



Comparing statistical process control charts for fault detection in wastewater treatment

H. L. Marais , V. Zaccaria and M. Odlare 

Future Energy Center, Mälardalen University, Västerås SE-721 23, Sweden

*Corresponding author. E-mail: heidi.marais@mdh.se

 HLM, 0000-0002-3097-459X; MO, 0000-0001-5480-0167

ABSTRACT

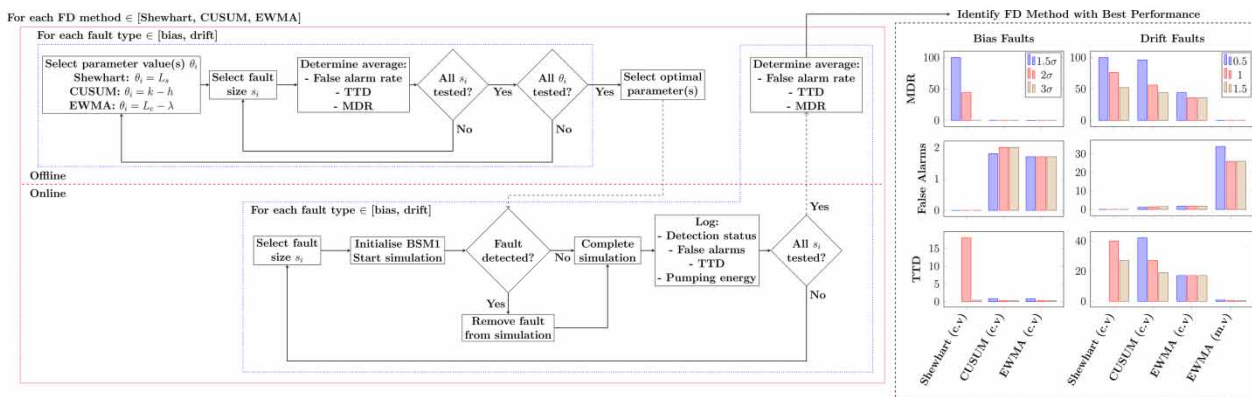
Fault detection is an important part of process supervision, especially in processes where there are strict requirements on the process outputs like in wastewater treatment. Statistical control charts such as Shewhart charts, cumulative sum (CUSUM) charts, and exponentially weighted moving average (EWMA) charts are common univariate fault detection methods. These methods have different strengths and weaknesses that are dependent on the characteristics of the fault. To account for this the methods in their base forms were tested with drift and bias sensor faults of different sizes to determine the overall performance of each method. Additionally, the faults were detected using two different sensors in the system to see how the presence of active process control influenced fault detectability. The EWMA method performed best for both fault types, specifically the drift faults, with a low false alarm rate and good detection time in comparison to the other methods. It was shown that decreasing the detection time can effectively reduce excess energy consumption caused by sensor faults. Additionally, it was shown that monitoring a manipulated variable has advantages over monitoring a controlled variable as set-point tracking hides faults on controlled variables; lower missed detection rates are observed using manipulated variables.

Key words: fault detection, process efficiency, process monitoring, statistical control charts, univariate, wastewater treatment

HIGHLIGHTS

- The best fault detection performance was obtained with the EWMA chart.
- Manipulated variable monitoring improves controlled variable sensor fault detection.
- Fault detection in wastewater treatment processes can improve the energy efficiency.

GRAPHICAL ABSTRACT



INTRODUCTION

Process monitoring is an important part of all industrial processes. Often process outputs are constrained by quality or safety requirements and continuous monitoring is necessary to ensure that those standards are met (Isermann 1997). In wastewater

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

treatment plants (WWTPs) this is becoming more relevant as environmental regulations result in more stringent requirements on the quality of the effluent (Olsson *et al.* 2005). An important part of process monitoring is fault detection (FD), as it involves identifying whether a process has deviated from the normal operating state (Rosen *et al.* 2003). The development and implementation of reliable FD methods has the potential to improve process control by ensuring that the control system has access to reliable data (Olsson & Newell 1999). This is vital for WWTPs that have historically lagged behind other industrial processes with regards to control and automation due to the push to move from treatment facilities to water resource recovery facilities; more precise and flexible control of the process will be required.

Available FD methods vary greatly in complexity (Venkatasubramanian *et al.* 2003), however, a common technique is to place limits on certain variables and declare a fault when the variable exceeds the limit (Chiang *et al.* 2000). This is called limit sensing and is a traditional process monitoring method. Limit sensing is not particularly sensitive or robust because it ignores interactions between different variables, this makes it a univariate technique (Chiang *et al.* 2000). When considering FD in an industrial process, like the WWTP, there are advantages to investigating simple univariate FD methods despite the high level of complexity of the system (Marais *et al.* 2020). Some of the primary benefits of evaluating univariate methods are to obtain a benchmark to evaluate the performance of more advanced and costly FD methods against, as well as in their potential to be combined with advanced methods. Furthermore, the use of automated FD in industrial WWTPs is limited; optimising and demonstrating the capabilities of simple univariate methods is an important step to increase the use of FD methods in industrial processes due to the slow acceptance rate of more advanced methods.

Common univariate FD methods are often based on statistical process control and include, but are not limited to, Shewhart charts, cumulative sum (CUSUM) charts, and exponentially weighted moving average (EWMA) charts (Chiang *et al.* 2000). While numerous modifications have been suggested to these chart-based methods, their base forms are more commonly seen in applications to WWTPs. Some of the earliest investigations of these methods applied to WWTPs include that of Thomann *et al.* (2002) and Thomann (2008), who suggested the use of the Shewhart control chart to give warnings and alarms for out-of-control behaviour in a WWTP for an ammonium sensor and a suspended solids sensor, as well as a nitrate and phosphate sensor. Schraa *et al.* (2006), however, claimed that the Shewhart and CUSUM charts should perform poorly for the WWTP due to seasonal effects and non-constant variance, and advocated for the use of the EWMA control chart due to its good performance for a slow varying mean.

A more thorough comparison is presented by Corominas *et al.* (2011), comparing the Shewhart chart, standard EWMA, EWMA performed on residuals, and other EWMA variants, utilising the Benchmark Simulation Model No.1 long-term (BSM1_LT) platform. They investigated both bias and drift sensor faults on dissolved oxygen sensors, one in a control loop and one not. Using a proposed evaluation index based on several factors (such as the number of false alarms, the missed detection rate (MDR), and the time to detection (TTD)) they found that the EWMA chart was prone to high false alarm counts and that better performance was achieved with the EWMA calculated on residuals and when monitoring the actuator instead of the controlled variable.

The use of control charts on residuals has been a popular application of the methods, with Baklouti *et al.* (2017), Baklouti *et al.* (2018b), and Baklouti *et al.* (2021) utilising EWMA charts, or variations there of (maximum adaptive or maximum double adaptive), on residuals calculated from state estimation. They used the Benchmark Simulation Model No.1 (BSM1) process and tested bias type sensor faults on a chemical oxygen demand sensor (Baklouti *et al.* 2017) and drift and bias faults on a dissolved oxygen sensor (Baklouti *et al.* 2018b, 2021). They too compared the results based on the MDR, the number of false alarms, and the TTD. Using the same BSM1 layout and sensors, Baklouti *et al.* (2018a) investigated Shewhart charts, Pareto optimised EWMA, and multiscale optimised EWMA also on calculated residuals. Their results showed high MDRs obtained with the Shewhart method coupled with fewer false alarms, and overall good results for the EWMA charts.

