ARTICLE

BIOTECHNOLOGY and BIOENGINEERING

Performance Evaluation of Fault Detection Methods for Wastewater Treatment Processes

Lluís Corominas,[1] Kris Villez,[1,2] Daniel Aguado,[3] Leiv Rieger,[1] Christian Rosén,[4] Peter A. Vanrolleghem[1]

[1]model*EAU*, Université Laval, Pavillon Pouliot, 1065 av. de la Médecine, Quebec, Que, Canada G1V 0A6; telephone: +1-418-656-2131 ext 8730; fax: +1-418-656-2928; e-mail: lcorominas@icra.cat

[2]Laboratory of Intelligent Process Systems (LIPS), Purdue University, West Lafayette, Indiana

[3]Instituto de Ingeniería del Agua y Medio Ambiente, Universidad Politécnica de Valencia, Valencia, Spain

[4]Veolia Water, Solutions & Technologies, Malmö, Sweden

Received 31 May 2010; revision received 6 September 2010; accepted 7 September 2010

Published online 29 September 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/bit.22953

ABSTRACT: Several methods to detect faults have been developed in various fields, mainly in chemical and process engineering. However, minimal practical guidelines exist for their selection and application. This work presents an index that allows for evaluating monitoring and diagnosis performance of fault detection methods, which takes into account several characteristics, such as false alarms, false acceptance, and undesirable switching from correct detection to non-detection during a fault event. The usefulness of the index to process engineering is demonstrated first by application to a simple example. Then, it is used to compare five univariate fault detection methods (Shewhart, EWMA, and residuals of EWMA) applied to the simulated results of the Benchmark Simulation Model No. 1 long-term (BSM1_LT). The BSM1_LT, provided by the IWA Task Group on Benchmarking of Control Strategies, is a simulation platform that allows for creating sensor and actuator faults and process disturbances in a wastewater treatment plant. The results from the method comparison using BSM1_LT show better performance to detect a sensor measurement shift for adaptive methods (residuals of EWMA) and when monitoring the actuator signals in a control loop (e.g., airflow). Overall, the proposed index is able to screen fault detection methods.

Biotechnol. Bioeng. 2011;108: 333–344.

KEYWORDS: activated sludge; data quality; mathematical modeling; monitoring; process control

Introduction

The use of on-line sensors in control and automation for optimized operation of wastewater treatment plants (WWTPs) is increasingly popular. As a result, large quantities of data are provided, which makes manual expert-based quality evaluation of these data impossible. Therefore, the development of methods that allow for automatic detection of faults (process monitoring) and identifying their root cause (fault diagnosis) are urgently needed. Such methods could be used to improve data quality and to account for the effects of faults in active controllers and overall plant performance (fault-tolerant control).

Fault detection is not a new topic in research. Today, a plethora of methods is available and widespread in different engineering areas (Venkatasubramanian et al., 2003a,b,c). In the wastewater treatment discipline process, history-based methods have been investigated most frequently. For instance, applied statistical process control (SPC) methods range from univariate methods like control charts (Schraa et al., 2006) to multivariate methods, for example, based on principal component analysis (PCA) (Lee and Vanrolleghem, 2003; Rosen and Lennox, 2001; Villez et al., 2008; Yoo et al., 2006). These methods work reasonably well with synthetic data sets but they have difficulties in handling real-life dynamics (e.g., seasonal changes) and non-linearities (e.g., temperature-dependent kinetics). For this reason, adaptive methods, in which some parameters can change over time, have been proposed to account for changing process behavior (Aguado and Rosen, 2008; Aguado et al., 2007; Lee et al., 2005; Lennox and Rosen, 2002). Expert knowledge-based methods have also been used for fault detection, for example, in Genovesi et al. (1999), fuzzy logic is used for monitoring anaerobic digestion.

Still, many other methods remain untested and only minimal practical guidelines exist for their selection and application. In order to objectively compare fault

Additional supporting information may be found in the online version of this article.
Correspondence to: Lluís Corominas
Contract grant sponsor: Canada Research Chair in Water Quality Modeling
Contract grant sponsor: NSERC Special Research Opportunities

detection methods, a simulation platform, the Benchmark Simulation Model No. 1 long-term (BSM1_LT), has been developed (Rosen et al., 2004) based on the work of the International Water Association (IWA) Task Group on Benchmarking of Control Strategies (BSM, 2009). Among other features, the BSM1_LT includes models that describe typical faults in WWTPs. One task that remains unaddressed within the BSM1_LT platform is the development of applicable criteria for objective comparison of fault detection algorithms.

Therefore, the goal of this article is to present a single index for evaluating fault detection performance that allows for screening different methods. The usefulness of this index to process engineering is illustrated first by application to a simple example and then it is applied to the BSM1_LT platform to evaluate the performance of different fault detection methods found in the literature. This example is made extra challenging for the detection methods and the performance index by introducing a gradual process change that is induced by a temperature profile over 1 year.

Materials and Methods

Fault Detection Evaluation Index

An objective index is proposed to evaluate the monitoring and diagnosis performance of different fault detection methods. This index assigns penalization points each time the fault detection method fails. The fault detection method evaluates at every sample whether a fault in a sensor or actuator is occurring or not. The output of the method is the estimated state of the sensor or actuator which is compared to the true state to determine if the output of the fault detection method is correct or not. When it is not correct, the fault detection method gets penalization points, so the worse the method performs the more points it gets. The index is designed in such a way that:

- False acceptance (FAC) is penalized by using a penalty function that is evaluated at each time instant. The later a fault event is detected the more penalization points are assigned to the fault detection method, so the proposed penalization approach is dynamic and considers the speed in fault detection.
- Intermittent detection of a fault within the duration of a fault event is extra penalized. This fact has been considered since undesirable switching from correct detection to non-detection can lead to a loss in trust on the fault detection system by the human operators.
- A constant penalization value is assigned to each false alarm (FAL).

The instantaneous penalties are summed over time to obtain the accumulated penalties. By taking into account that later detections are worse than earlier detections and by accounting for a loss in trust of human operators that could be caused by intermittent detections, the outcome of the evaluation index is a measure of reliability which ranges from 0% (not reliable) to 100% (reliable).

