical oxygen demand standard. *Sci. China Earth Sci.* **2022**, *65*, 87–95. [CrossRef]
63. Nguyen, T.K.L.; Ngo, H.H.; Guo, W.; Nguyen, T.L.H.; Chang, S.W.; Nguyen, D.D.; Varjani, S.; Lei, Z.; Deng, L. Environmental impacts and greenhouse gas emissions assessment for energy recovery and material recycle of the wastewater treatment plant. *Sci. Total Environ.* **2021**, *784*, 147135. [CrossRef]
64. Mannina, G.; Ekama, G.; Caniani, D.; Cosenza, A.; Esposito, G.; Gori, R.; Garrido-Baserba, M.; Rosso, D.; Olsson, G. Greenhouse gases from wastewater treatment—A review of modelling tools. *Sci. Total Environ.* **2016**, *551–552*, 254–270. [CrossRef]
65. Szulc, P.; Kasprzak, J.; Dymaczewski, Z.; Kurczewski, P. Life cycle assessment of municipal wastewater treatment processes regarding energy production from the sludge line. *Energies* **2021**, *14*, 356. [CrossRef]
66. Raghuvanshi, S.; Bhakar, V.; Sowmya, C.; Sangwan, K.S. Waste water treatment plant life cycle assessment: Treatment process to reuse of water. *Procedia CIRP* **2017**, *61*, 761–766. [CrossRef]
67. Corominas, L.; Foley, J.; Guest, J.S.; Hospido, A.; Larsen, H.F.; Morera, S.; Shaw, S. Life cycle assessment applied to wastewater treatment: State of the art. *Water Res.* **2013**, *47*, 5480–5492. [CrossRef] [PubMed]
68. Vallero, D.A. Air pollution biogeochemistry. In *Air Pollution Calculations*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 175–206. ISBN 978-0-12-814934-8. [CrossRef]
69. Listowski, A.; Ngo, H.H.; Guo, W.S.; Vigneswaran, S.; Shin, H.S.; Moon, H. Greenhouse gas (GHG) emissions from urban wastewater system: Future assessment framework and methodology. *J. Water Sustain.* **2011**, *1*, 113–125.
70. Prendez, M.; Lara-González, S. Application of strategies for sanitation management in wastewater treatment plants in order to control/reduce greenhouse gas emissions. *J. Environ. Manag.* **2008**, *88*, 658–664. [CrossRef] [PubMed]
71. Intergovernmental Panel on Climate Change (IPCC). Guidelines for National Greenhouse Gas Inventories. In *Intergovernmental Panel on Climate Change*; IPCC Guidelines; Table 6.11; Intergovernmental Panel on Climate Change (IPCC): Geneva, Switzerland, 2006; p. 5. Available online: http://www.ipcc-nggip.iges.or.jp/public/2006gl/index.html (accessed on 14 December 2022).
72. Gori, R.; Jiang, L.-M.; Sobhani, R.; Rosso, D. Effects of soluble and particulate substrate on the carbon and energy footprint of wastewater treatment processes. *Water Res.* **2011**, *45*, 5858–5872. [CrossRef]
73. Akbarjon, N.; Huang, F.; Shen, W. Calculation of N_2O emissions in the wastewater treatment process of paper mill. In Proceedings of the 2nd International Conference on Sustainable Energy, Environment and Information Engineering (SEEIE 2019), Beijing, China, 24–25 March 2019. [CrossRef]
74. Baresel, C.; Andersson, S.; Yang, J.; Andersen, M.H. Comparison of nitrous oxide (N_2O) emissions calculations at a swedish wastewater treatment plant based on water concentrations versus off-gas concentrations. *Adv. Clim. Chang. Res.* **2016**, *7*, 185–191. [CrossRef]
75. IEA (International Energy Agency). Projected Costs of Generating Electricity. Available online: https://www.iea.org/reports/projected-costs-of-generating-electricity-2020 (accessed on 14 December 2022).