Sánchez-Fernández *et al.* (2018) considered dissolved oxygen sensor faults in the Benchmark Simulation Model No.2 (BSM2) and found that the univariate EWMA chart on residuals performed as well, or better, than multivariate monitoring using principal component analysis (PCA) with T^2 and Q statistics. Similarly, Kazemi *et al.* (2021) reached the same conclusion but for the univariate CUSUM chart when monitoring residuals; they tested a process fault, as well as bias and drift faults on a pH sensor, and showed that the CUSUM chart outperformed the multivariate monitoring for small fault sizes, and in the cases with missing data. Finally, Riss *et al.* (2021) used a modified CUSUM chart, based on the median, and combined it with the random forest method to successfully perform fault detection on real data from a WWTP.

Evidently these univariate chart-based methods have potential to be used in WWTPs, however, none of the aforementioned studies provide a rigorous comparison of the different methods for a range of fault types and sizes. These methods are known

to perform differently depending on the type of fault that is occurring; for example, theoretically, CUSUM and EWMA charts are good at detecting small and recurring changes in the data, while Shewhart charts are better for detecting large and sudden shifts in the data (Montgomery 2009). As a result, it is difficult to say which of these methods is best suited for use in process monitoring and FD in a WWTP.

In this study the goal was to evaluate which of these simple FD strategies performs best for specific faults, as well as for the general case. To achieve this a range of faults and fault sizes were tested, using the BSM1 platform, to get an idea of the overall performance of each method. The three methods mentioned (Shewhart, CUSUM, and EWMA charts) were compared for two types of sensor faults, bias and drift, on the nitrate sensor in the second bioreactor in the BSM1 process. Bias and drift faults are common forms of sensor faults, where the former represents a sudden and sustained deviation from the true value while the latter is a gradual deviation and is generally harder to detect. Both the nitrate sensor in the second bioreactor (a controlled variable) and the flow rate of the internal recycle (a manipulated variable) were considered for process monitoring. In this investigation no modifications were made to the base methods, however, a systematic approach to optimal parameter selection for the methods was proposed. This work demonstrates the effectiveness of a simple univariate FD method and identifies potential areas for improvements and future work.

METHODS

Description of fault detection methods

In univariate FD, based on statistical process control charts, the effectiveness of the method depends on the upper and lower threshold values, or control limits, that are selected. These are chosen to minimise the rate of false alarms, to reduce the TTD, and to minimise the MDR. Naturally, there is a trade-off that must be made between these criteria. Each method (Shewhart, CUSUM, EWMA) has parameters that determine the threshold limits, the behaviour of the statistic being monitored. In some cases, there is also a dependence on historical data or knowledge about the expected values of the samples. These control chart methods perform best when the data is stationary; WWTPs are not completely stationary and have diurnal and seasonal variations. Diurnal variations can be taken into consideration with the variance, but addressing seasonal variations may require utilising different chart parameters at different times of the year (Olsson & Newell 1999). This is shown in the work of Corominas *et al.* (2011) who suggest adapting the control limits based on the temperature, or simply adopting summer and winter control limits. In this work only diurnal variations are considered as considering limits for different seasons should not affect the analysis of which method performs best given different fault types and sizes. The data required to determine optimal parameter sets is not extensive and should only be representative of the normal operating state of the process.

When considering the Shewhart control chart, the form for individual measurements is appropriate for the sensor monitoring problem. The control chart is constructed on the sensor output, x_i , has a centreline at the mean, μ_0 , and has upper and lower control limits (UCL , LCL) calculated as per Equations (1) and (2) (Montgomery 2009), where \overline{MR} is the average moving range for the data of interest, L_s is the limit size, and d_2 is a function of the sample size which, for the current application, is always equal to 1.128 (Montgomery 2009).

$$UCL = \mu_0 + L_s \frac{\overline{MR}}{d_2} \quad (1)$$

$$LCL = \mu_0 - L_s \frac{\overline{MR}}{d_2} \quad (2)$$

The mean and average moving range are based on the expected behaviour of the data, while the limit size can be adjusted to influence the number of false alarms and the MDR.

The CUSUM chart is particularly well suited to applications involving individual measurements, and also performs well when detecting small shifts in the variable of interest (Montgomery 2009). The tabular CUSUM (one of the more common forms) involves the summation of deviations of the samples from the target, both above and below a certain threshold. Because of the cumulative nature it incorporates more information about the behaviour of the sample than the Shewhart

method. The deviations, which are the statistics being monitored, are calculated in Equations (3) and (4) (Montgomery 2009),

$$C_i^+ = \max [0, x_i - (\mu_0 + K) + C_{i-1}^+] \quad (3)$$

$$C_i^- = \max [0, (\mu_0 - K) - x_i + C_{i-1}^-] \quad (4)$$

where K is called the reference, or allowance, value and relates to the size deviation that is desirable to detect, and $C_0^+ = C_0^- = 0$. The threshold value, H , defines the control limits. It is common practice to define both K and H as multiples of the standard deviation (σ) of the data: $K = k\sigma$ and $H = h\sigma$ (Montgomery 2009). Both C_i^+ and C_i^- are monitored which results in two lines on the CUSUM control chart; the UCL (H) can be violated by C_i^+ , or the LCL ($-H$) by C_i^- .

The final method is the EWMA method, this control chart is good for detecting small shifts and it is typically used with individual observations. The threshold values are defined according to Equations (5) and (6) (Montgomery 2009),

$$UCL = \mu_0 + L_e \sigma \sqrt{\left(\frac{\lambda}{2 - \lambda}\right) [1 - (1 - \lambda)^{2i}]} \quad (5)$$

$$LCL = \mu_0 - L_e \sigma \sqrt{\left(\frac{\lambda}{2 - \lambda}\right) [1 - (1 - \lambda)^{2i}]} \quad (6)$$

where λ , a smoothing parameter, is $\in (0, 1)$, and L_e represents the width of the control limits. The monitored statistic is shown in Equation (7),

$$z_i = \lambda x_i + (1 - \lambda) z_{i-1} \quad (7)$$

where $z_0 = \mu_0$ (Montgomery 2009). The parameters λ and L_e are tuned to determine the performance, with small values of λ detecting smaller deviations in the variable of interest but requiring narrower control limits (Montgomery 2009).