The details on how the fault detection index works are explained through Figure 1. This figure describes the evaluation of a hypothetical fault detection system to identify normal (state 1) and faulty (state 2) operation of a sensor along 1 day. As can be seen in this figure, the sensor fails from 0.5 to 0.8 days (true status—x—in Fig. 1a) and the hypothetical fault detection system provides an estimation of the sensor state (estimated status—$\hat{x}$—in Fig. 1a) that is not correct all the time: the fault detection system indicates a false alarm (FAL) between 0.2 and 0.3 days and a false acceptance (FAC) between 0.5 and 0.6 days. To evaluate the fault detection performance different steps are followed:

Definition of a timer (k): An artificial timer is initialized at the beginning of a fault event ($k(1) = 1$). The timer is switched On ($k(t) = k(t-1) + 1$) when a fault event starts (0.5 days in this example) and is switched Off ($k(t) = 1$) at the end of the event as can be seen in Figure 1b. As mentioned before, methods that switch back from a positive detection to non-detection (intermittent detection) are penalized extra (intermittent detection is illustrated in Fig. 2). To do so the time from the start of the fault, $k(t)$ is artificially increased with a positive constant k_{switch} whenever the alarm switches from correct detection to false acceptance (i.e., $k(t)$ becomes $k(t) + k_{\text{switch}}$).

Calculation of the penalty function (P(t)): Figure 1c shows the penalty function for false acceptance (P_{FAC}), which is calculated from Equation (1) using the artificial timer values from Figure 1b. The penalty function is exponentially increasing and reaches the maximal penalty level ($P_{\text{FAC,sat}}$) after a given time (τ_{FAC}) which reflects the urgency of penalizing for late detection (in this example $\tau_{\text{FAC}} = 45$ min). The penalty function for false alarms (P_{FAL}) is computed by using Equation (2) (Fig. 1d).

Calculation of penalization points (G): The penalties assigned to the evaluated fault detection system are obtained as follows. First, the difference between the true state ($x(t)$) and the estimated state ($\hat{x}(t)$) (Eq. 3, Fig. 1e) is computed, obtaining 0 if they coincide and 1 if they are different. Then, this sequence of 0 and 1s is multiplied with the penalty functions ($P_{\text{FAC}}(t)$ and $P_{\text{FAL}}(t)$). The area under the function shown in Figure 1f corresponding to the FAC (Eq. 4) and FAL (Eq. 5) situations are the accumulated penalties (G_{FAC} and G_{FAL}). Equation (6) gives the total penalization.

Calculation of maximum penalty: The total maximum penalty (G) and the maximum penalties for FAC and FAL ($G_{\text{FAC,max}}$ and $G_{\text{FAL,max}}$, respectively; see Fig. 1c and d) are obtained by setting $d(t)$ to 1 for all time instants (see Eqs. 7–9). The sum of P_{FAC} and P_{FAL} penalty functions represents the reference case that results from maximum penalization at all times.

Measure of reliability (J): J represents how reliable the system is to detect faults and it is determined by subtracting

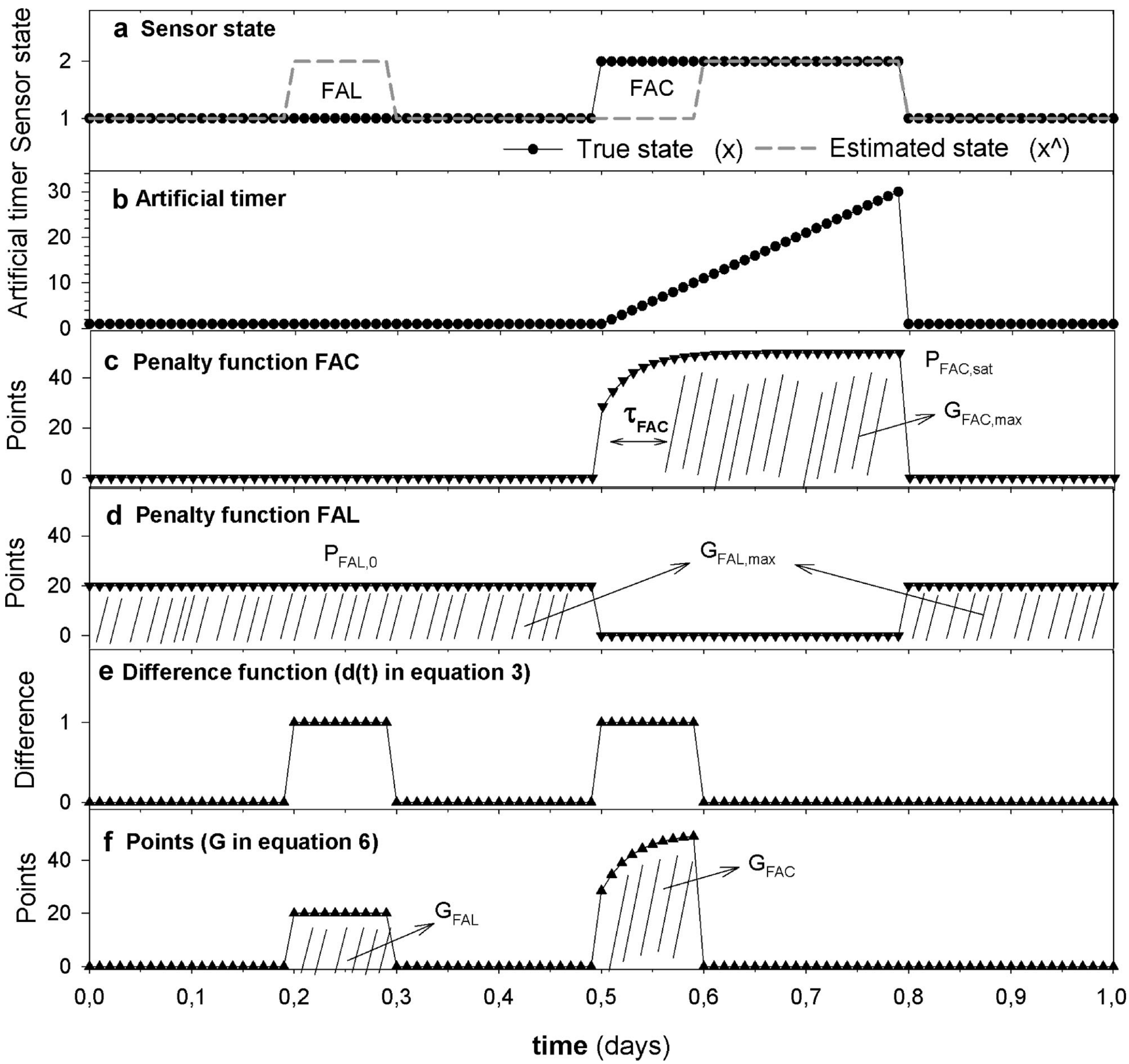

Figure 1. Representation of the fault detection index concept. (**a**) Sensor state, (**b**) artificial timer, (**c**) penalty function FAC, (**d**) penalty function FAL, (**d**) difference function ($d(t)$ in Eq. 3), and (**e**) points (G in Eq. 6).

the accumulated penalty divided by the maximal accumulated penalty from one and multiplying by 100 (Eq. 10). The closer to 100 the more reliable the fault detection system is judged to be. More information can be extracted from J_{FAC} and J_{FAL}, which indicate the reliability of the system against generation of false acceptance and false alarms, independently (Eqs. 11 and 12).

$$P_{FAC}(t) = P_{FAC,0} + (P_{FAC,sat} - P_{FAC,0})\left(1 - e^{(-k(t)/\tau_{FAC})}\right) \quad (1)$$

$$P_{FAL}(t) = P_{FAL,0} \quad (2)$$

$$d(t) = \max[0, \mathrm{abs}(x(t) - \hat{x}(t))] \quad (3)$$

$$G_{FAC} = \sum_{t_0}^{t_{end}} [P_{FAC}(t)\, d(t)] \quad (4)$$

$$G_{FAL} = \sum_{t_0}^{t_{end}} [P_{FAL}(t)\, d(t)] \quad (5)$$

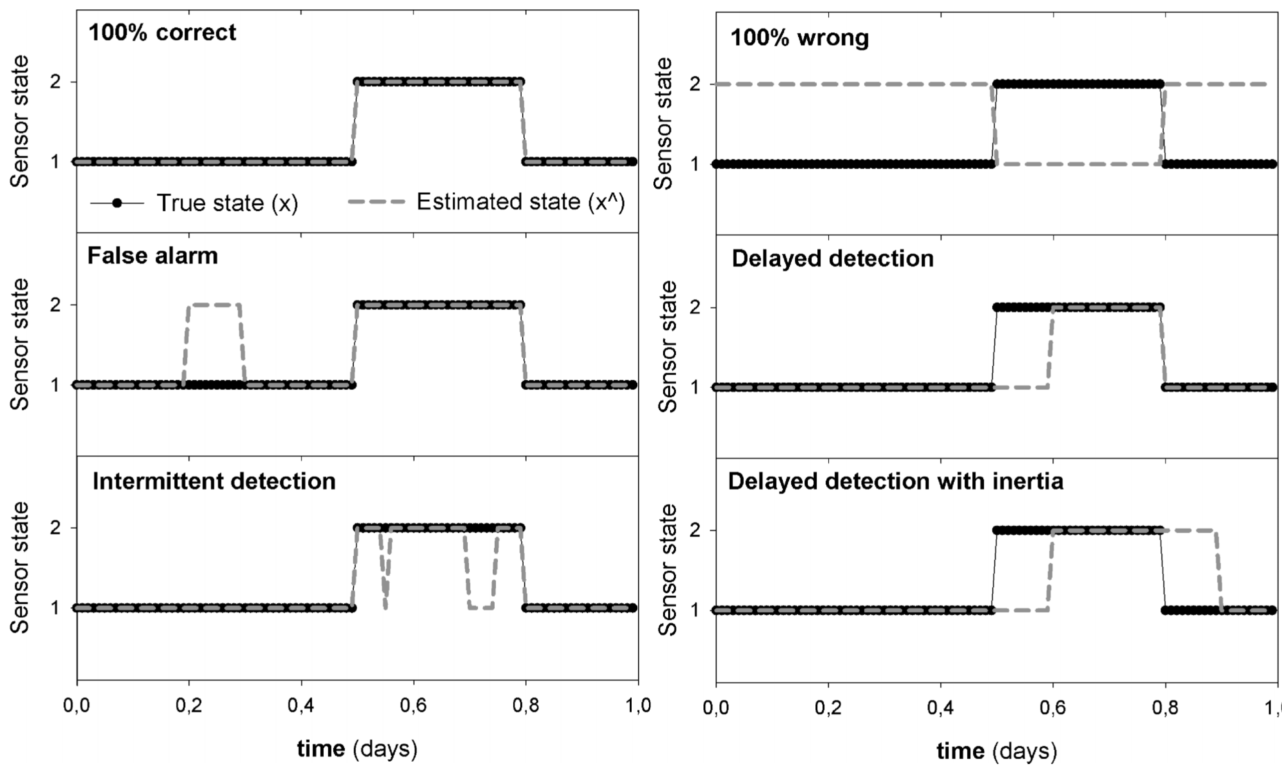

Figure 2. System and diagnosed sensor states for six artificial methods (1 is normal operation and 2 is faulty operation).

$$G = G_{FAC}(t) + F_{FAL}(t) \tag{6}$$

$$G_{FAC,max} = \sum_{t_0}^{t_{end}} [P_{FAC}(t)] \tag{7}$$

$$G_{FAL,max} = \sum_{t_0}^{t_{end}} [P_{FAL}(t)] \tag{8}$$

$$G_{max} = G_{FAC,max} + G_{FAL,max} \tag{9}$$

$$J = \left(1 - \frac{G_{FAC} + G_{FAL}}{G_{FAC,max} + G_{FAL,max}}\right) \times 100 \tag{10}$$

$$J_{FAC} = \left(1 - \frac{G_{FAC}}{G_{FAC,max}}\right) \times 100 \tag{11}$$

$$J_{FAL} = \left(1 - \frac{G_{FAL}}{G_{FAL,max}}\right) \times 100 \tag{12}$$

In summary, the inputs to the fault detection evaluation index are the true $x(t)$ and the estimated $\hat{x}(t)$ sensor/actuator states provided by the fault detection system that is being evaluated and the outputs are expressed as a measure of reliability of the fault detection system (J). Five parameters need to be specified for the index: $P_{FAC,0}$, $P_{FAC,sat}$, τ_{FAC}, $P_{FAL,0}$, and k_{switch}. The base level of penalization which is time independent is indicated by $P_{FAC,0}$ and $P_{FAL,0}$. $P_{FAC,sat}$ is used to indicate the maximal penalty level and τ_{FAC} reflects the urgency of penalizing for late detection (in the example of Fig. 1 $\tau_{FAC} = 45$ min). At this stage, it is suggested to use a trial and error approach to tune these performance evaluation parameters. However, the selection of the parameters should reflect the benefits and costs associated with the implementation of a fault detection system. Although in the given example only two types of sensor states are considered (normal and faulty), it is possible to extend this index to different types of faults by defining Equations (1) and (2) for each type of fault.

Illustrative Simple Example

The validity of the global index is checked against a simple illustrative example that covers six types of detection:

- Perfect detection (100% correct).
- Completely wrong detection (100% wrong).
- Equal to perfect detection but including a false alarm event (false alarm).
- Equal to perfect detection but with delayed detection of the fault (delayed detection).

- Equal to perfect detection but including a switch back detection (intermittent detection).
- Equal to delayed detection but including inertia (delayed detection with inertia).

Figure 2 presents the true (x) and the estimated ($\hat{x}$) sensor states for the six examples of detection.