Performance evaluation criteria

As mentioned, the various control chart parameters affect the number of false alarms that occur, the MDR, and the TTD. It is necessary to define these criteria. Consider the case where a single fault occurs during a simulation, we define the interval within the control limits as $CL = [LCL, UCL]$, and the set of points outside the control limits as its complement \overline{CL} . When a control limit is exceeded by a monitored statistic, J_i , where J_i is x_i , C_i^+ , C_i^- , or z_i , this is equivalent to $J_i \in \overline{CL}$. We define v_i which evaluates if a control limit is violated by the monitored statistic at timestep

$$v_i = \mathbb{1}_{\overline{CL}}(J_i) \quad (8)$$

where $\mathbb{1}$ is the indicator function. We also define f_i which shows if a fault is occurring, and c_i which tracks if a new false alarm has occurred:

$$f_i = \begin{cases} 1 & \text{if fault occurring} \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$$c_i = \begin{cases} i & \text{if } v_i = 1 \text{ and } f_i = 0 \\ c_{i-1} & \text{otherwise} \end{cases} \quad (10)$$

where $c_1 = 0$. We define a false alarm as a control limit violation in the absence of a fault. The number of false alarms is the number of times a threshold is incorrectly crossed, two successive erroneous violations are thus considered a single false alarm. The number of false alarms for a simulation period can be calculated as

$$\text{Number of False Alarms} = \sum_{i=1} v_{i+1} (1 - \mathbb{1}_i(c_i)) \quad (11)$$

Since each simulation contains a single fault, as previously stated, any control limit violation within the duration of the fault indicates the detection of the fault. This corresponds to the set of correct detections, $F = \{i:v;f_i = 1\}$, being non-empty. If this set is empty, it indicates that the monitored statistic failed to violate a control limit during a fault event and the fault was undetected. The MDR is determined for several simulations as the percentage of simulations where the fault was undetected. Finally, the TTD is the time between the fault event occurring (which in a real fault case can be hard to pinpoint) and the time at which the monitored statistic correctly violates a control limit. It can be calculated as

$$TTD = \underset{i \in F}{\operatorname{argmin}} i - i_{\text{fault start}} \quad (12)$$

The TTD in the case where a fault is undetected is undefined.

WWTP used for simulations

The MATLAB Simulink implementation of the BSM1 was used to simulate the WWTP. This setup consists of five biological reactors placed in series and modelled with the Activated Sludge Model No. 1 (ASM1). Two of the reactors are treated as anoxic and three as aerated, resulting in a total anoxic volume of 2,000 m³ and a total aerated volume of 3,999 m³. Following the reactors is a non-reactive ten-layer settler, modelled with the Takács model, which has a volume of 6,000 m³. The model assumes a constant temperature of 15 °C. A detailed description of the BSM1 can be found in [Gernaey *et al.* \(2014\)](#). The layout is shown in [Figure 1](#).

The primary sensor of interest in this study was the nitrate sensor ($S_{NO,2}$) in the second biological reactor, as shown in [Figure 1](#). This is a standard sensor in the BSM1 layout and has a setpoint of 1 gN·m⁻³ in the control loop. All sensor faults investigated were introduced to this sensor, $S_{NO,2}$. An additional sensor was added to the BSM1 to measure the flow rate of the internal recycle, Q_a . This was a class A sensor with a measurement range of 0–1,00,000 m³·day⁻¹, based on the recommendations in [Gernaey *et al.* \(2014\)](#). Sensor noise was obtained from the [IWA Modelling & Integrated Assessment Specialist Group \(2020\)](#). Both the $S_{NO,2}$ and the Q_a sensors were used for process monitoring, with the focus being on the former.

All simulations were run using the provided dry-weather influent which simulates for a 14-day period. The simulations were initialised as per [Gernaey *et al.* \(2014\)](#).

Determination of optimal parameters

The simulation work was divided into two parts, where the first involved the calibration and optimisation of the parameters used in the control charts. For this, the BSM1 was simulated under normal operating conditions with no added faults, and the response of the $S_{NO,2}$ sensor recorded. Bias and drift faults were then added in post-processing to this pre-obtained sensor signal to generate a wide range of fault conditions in the $S_{NO,2}$ sensor that could be used to determine the parameters of the control chart FD methods.

The faults were added with start times between days 7 and 12 of the simulation, with durations between 4 and 48 h. Latin hypercube sampling was used to generate start times and durations within these limits for each fault size of interest. The bias

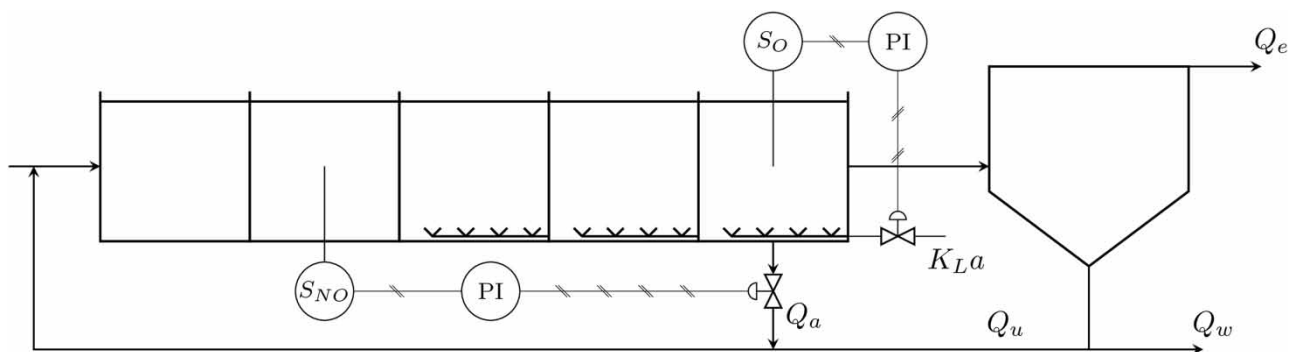


Figure 1 | The layout of the BSM1 WWTP, adapted from [Gernaey *et al.* \(2014\)](#), showing the two standard control loops.

fault sizes ranged from $1\sigma \text{ gN}\cdot\text{m}^{-3}$ to $5\sigma \text{ gN}\cdot\text{m}^{-3}$, where $\sigma = 0.565 \text{ gN}\cdot\text{m}^{-3}$, while the drift fault slopes were between $0.25 \text{ gN}\cdot\text{m}^{-3}\cdot\text{day}^{-1}$ and $3 \text{ gN}\cdot\text{m}^{-3}\cdot\text{day}^{-1}$.

The procedure was as follows and is shown in Figure 2:

- (1) For a given FD method, the relevant parameters were selected from a range.
- (2) A fault of specific type and size was added to the sensor data, with characteristics described in the preceding paragraph.
- (3) The FD method was used to detect faults on the generated faulty data.
- (4) The number of false alarms, TTD, and MDR were logged.
- (5) Steps 2–4 were repeated 20 times for the same fault type and size (different start times and durations).
- (6) The false alarms, MDR, and TTD were averaged for the given fault type, size, and parameter combination.
- (7) Fault size was adjusted and steps 2–6 repeated.
- (8) After iterating through all fault sizes and types, the parameters were changed and steps 1–7 repeated.

This was done for all three FD methods, where parameter ranges were selected based on recommended values from theory (Montgomery 2009):

- Shewhart: $1 \leq L_s \leq 5$ with 100 samples.
- CUSUM: $0.5 \leq k \leq 2$ and $1.5 \leq h \leq 5$ with 25 samples in each range.
- EWMA: $0.5 \leq L_e \leq 4.5$ and $0.05 \leq \lambda \leq 0.5$ with 25 samples in each range.

The results from this were used to select an optimal parameter combination for each FD method.

Online FD tests

The BSM1 was used for online FD tests with the control chart methods and a schematic of the procedure is given in Figure 3. Each simulation was initialised from the same starting point, and a unique fault added during the simulation. The faults were added with start times between days 7 and 12 of the simulation and with durations between 4 and 48 h, where again Latin hypercube sampling was used to generate sample points within these limits. The bias fault sizes tested were $1.5\sigma \text{ gN}\cdot\text{m}^{-3}$, $2\sigma \text{ gN}\cdot\text{m}^{-3}$, and $3\sigma \text{ gN}\cdot\text{m}^{-3}$, and the drift fault slopes were $0.5 \text{ gN}\cdot\text{m}^{-3}\cdot\text{day}^{-1}$, $1 \text{ gN}\cdot\text{m}^{-3}\cdot\text{day}^{-1}$, and $1.5 \text{ gN}\cdot\text{m}^{-3}\cdot\text{day}^{-1}$.

Blocks were added into the Simulink model to perform the FD, according to the descriptions of the methods provided previously, with the calibrated parameters. If the fault was correctly detected the fault was removed from the system at the time of detection to evaluate the effect which improved FD can have on the performance of the WWTP.

The dependent variables, or criteria for evaluation, were the number of false alarms that occurred, the MDR, and the TTD, as well as the signals from the sensors being considered and the pumping energy evaluation index built into the BSM1 platform. The pumping energy in BSM1 is calculated based on the flow rates of the external recycle, the internal recycle, and the waste sludge (Gernaey *et al.* 2014), it is given in $\text{kWh}\cdot\text{day}^{-1}$ and makes up part of the operational cost index in the BSM1 performance evaluation (Gernaey *et al.* 2014). The pumping energy is selected as an evaluation criterion due to its dependence on the internal recycle, Q_a , which is directly affected by faults in the nitrate sensor in the second biological reactor.

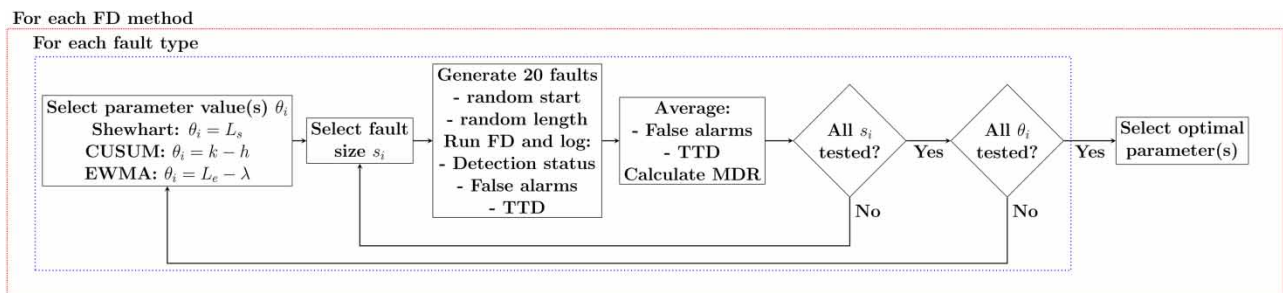


Figure 2 | Procedure for optimal parameter selection, performed for each FD method (Shewhart, CUSUM, EWMA) with two fault types (bias and drift).

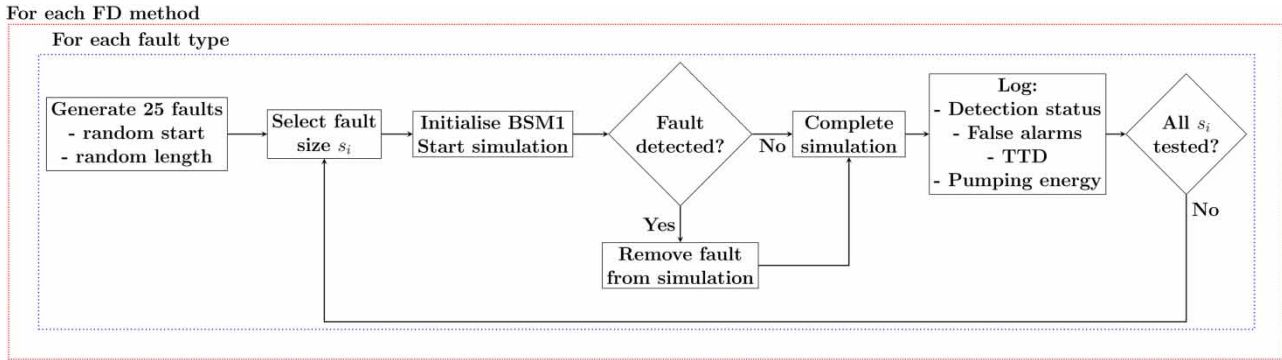


Figure 3 | Procedure for online testing of FD methods to determine an average performance for each fault size – fault type – FD method combination.

RESULTS AND DISCUSSION

Determination of optimal parameters

As mentioned, the determination of optimal parameters was done using the offline artificially generated faulty data. The faults were introduced to the $S_{NO,2}$ sensor in the second biological reactor. Figure 4 shows the results from the offline bias, and drift, fault tests using the Shewhart control chart for FD.

What is immediately evident is that attempting to minimise the TTD results in a rapid increase in the number of false alarms. As the aim is to minimise these two criteria, and the MDR, any point on the Pareto front with a low MDR is a candidate optimal solution for some trade-off between the criteria. It can also be noted that the drift faults have significantly longer TTD values (unnormalised bounds of 0.45–20.2 h compared to 0.045–3.7 h for bias faults), as well as higher MDRs. The false alarms for both methods vary between 0 and 135 when unnormalised.

Three points are marked on Figure 4: the points which have the smallest distance to the origin for the bias faults (upward facing triangle) and the drift faults (downward facing triangle) respectively, the right facing triangle is the point using the standard limit size. It is perhaps not unexpected that the different fault types have different optimal parameters. The limit sizes (L_s) that correspond to these three points are 2.37, 2.05, and 3, respectively. While the smaller limit sizes give shorter detection times, it was decided to proceed with the limit size of 3 due to the lower false alarm rate without sacrificing significantly on the MDR (0% for bias faults, 4% for drift faults). The Shewhart chart is well established and commonly used with a limit size of 3 which, by selecting this limit size, makes it a baseline method that can be used for comparison.

A similar line of reasoning can be followed for selecting the optimal parameters for the CUSUM chart, which has results shown in Figure 5. The limits of the unnormalised data are 0–27 and 0–30 false alarms for bias and drift faults

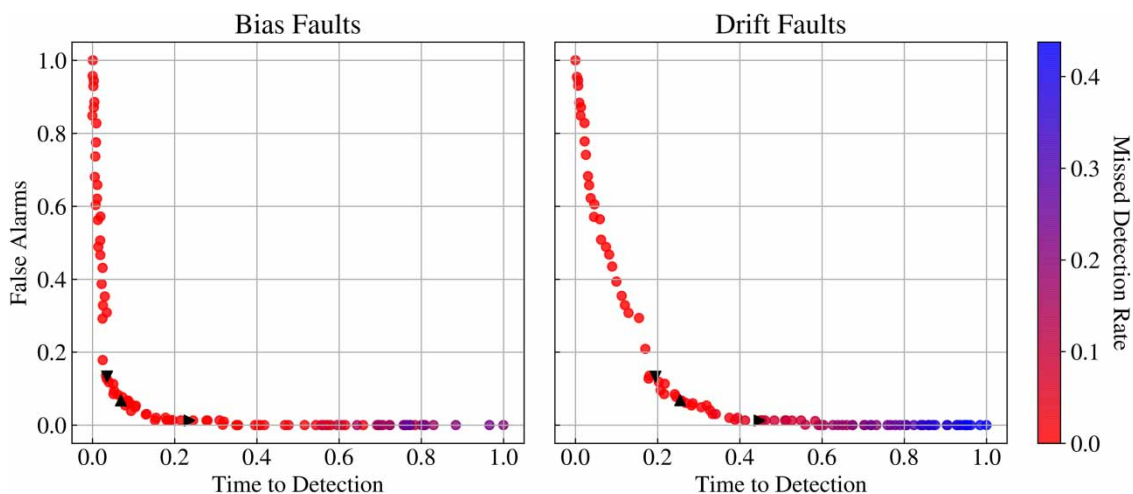


Figure 4 | Normalised performance of offline Shewhart FD.

respectively, and 0.21–3.1 h and 1.2–17.7 h for the detection times for the bias and drift faults. The upward and downward facing triangles have the same meaning as before, while the right facing triangle represents a third parameter set. The reference and threshold values, k and h , corresponding to these points are 1–2.38, 0.62–4.85, and 1.5–1.6, respectively (written as k – h).