BSM1_LT System

The simulation platform used is the BSM1_LT (Rosen et al., 2004). This platform includes model, process configuration (pre-denitrification plant with five activated sludge units in series, two anoxic—ASU1 and ASU2—and three aerobic—ASU3 to ASU5), control systems, benchmarking procedures, and evaluation criteria (for process and controller performance). It comprises a 1-year evaluation period with a dynamic influent and includes temperature-dependent kinetics. Given the focus on sensors and actuators, realistic models for both sensors and actuators are included (Rieger et al., 2003, 2006), as well as descriptions for sensor and actuator faults (Rosen et al., 2008a). Models for inhibition and toxicity (Rosen et al., 2008b) are also included in the platform. Thus, emulating often-encountered problems in real WWTP data, fault detection methods can be benchmarked within a realistic environment. The main motivation of using this platform is that these methods cannot be tested with real data because it is unfeasible to know the true value of a sensor over the long term.

Sensor Model Parameters

The NO_3-N measurement in the second anoxic compartment is of class B_0, which includes a response time of 10 min (see Rieger et al., 2003) with a measurement range of 0–20 g N m^{-3}. The measurement noise standard deviation is equal to 0.5 g N m^{-3}. For the dissolved oxygen sensors in the ASU4 and ASU5 aerated compartments, the probes are assumed to be of class A, which includes a response time of 1 min, with a measurement range of 0–10 g O_2 m^{-3} and a measurement noise standard deviation of 0.25 g O_2 m^{-3}.

Control System

The primary control objectives for the default control strategies in the BSM1_LT platform are to maintain (1) the NO_3-N concentration in the ASU2 (anoxic) at a predetermined set point value (1 g N m^{-3}) and (2) the dissolved oxygen concentration in the ASU5 (aerobic) at a predetermined set point value (2 g O_2 m^{-3}). The manipulated variables are the internal recycle flow rate from the last aerated compartment back to the first compartment and the oxygen transfer coefficient in the ASU5, K_La_ASU5 (which is the model equivalent to the airflow). An external carbon source (400,000 g COD m^{-3}) is added to the first reactor at a fixed rate Q_{carb} of 2.0 m^3 day^{-1}. In BSM1_LT, two different wastage flow rates are imposed depending on the time of the year in order to keep more biomass in the system during the winter period: 300 m^3 day^{-1} from October 30 to April 30 (days 0–182 of the dynamic simulation) and from October 29 to April 29 (days 364–546). For the rest of the simulation time, the wastage flow rate is 400 m^3 day^{-1}.

Faults Modeling

The illustrative example given here focuses on sensor faults and not on process faults (inhibition and toxicity kinetics of BSM1_LT are switched off in this study). The sensor models from Rieger et al. (2003) were extended to include seven sensor states, reflecting different sensor faults, as defined in Rosen et al. (2008a): (1) fully functional, (2) shift, (3) drift, (4) fixed value, (5) complete failure, (6) wrong gain, and (7) calibration. The occurrence of the faults is modeled by means of a Markov chain process (Rosen et al., 2008a). This model is run separately for each sensor and the outcomes are stored in files that are used as inputs for the simulations (Fig. 3 top presents the sequence of faults used for the DO sensors). For this study, the sensors used belong to the class "bad," which means that on average, a failure occurs every 2 weeks. In the case of the DO sensor it was fully functional for 74% of the total time (state 1), while for the remainder of time (26%) it suffered from shift problems (state 2). The other sensor faults were switched off in this study.

The implementation of the phenomenological fault descriptions was largely based on the approach described in Rosen et al. (2008a,b). Fault vectors with five elements (v1, v2, v3, v4, and v5) were created for the different types of faults (Table I). These elements sum or multiply to the true value in different places of the sensor model as shown in Figure 4 for the class A sensor of Rieger et al. (2003). Default parameters used in the BSM1_LT for the DO sensors are: $f_b = 2.0$, $f_g = 2.0$ (doubling of the slope of calibration curve), $f_r = 0.25\ (7\ \text{days})^{-1}$ (drift speed), and c_0 (the calibration point) = 2 mg L^{-1} for the DO sensor. The input and output from the DO sensor model (ASU4) are plotted in Figure 3 (bottom).

Simulation Protocol

The simulation protocol for BSM1_LT is as follows: First, the model is run to steady state for 200 days using a constant influent, without any faults. Afterwards, dynamic simulation is conducted using dynamic influent data (flows, concentrations, and temperature) for a period of 609 days at 15 min interval and with sensor and fault models active (the input file contains for each sensor the sequences of states, from 1 to 2, that were obtained from the Markov chains models). The first 245 days of dynamic simulation are used for training the fault detection methods and the following 364 days (245–609 days) are used to evaluate the performance of the methods. Simulations are conducted with WEST® (MOSTforWATER NV, Kortrijk, Belgium) and output data are stored every 15 min.

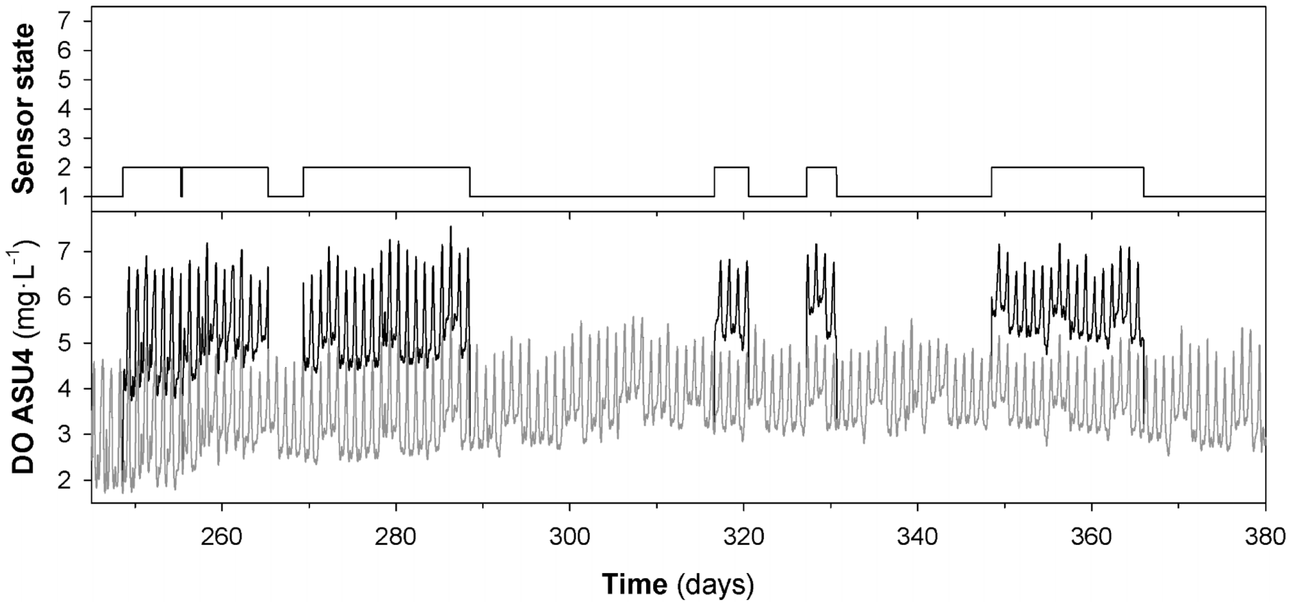

Figure 3. DO sensor states (**top**), true (gray) and measured (black) DO during the evaluation period (scenario 1) (**bottom**).