An interesting observation is the significantly higher MDR rate for the drift faults when compared to the bias faults; decreasing the size of the threshold value (h) appears to have the largest effect on the MDR considering the three highlighted points. This suggests that knowing what sizes of faults are dangerous or costly is important to optimal calibration of the control chart. However, as before, reducing the number of false alarms clearly comes at the cost of increasing the MDR. The point closest to the origin in the case of the drift faults has a 2.5% MDR, with 4 false alarms on average (unnormalized), while the third parameter set has a 6% MDR with on average 0.85 false alarms (unnormalized); the false alarms decreased by a factor of 4.7 while the MDR increased by a factor of 2.4. For this reason, the third parameter set, 1.5–1.6, was selected for this method.

Finally, the EWMA chart results are shown in Figure 6. An interesting difference can be observed compared to Figures 4 and 5; the TTD appears to be able to reach a minimum while the number of false alarms remains acceptably low. This is slightly deceptive, however, due to the normalisation. In Figure 5 the absolute values for maximum number of false alarms were 27 and 30 for bias and drift respectively, while in Figure 6 the absolute value is 125 for both fault cases, which means that the 0.3 point in Figure 6 corresponds to 37.5 false alarms, and is significantly higher than what was obtained for the CUSUM approach. The unnormalised detection times have bounds of 0.018–2.5 h for bias faults, and 0.13–14.6 h for drift faults.

The three parameter sets for the upward, downward, and right facing triangles correspond to limit size – smoothing factor pairs (L_e – λ) of 2.5–0.48, 1.33–0.05, and 3.05–0.4. The effects that the parameters have on the performance for detecting bias faults is significantly lower than on the drift faults. Due to the significantly higher rate of false alarms when compared to the previous methods the third parameter set was chosen as it reduced the false alarms without reaching high MDRs or unacceptable TTDs. In all three cases, it was prioritised to reduce the number of false alarms over minimising the TTD. This choice was motivated by the relatively slow dynamics in a WWTP, while false alarms can disrupt operations and cause unnecessary cost to the process.

To summarise the parameter selection and the performance metrics obtained:

- Shewhart: $L_s = 3$, False alarms = 1.75 (bias and drift), TTD=0.825 h (bias), 10.2 h (drift), MDR=0% (bias), 4.25% (drift).
- CUSUM: $k = 1.5$, $h = 1.6$, False alarms=0.830 (bias), 0.845 (drift), TTD=0.903 h (bias), 10.5 h (drift), MDR=0% (bias), 6% (drift).
- EWMA: $L_e = 3.05$, $\lambda = 0.4$, False alarms=4.72 (bias), 4.75 (drift), TTD=0.569 h (bias), 8.30 h (drift), MDR=0% (bias), 2.25% (drift).

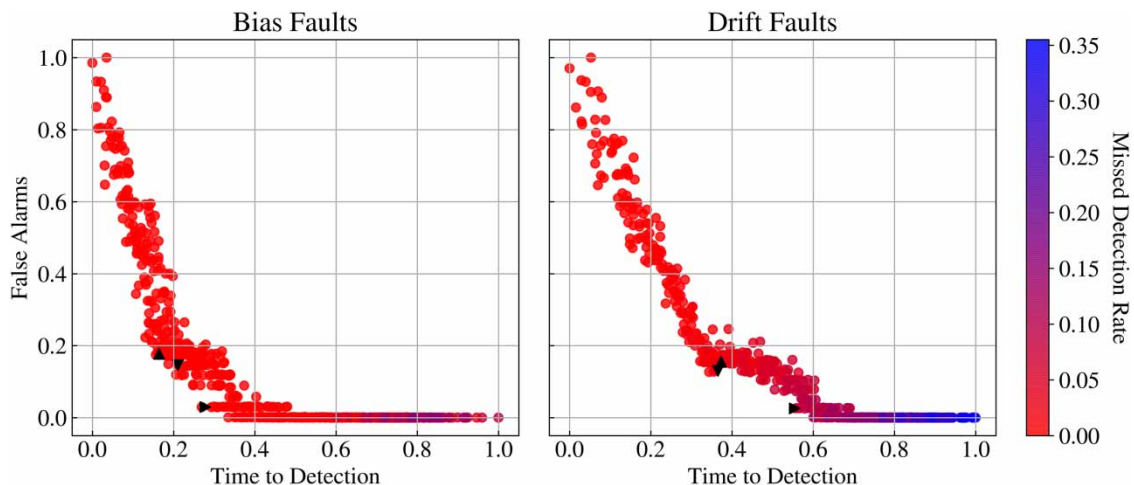


Figure 5 | Normalised performance of offline CUSUM FD.

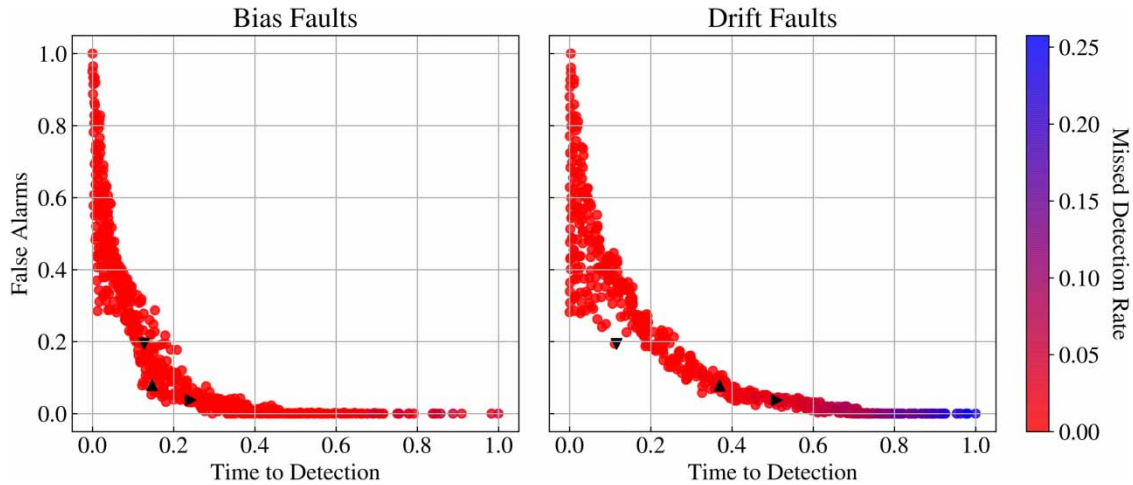


Figure 6 | Normalised performance of offline EWMA FD.

In these offline tests the EWMA method had the shortest detection times and lowest missed detection rates, while the CUSUM method had the fewest false alarms.

Online FD: monitoring the nitrate sensor

The parameters selected in the prior section were used in the online FD tests for monitoring the nitrate sensor. In Figure 7 examples of the control charts are shown for the two types of faults. Note that the faults are different due to the random start and end times. In the Shewhart method the statistic that is monitored is the sensor output itself, as discussed previously; this results in a noisy output shown in the Shewhart charts in Figure 7 due to the absence of additional filtering of the data. The CUSUM chart shows both the positive and negative statistics (C_i^+ and C_i^-) on the same plot and is difficult to interpret due to sudden spikes in the statistic when faults are detected, note that the y-axis is bounded as the absolute values of the statistics are not of interest, only the period for which it is above or below the threshold values. In the CUSUM chart the size and severity of the fault is unclear from the chart, however, the simplicity and the few numbers of false alarms produced is appealing. The EWMA method appears to be a good compromise between the aforementioned two charts, due to the ability to adjust the noise and rate of response with the smoothing factor but being able to visually judge the severity of the fault from the chart.

The analysis of the results will focus on the three performance evaluation criteria. As mentioned in the descriptions of the simulations, three fault sizes were tested for each fault type. Table 1 shows the results for the bias faults.

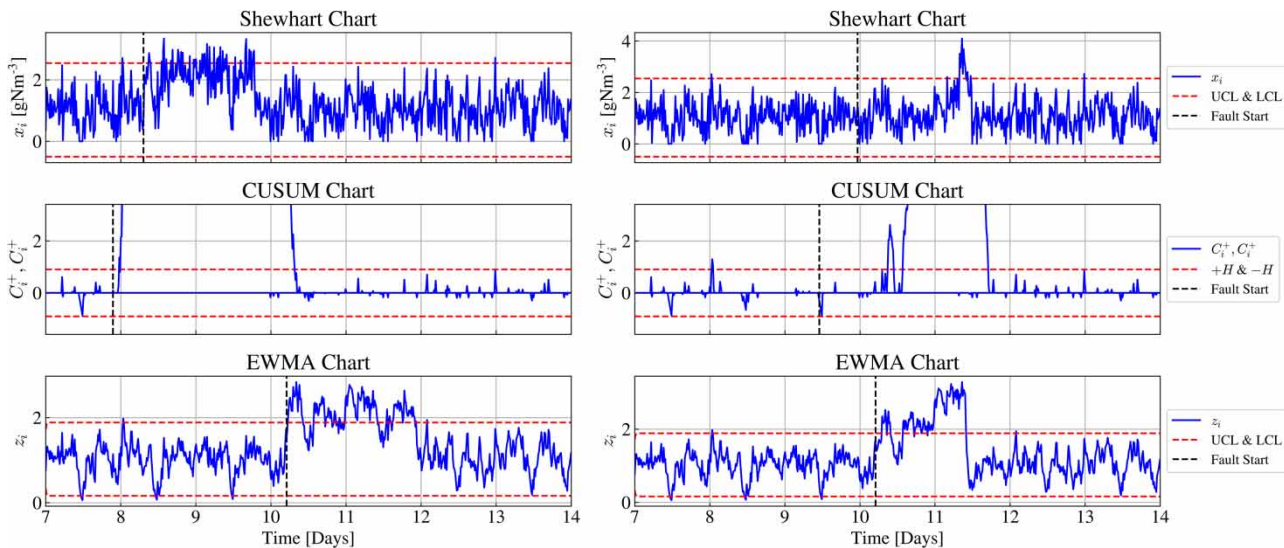


Figure 7 | Examples of control charts for online FD of left, bias faults and right, drift faults.

Table 1 | Average performance for online FD tests with bias faults

	Shewhart			CUSUM			EWMA		
	1.5σ	2σ	3σ	1.5σ	2σ	3σ	1.5σ	2σ	3σ
Fault size [$\text{gN}\cdot\text{m}^{-3}$]	1.5σ	2σ	3σ	1.5σ	2σ	3σ	1.5σ	2σ	3σ
MDR [%]	100	44	0	0	0	0	0	0	0
False alarms	0.0	0.0	0.0	1.8	2.0	2.0	1.7	1.7	1.7
TTD [hours]	NA	18	0.4	0.85	0.25	0.25	0.8	0.25	0.25

The Shewhart chart, despite its good performance in the offline tests, had extremely large MDRs for the two smaller fault sizes, failing to detect any faults with a bias size of $1.5\sigma \text{ gN}\cdot\text{m}^{-3}$. However, the CUSUM and EWMA methods had no missed detections for the bias faults, as expected based on the results of the offline tests, and were able to successfully detect all the faults.

The method with the highest MDR corresponds to that with the lowest number of false alarms and indicates that the limit selected for the Shewhart chart was perhaps too conservative for the actual behaviour of the data. The false alarms of the CUSUM and EWMA charts are satisfactorily low. What is surprising, however, is that the EWMA had fewer false alarms than the CUSUM method, when operating online. This is contradictory to the expected behaviour based on the offline tests. Finally, to address the TTD, it should be noted that 0.25 h is the sampling rate in the BSM1 simulation platform and therefore the reported TTD cannot be lower than this. It is likely that the CUSUM and EWMA charts would successfully detect the larger bias faults in a shorter time were a higher sampling rate used. However, due to the relatively slow dynamics in a WWTP and the appreciable sensor noise, a higher sampling rate would not provide much benefit to the control chart analysis as currently presented. The benefits of a higher sampling rate would be in the potential to apply filtering and smoothing algorithms to the sensor signal.

The results from the drift FD tests are shown in Table 2, and look drastically different to those of the bias faults. The main difference is seen in the high MDRs across all chart types, along with the long TTDs. In the offline FD tests the highest MDR was 6%, whereas in Table 2 the lowest value is 36%. This can be explained by the presence of active online control that was present in the simulations. The nitrate sensor that was monitored, and was the faulty sensor, is the controlled variable in one of the two default BSM1 control loops. This results in the sensor output being pushed back to the set-point value and, in the case of the slow drift faults, this happens before a control limit is crossed and an alarm triggered.

Based on what is seen in Table 2, the EWMA method appears best suited to the drift faults as it can consistently detect more than 50% of the fault cases, a claim none of the other methods can make. It is not unexpected that the number of false alarms are similar in magnitude to those seen in the bias fault tests – these are dependent on the parameters of the FD methods and the variations in the data, not the faults that are occurring. The long TTDs are likely due to the faults reaching a size where the controller can no longer compensate for them, and a limit is then violated. This is suboptimal as at this stage the fault has already had a large effect, likely negative, on the process. Quicker detection of these drift faults is essential. An interesting observation, however, is the apparent lack of dependence on fault size that the EWMA has in its TTD. This is a desirable characteristic of a FD method.

From the above discussion it can be concluded that for the selected parameters used in the control charts, the EWMA FD method outperforms the other two methods for the tested faults, bias and drift, and sizes.