The outputs of the sensor models obtained during the dynamic simulation are fed to the fault detection methods implemented in Matlab® (Mathworks, Inc.) that generate the estimated sensor states ($\hat{x}$). The fault detection index is computed for each evaluated method by introducing the true states (x) imposed by the input file together with the estimated states ($\hat{x}$) obtained by the fault detection methods.

Simulated Fault Scenarios

The system is simulated for six scenarios with different combinations of faulty instruments and monitored variables (in cases 3 and 6 the fault is on the DO ASU5 sensor but the data used for monitoring is the K_La—i.e., airflow—instead of the direct DO measurement) and with different temperature profiles (see Table II). The fault sequences applied to the DO sensors in ASU4 and ASU5 are the same. The DO sensor in ASU5 is involved in a feed-back control loop. The fact that the DO sensor in ASU4 is not used for control allows for comparing performance of fault detection methods applied to controlled and non-controlled variables.

Table I. Vectors used to describe faults phenomenology (adapted from Rosen et al., 2008a).

	Fault vector				
Sensor status	[v1	v2	v3	v4	v5]
1. Fully functional	[1	0	0	1	0]
2. Shift	[1	f_b	0	1	0]
3. Excessive drift	[1	$-(t-t_0)\cdot f_r$	0	1	0]
4. Fixed value	[0	0	0	0	1]
5. Complete failure	[0	0	0	1	0]
6. Wrong gain	[f_g	0	$(1-f_g)\cdot c_0$	1	0]
7. Calibration	[0	0	0	0	1]

Fault Detection Methods

Several univariate control charts were tested and compared in this study. In order to ensure objective comparison the methods were calibrated and tuned with the same "fault-free" data set. All methods were implemented as defined in Montgomery (2009).

A control chart is a graphical display of a quality statistic that is computed from a univariate time series. A center line represents the average value (or no error for residual plots) and boundaries are set to the quality statistic (usually both an upper control limit—UCL and a lower control limit—LCL). Five methods were tested in this study:

- The Shewhart control chart.
- The exponentially weighted moving average (EWMA) control chart.
- Three variations on the Shewhart control chart on the residuals of EWMA (resEWMA):
  - resEWMA: Standard method as defined in Montgomery (2009).
  - resEWMA*: The EWMA filter is not updated if the quality statistic is out-of-control.
  - resEWMA*T: Same as resEWMA* but including temperature dependency for UCL and LCL ($UCL^T = f \cdot UCL$, f being an exponential function with $f = 1$ at 20°C and $f = 0.8$ at 10°C). The function was obtained using Hermite curve interpolation

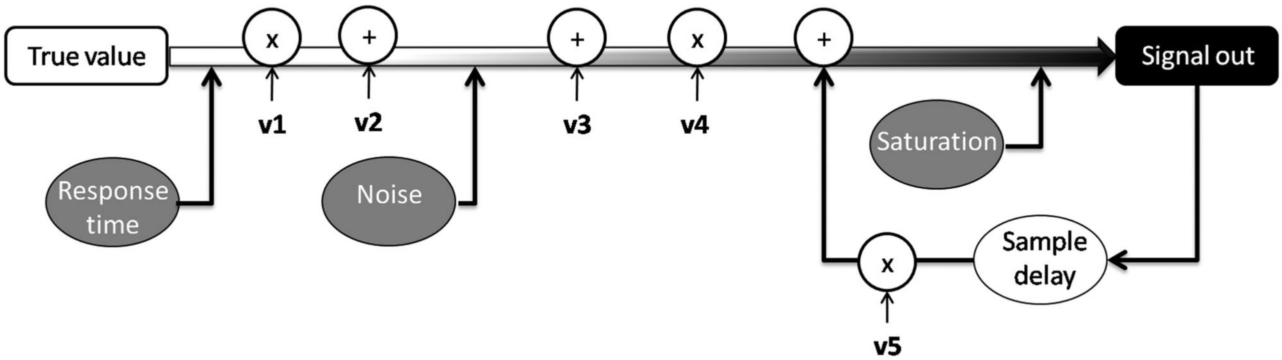

Figure 4. A sensor model including faults (class A).

Table II. Simulation scenarios.

Scenario	Faulty instruments	Monitored variable	Temperature	Manipulated variable
1 (DO_ASU4)	DO ASU4	DO ASU4	Dynamic	—
2 (DO_ASU5)	DO ASU5	DO ASU5	Dynamic	$K_{L}a$ ASU5
3 ($K_{L}a$_ASU5)	DO ASU5	$K_{L}a$ ASU5	Dynamic	$K_{L}a$ ASU5
4 (DO_ASU4_T)	DO ASU4	DO ASU4	Constant (15°C)	—
5 (DO_ASU5_T)	DO ASU5	DO ASU5	Constant (15°C)	$K_{L}a$ ASU5
6 ($K_{L}a$_ASU5_T)	DO ASU5	$K_{L}a$ ASU5	Constant (15°C)	$K_{L}a$ ASU5

(Spiegel and Liu, 1999) and the f values were obtained by trial and error.

Results

Illustrative Simple Example

Table III presents the results of the proposed indices for the six simulated cases depicted in Figure 2. $P_{FAC,0}$, $P_{FAC,sat}$, and τ_{FAC} values for false acceptance were 20, 50, and 3 samples (45 min), respectively, and k_{switch} was equal to 1. $P_{FAL,0}$ for false alarms was set to 20.

The evaluation system is set up that a *perfect detection* case receives 0 penalization points (G) and the *completely wrong* detection case receives the maximum penalty G_{max}. Accordingly, the *perfect detection* case gives 100% reliability (J) and *completely wrong* detection 0% reliability (J). If the decision of the fault detection system results in *false alarm*, *delayed detection*, *delayed with inertia*, or *intermittent detection*, there is an increase in the penalization points compared to the *perfect detection* case. This results in a reduction in reliability of 7%, 14%, 21%, and 11%, respectively. When only evaluating the reliability of the system against false acceptance (J_{FAC}), the *perfect detection* and the *false alarm* cases are 100% reliable because they are capable to identify the faulty event (from 0.5 to 0.8 days). The cases *completely wrong*, *delayed detection* (with and without inertia), and *intermittent detection* are 0%, 70%, and 79% reliable against false acceptance because the corresponding fault detection systems do not identify the faulty event properly. When focusing on false alarms, just the cases *completely wrong* and *false alarm* get penalization points (G_{FAL} of 1,380 and 200 points, respectively), obtaining a 0%

Table III. Index results for the artificial case study to verify the point awarding system.

	Unit	100% correct	100% wrong	False alarm	Delayed detection	Delayed detection inertia	Intermittent detection
G	Points	0	2,804	200	427	627	297
G_{max}	Points	2,804	2,804	2,804	2,804	2,804	2,804
G_{FAC}	Points	0	1,424	0	427	427	297
$G_{FAC,max}$	Points	1,424	1,424	1,424	1,424	1,424	1,424
G_{FAL}	Points	0	1,380	200	0	0	0
$G_{FAL,max}$	Points	1,380	1,380	1,380	1,380	1,380	1,380
J	%	100	0	93	85	78	89
J_{FAC}	%	100	0	100	70	70	79
J_{FAL}	%	100	0	86	100	100	100

and 86% reliability against false alarms (J_{FAL}). In the rest of the cases, no false alarms are generated and therefore these fault detection systems are 100% reliable against false alarms.

From the described results, it can be concluded that the developed penalty function evaluates several desirable aspects of a good fault detection system. Indeed, the penalty function is indicative of the detection performance in terms of false alarms, false acceptances, and intermittent detection. As indicated, this penalty function can be decomposed into penalties for false alarms and false acceptances.

Benchmark Case Study

The BSM1_LT is used to evaluate the performance of the univariate fault detection methods described above to detect shift faults in the DO probes in ASU4 and ASU5. The selection of the monitored variable (DO or K_La as a means for airflow) and the effect of temperature on the fault detection performance are evaluated.

Control Charts: Qualitative Evaluation

For all fault detection methods, the theoretical confidence interval corresponding to $\alpha = 0.0063\%$ (which is the very conservative 4-σ two-sided confidence interval for normally distributed data) was used. A value of 0.01 for λ (forgetting factor) is selected (high λ corresponds to short-term memory) for the EWMA-related methods.

An example of a control chart (resEWMA* for scenario 2, monitoring DO) is presented in Figure 5 left. It can be seen in Figure 5a that the DO signal oscillates around the DO

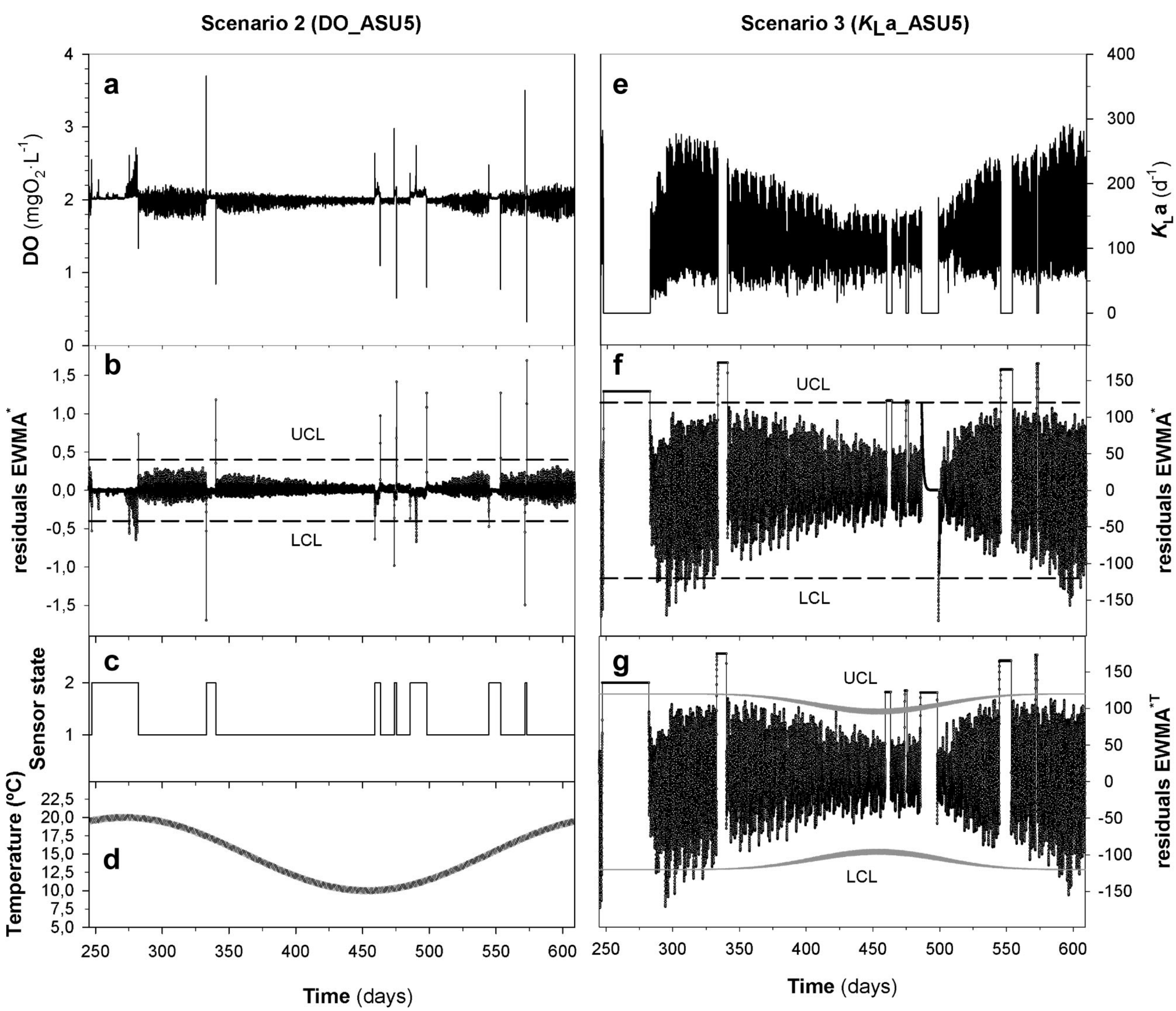

Figure 5. Control charts for scenario 2 (**left**) and scenario 3 (**right**). **a**: DO, (**b**) resEWMA*, (**c**) sensor states, (**d**) temperature, (**e**) K_La ASU5, (**f**) resEWMA*, and (**g**) resEWMA*T.

set-point of 2 mg L^{-1} with less variability during the winter period (see temperature profile in Fig. 5d) because the controller is well tuned for the winter period, with slower process dynamics. The shift faults provoke sudden changes in the signal at the beginning and at the end of the faulty events. In Figure 5b, the residuals of the EWMA are given with the upper and lower control limits (UCL and LCL, respectively), and in Figure 5c, the state of the sensor is presented (1 for functional and 2 for shift fault). This method detects the shift faults after a long delay (see in Fig. 5b that the residuals of EWMA are out of the UCL and LCL at the very end of the faulty events, Fig. 5c). The control charts for resEWMA* and resEWMA*T for scenario 3 (monitoring K_La) are shown in Figure 5 right. The imposed shifts are reflected with sharp decreases in the K_La signal, because the controller compensates for the shift (Fig. 5e). The K_La variations are bigger during the summer period. The resEWMA* and resEWMA*T methods are able to detect the shift faults because the residuals are outside the UCL and LCL for almost all fault events (Fig. 5f). Adjusting the UCL and LCL with temperature, improves the fault detection during the winter period (Fig. 5g). From the visual inspection of the control charts in Figure 5, it is clear that monitoring K_La gives better detection results than monitoring the DO signal.