Online FD: effect of control

As mentioned, the presence of online control had a large effect on the ability to detect faults in the sensor measuring the controlled variable. Figure 8 illustrates the controller concealing such a sensor fault; the top plot shows the sensor output, and

Table 2 | Average performance for online FD tests with drift faults

	Shewhart			CUSUM			EWMA		
	0.5	1	1.5	0.5	1	1.5	0.5	1	1.5
Fault size [$\text{gN}\cdot\text{m}^{-3}\cdot\text{day}^{-1}$]	0.5	1	1.5	0.5	1	1.5	0.5	1	1.5
MDR [%]	100	76	52	96	56	44	44	36	36
False alarms	0.0	0.0	0.0	1.1	1.4	1.6	1.7	1.8	1.7
TTD [hours]	NA	40	27	42	27	19	17	17	17

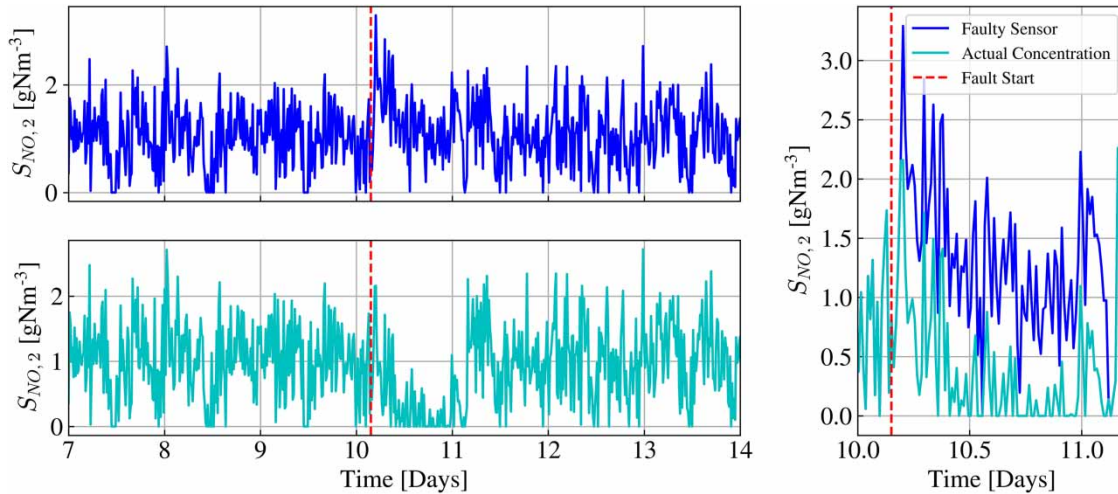


Figure 8 | Example of the controller concealing the sensor fault. Top: faulty sensor output, bottom: actual concentration in reactor, right: both outputs for the duration of the fault.

while there is a small spike, the controller quickly addresses the deviation from the set-point. The bottom plot shows the actual concentration in the reactor, which has deviated from the desired value due to the sensor fault.

An obvious step to take, based on these results, is to consider how the results change if the manipulated variable is monitored, with the goal of detecting concealed faults in the controlled variable sensor. The manipulated variable, in this case, is the internal recycle flow rate (Q_a). The behaviour of this variable is more cyclic in nature than that of the $S_{NO,2}$ sensor, and so it is less suited for use with control charts. However, to illustrate the potential improvement in MDR that can be gained, simulations were performed for the drift faults monitoring Q_a with the EWMA method. The EWMA method was not re-optimised with the same meticulousness for the new variable, due to the reliance on active process control for this FD scheme to work the offline optimisation could not be done in the same way. However, for deviations in the manipulated variable itself it was observed that the same parameter set, $L_e-\lambda$ of 3.05–0.4, provided a reasonable trade-off between the false alarms and TTD while minimising the MDR. This supported the use of this parameter set for these preliminary tests. Table 3 shows the results of the online FD tests, comparing to the prior results from monitoring the nitrate concentration.

It is clear from Table 3 that, despite the higher number of false alarms, monitoring the manipulated variable is significantly more effective at detecting drift faults in the controlled variable sensor. All the faults were successfully detected, and the TTD was significantly lower than when monitoring the controlled variable.

Effect of faults on process performance

To motivate the need for good FD, the effect of the $S_{NO,2}$ sensor faults on the pumping energy (built in BSM1 performance evaluation index) is considered. This index was selected as the internal recycle flow rate is significantly affected by sensor faults in the $S_{NO,2}$ sensor and is a major factor in the pumping energy index. The changes in the pumping energy from the baseline value are shown versus the time taken to detect a particular fault, for varying fault sizes, in Figure 9.

From the Figure it can be seen that the duration of the fault, rather than its size, determines the effect that it will have on the process. Of course, this is not always the case, but for the sensor faults considered it is the dominant factor. The smaller fault

Table 3 | Average performance for online FD tests with drift faults using the EWMA method, comparing monitoring the manipulated variable (Q_a) to the controlled variable ($S_{NO,2}$)

	Monitor Q_a			Monitor $S_{NO,2}$		
Fault size [$\text{gN}\cdot\text{m}^{-3}\cdot\text{day}^{-1}$]	0.5	1	1.5	0.5	1	1.5
MDR [%]	0	0	0	44	36	36
False alarms	34	26	26	1.7	1.8	1.7
TTD [hours]	0.80	0.35	0.35	17	17	17

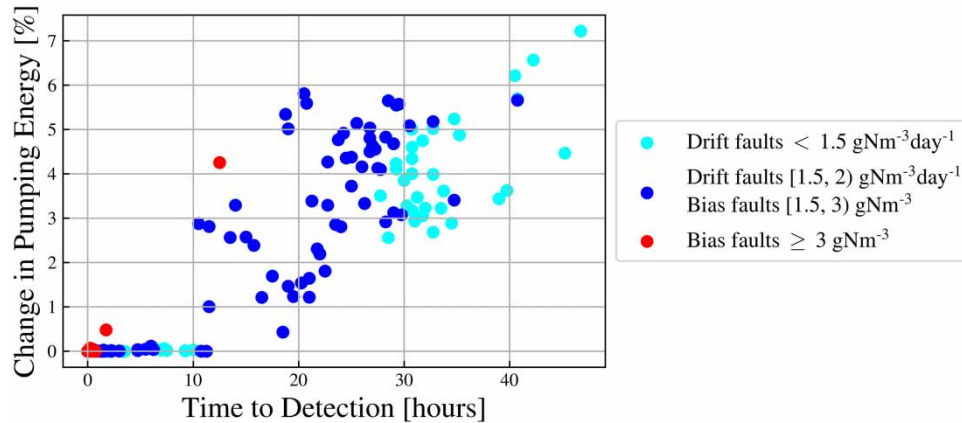


Figure 9 | The change in pumping energy plotted against the time taken to detect a fault.

sizes have the largest impact on the pumping energy requirements due to their extended presence in the process as a result of being more difficult to detect than larger faults. Another observation is that the first 10 h do not have a significant effect on the pumping energy; this suggests that accepting slower detection times in order to lower the number of false alarms is not expected to influence the process significantly. This is expected due to the relatively slow dynamics present in the WWTP.

CONCLUSIONS

It is clear that faults, even small faults, can have a significant effect on the performance and efficiency of the process, as seen in the up to 7% change in pumping energy. Rapid detection, and correction, of faults is crucial to optimising the operation of the process and ensuring that the quality limits placed on the effluent are not compromised. The three univariate control chart-based FD methods that were tested and compared are well established and have proved valuable in the field of FD applied to WWTPs. It was shown, however, that diligent calibration of the methods is vital due to the large effect that their respective parameters have on their performance. An approach for determining optimal parameters was developed and it was seen to be highly dependent on the criteria considered.