Penalization System

The penalization system was used to evaluate the performance of the monitoring methods. The values for $P_{FAC,0}$, $P_{FAL,0}$, $P_{FAC,sat}$, and k_{switch} were the same ones used in the illustrative simple example. Here, the value for τ_{FAC} was 288 samples (3 days). The penalization system utilization is presented in Figure 6 (resEWMA* control chart for the scenario 3). At the top, the difference function is presented, which indicates the deviations between the true (x) and the estimated ($\hat{x}$) sensor states. In the middle, the penalty function (sum of Eqs. 1 and 2) describes the assignment of penalization points for false alarms and false acceptance. A more or less constant line at 20 points is observed that corresponds to the false alarm generation. Also, several sequences are observed for false acceptance in which the penalty smoothly varies from 20 to saturation (at 50 points). At the bottom, the actual points that the resEWMA* method obtained for Scenario 3 are presented. As seen in Figure 6 the method cannot detect the shift fault event from 490 to 500 days, and therefore, penalization points are received. The method can better detect the other faults and therefore few penalization points are accumulated. Two periods of false alarms can be identified, time from 280 to 340 days and from 580 to 610 days.

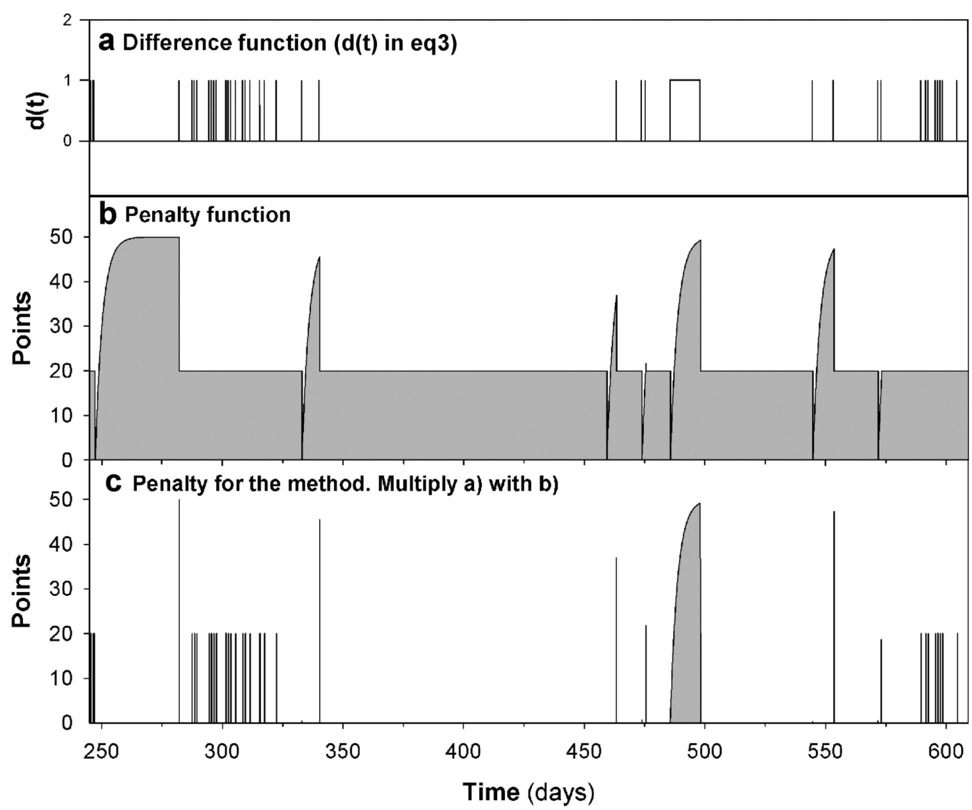

Figure 6. Penalization points for the resEWMA* method applied to scenario 3. **a**: Difference $(x-\hat{x})$ (Eq. 3). **b**: Penalty function. **c**: Penalty for the method. Multiply (a) with (b).

Comparison of Methods and Scenarios by Using the Fault Detection Index

Figure 7 presents the results of the proposed index for the evaluated methods applied to the defined scenarios. The results for the scenarios 1–3 (varying temperature) are shown on the left. The J, J_{FAL}, and J_{FAC} indicate the reliability of the fault detection system. In this case, J (Fig. 7, top left) shows that the methods are between 40% and 100% reliable to detect the imposed shift. The worst result (40% reliability) corresponds to the scenario 1 when monitoring DO that is not involved in a control loop and the 100% reliability is obtained for scenario 3 when monitoring K_La with DO involved in the control loop. Looking only at J_{FAL} (Fig. 7, middle left) all methods (except for EWMA) are between 90% and 100% reliable against false alarms in the different scenarios. Regarding false acceptance (J_{FAC}), only the methods EWMA, resEWMA*, and resEWMA^{*T} are more than 80% reliable. However, in the case of EWMA (scenarios 1 and 3) high reliability against false acceptance (J_{FAC} around 100%) is achieved at the expense of low reliability against false alarms (J_{FAL} around 20%, indicates that the system generates lots of false alarms). Overall, better performance is obtained when monitoring K_La (K_La_ASU5), than when monitoring the controlled DO (DO_ASU5) and finally when monitoring non-controlled DO (DO_ASU4).

Effect of Temperature Dynamics on the Monitoring Performance

The results for scenarios 4–6 (constant temperature at 15°C) are presented in Figure 7 right. It can be seen that the reliability of the methods both in terms of false acceptance (J_{FAC}) and false alarms (J_{FAL}) increases at constant temperature. Regarding the resEWMA^{*T} it is possible to obtain very similar performance for the K_La_ASU5 during dynamic temperature (Fig. 7, left) compared to the same case for constant temperature (Fig. 7, right), thus showing the benefits of including temperature to adapt the control limits (Fig. 5g).