When the optimised chart methods were applied to online FD using the BSM1 the EWMA method outperformed the other methods for both bias and drift faults. The Shewhart method performed poorly in comparison to the CUSUM and EWMA methods for both fault types. When considering the bias faults, the CUSUM chart performed well, similarly to the EWMA chart. However, the CUSUM chart failed on the drift faults with high MDRs and TTDs. The EWMA chart was able to consistently detect more than 50% of the fault cases. The discrepancies between the results of the offline and online tests are explained by the active process control which was able to hide many of the drift sensor faults through the controller compensating for the apparent change in the controlled variable. It was shown that monitoring a manipulated variable can be a solution to identifying these faults.

The EWMA method should be further investigated, considering modifications that have been suggested to the base method, to further reduce the MDR and TTD in the case of drift faults. This method may be improved further through the use of filtering methods on the sensor signals, such as moving average windows, due to the apparent lack of urgency in detecting these faults shown in Figure 9.

ACKNOWLEDGEMENT

This study has been done within the international project Control4Reuse with partners from Sweden, France and Brazil. The project is part of the IC4WATER programme, in the frame of the collaborative international consortium of the 2017 call of the Water Challenges for a Changing World Joint Programme Initiative (Water JPI). The authors would like to thank Formas (Project No 2018-02213) for funding the Swedish part of this project, within the above mentioned initiative.

The authors also acknowledge and express gratitude to Dr Ulf Jeppsson and the other IWA Task Group members for making the BSM1 code available.

Dr Eva Nordlander is also thanked for her contributions in the form of supervision and valuable discussions related to the work.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing interests or personal relationships that could have influenced the work reported in this paper.

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

REFERENCES

- Baklouti, I., Mansouri, M., Nounou, H., Slima, M. B. & Hamida, A. B. 2017 Fault detection in a wastewater treatment plant. In: *Proc of the 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*. pp. 1–5. <https://doi.org/10.1109/ATSIP.2017.8075523>.
- Baklouti, I., Mansouri, M., Hamida, A. B., Nounou, H. & Nounou, M. 2018a [Monitoring of wastewater treatment plants using improved univariate statistical technique](#). *Process Safety and Environmental Protection* **116**, 287–300. <https://doi.org/10.1016/j.psep.2018.02.006>.
- Baklouti, I., Mansouri, M., Hamida, A. B., Nounou, H. & Nounou, M. 2018b [Max-double adaptive EWMA for fault detection of wastewater treatment plants](#). In: *Proc of the 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*. pp. 1–6. <https://doi.org/10.1109/AICCSA.2018.8612891>.
- Baklouti, I., Mansouri, M., Hamida, A. B., Nounou, H. N. & Nounou, M. 2021 [Enhanced operation of wastewater treatment plant using state estimation-based fault detection strategies](#). *International Journal of Control* **94** (2), 300–311. <https://doi.org/10.1080/00207179.2019.1590735>.
- Chiang, L. H., Russell, E. L. & Braatz, R. D. 2000 *Fault Detection and Diagnosis in Industrial Systems*. Springer Science & Business Media, London, UK.
- Corominas, L., Villez, K., Aguado, D., Rieger, L., Rosén, C. & Vanrolleghem, P. A. 2011 [Performance evaluation of fault detection methods for wastewater treatment processes](#). *Biotechnology and Bioengineering* **108** (2), 333–344. <https://doi.org/10.1002/bit.22953>.
- Germaey, K. V., Jeppsson, U., Vanrolleghem, P. A. & Copp, J. B. 2014 *Benchmarking of Control Strategies for Wastewater Treatment Plants*. IWA Publishing, London, UK. <https://doi.org/10.2166/9781780401171>.
- Isermann, R. 1997 [Supervision, fault-detection and fault-diagnosis methods – an introduction](#). *Control Engineering Practice* **5** (5), 639–652. [https://doi.org/10.1016/S0967-0661\(97\)00046-4](https://doi.org/10.1016/S0967-0661(97)00046-4).
- IWA Modelling & Integrated Assessment Specialist Group 2020 [Benchmarking – Modelling & Integrated Assessment](#). Available from: <http://iwa-mia.org/benchmarking/> (accessed 23 March 2020).
- Kazemi, P., Bengoa, C., Steyer, J.-P. & Giralt, J. 2021 [Data-driven techniques for fault detection in anaerobic digestion process](#). *Process Safety and Environmental Protection* **146**, 905–915. <https://doi.org/10.1016/j.psep.2020.12.016>.
- Marais, H. L., Nordlander, E., Thorin, E., Wallin, C., Dahlquist, E. & Odlare, M. 2020 [Outlining process monitoring and fault detection in a wastewater treatment and reuse system](#). In: *Proc of the 2020 European Control Conference (ECC)*, pp. 558–563.
- Montgomery, D. C. 2009 *Introduction to Statistical Quality Control*, 6th edn. John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Olsson, G. & Newell, B. 1999 *Wastewater Treatment Systems*. IWA Publishing, London, UK.
- Olsson, G., Nielsen, M., Yuan, Z., Lynggaard-Jensen, A. & Steyer, J.-P. 2005 *Instrumentation, Control and Automation in Wastewater Systems*. IWA Publishing, London, UK.
- Riss, G., Romano, M., Memon, F. A. & Kapelan, Z. 2021 [Detection of water quality failure events at treatment works using a hybrid two-stage method with CUSUM and random forest algorithms](#). *Water Supply* **21**, 3011–3026. <https://doi.org/10.2166/ws.2021.062>.
- Rosen, C., Röttorp, J. & Jeppsson, U. 2003 [Multivariate on-line monitoring: challenges and solutions for modern wastewater treatment operation](#). *Water Science and Technology* **47** (2), 171–179. <https://doi.org/10.2166/wst.2003.0113>.
- Sánchez-Fernández, A., Baldán, F. J., Sainz-Palmero, G. I., Benítez, J. M. & Fuente, M. J. 2018 [Fault detection based on time series modeling and multivariate statistical process control](#). *Chemometrics and Intelligent Laboratory Systems* **182**, 57–69. <https://doi.org/10.1016/j.chemolab.2018.08.003>.
- Schraa, O., Tole, B. & Copp, J. B. 2006 [Fault detection for control of wastewater treatment plants](#). *Water Science and Technology* **53** (4–5), 375–382. <https://doi.org/10.2166/wst.2006.143>.
- Thomann, M. 2008 [Quality evaluation methods for wastewater treatment plant data](#). *Water Science and Technology* **57** (10), 1601–1609. <https://doi.org/10.2166/wst.2008.151>.
- Thomann, M., Rieger, L., Frommhold, S., Siegrist, H. & Gujer, W. 2002 [An efficient monitoring concept with control charts for on-line sensors](#). *Water Science and Technology* **46** (4–5), 107–116. <https://doi.org/10.2166/wst.2002.0563>.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K. & Kavuri, S. N. 2003 [A review of process fault detection and diagnosis: part I: quantitative model-based methods](#). *Computers & Chemical Engineering* **27** (3), 293–311. [https://doi.org/10.1016/S0098-1354\(02\)00160-6](https://doi.org/10.1016/S0098-1354(02)00160-6).

First received 14 December 2021; accepted in revised form 18 January 2022. Available online 2 February 2022