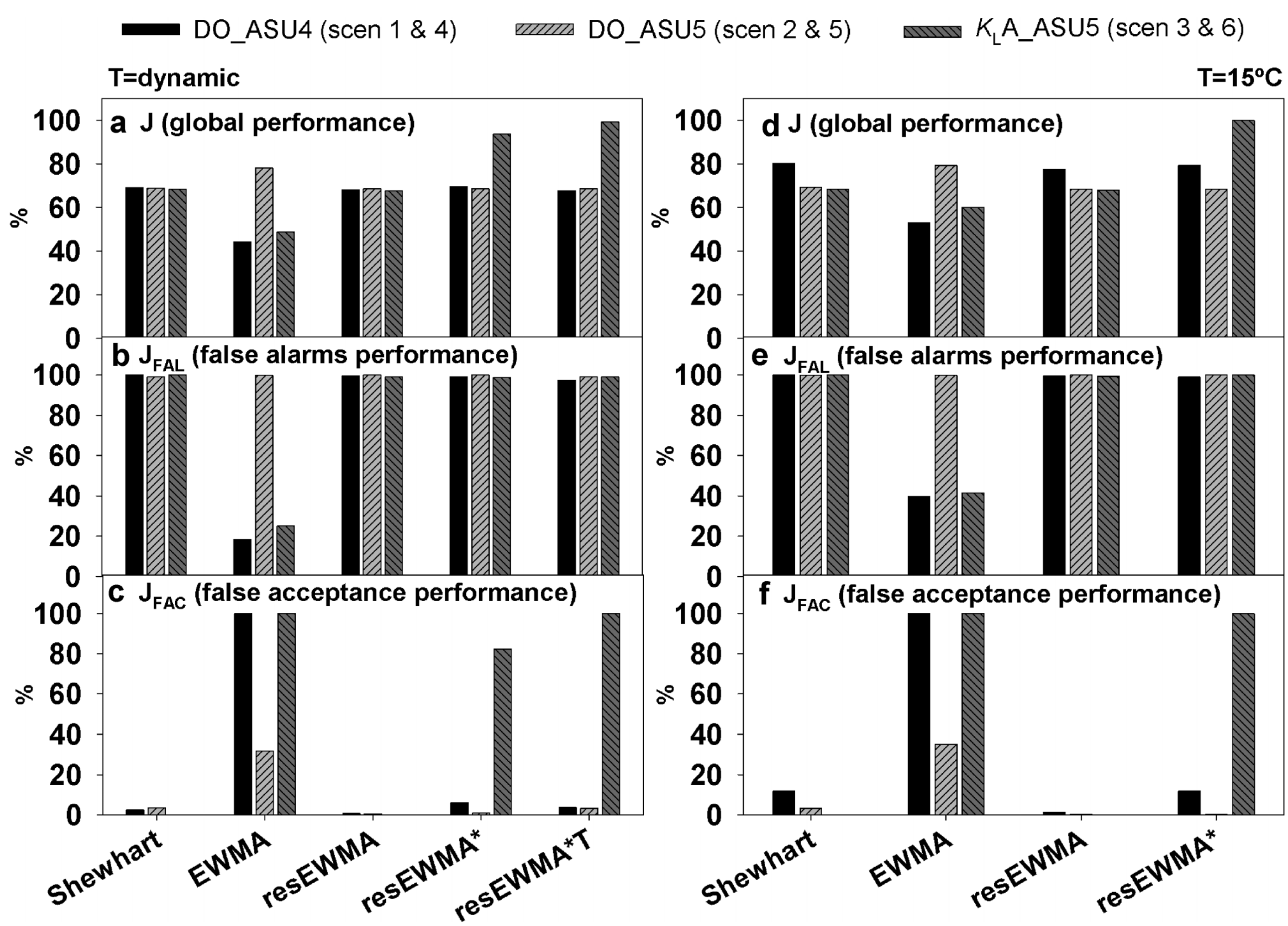

Figure 7. Evaluation of different monitoring methods (Shewhart, EWMA and three types of residuals on EWMA) applied to different scenarios (monitoring non-controlled DO, controlled DO and kLa at constant and dynamic temperature) using the proposed fault detection evaluation index (dynamic temperature on the left—a,b,c and constant temperature on the right—d,e,f). **a**: J (global performance). **b**: J_{FAL} (false alarms performance). **c**: J_{FAC} (false acceptance performance). **d**: J (global performance). **e**: J_{FAL} (false alarms performance). **f**: J_{FAC} (false acceptance performance).

Discussion

Fault Detection Methods

A good fault detection method should be accurate (correctness in classification) and fast in detecting the faults (speed). Moderate performance was observed for all methods tested in this study to detect shift faults, except for the resEWMA* method when monitoring K_La controlled by the faulty DO sensor. Better detection methods should be found that can cope with the real-life aspects of the BSM1_LT platform, such as time-varying and non-linear process behavior (e.g., effects of temperature). Future research should be aimed at the evaluation of methods that account for such characteristics and to evaluate other types of faults.

The poor performance of the resEWMA method has been tackled by means of two adjustments. First, in the resEWMA* method the EWMA filter is not updated if the quality statistic is out-of-control, which decreases the false acceptance rate of the method. Second, the resEWMA^{*T} increases the performance since more faults are detected when the limits are adjusted to temperature variation. It is worth noting that aeration-related variables (DO and K_La) present smaller variations in winter periods because of the lower dynamics at lower temperatures. Different control limits for winter and summer periods thus increase the overall performance of the method.

From the results obtained in this study, it is apparent that the choice of the monitored variable affects the performance of the fault detection methods. The action of the controller masks the effect of some faults (e.g., shift) in the sensor signal and therefore monitoring the actuator variable (K_La in this case) improves the fault detection performance (K_La_ASU5).

Further work will be conducted to evaluate other fault detection methods. For instance, multivariate methods (PCA methods) will be investigated. Moreover, already existing methods will be included in the analysis such as the method presented in Aguado and Rosen (2008), where adaptive and multivariate features are combined. Overall, an effort has to be made to develop methods for the wastewater treatment field that can account for daily, weekly, and seasonal patterns in the data and that are able to use redundant information.

Benchmark Platform

The BSM1_LT platform allows for testing the fault detection algorithms under standardized, realistic environmental conditions. The simulated measurements are sufficiently realistic to allow for adequate testing of monitoring methods before bringing them into practice. Although this study focused on sensor faults the BSM1_LT platform can also be used to model actuator and process faults (e.g., inhibition and toxicity detection) and therefore allows for testing of various fault detection methods for sensors, actuators, and process faults.

The proposed fault detection evaluation index is an important contribution to process engineering. However, further investigation on the parameter values of the index is needed ($P_{FAC,0}$, $P_{FAL,0}$, $P_{FAC,sat}$, τ_{FAC}, and k_{switch}). These parameters can highlight different aspects of fault detection performance, such as false acceptance, false alarms, and speed of detection. It is to be expected that the relative performance of monitoring methods is sensitive to the choice of the values of these parameters of the index. For practical relevance, the selection of parameters per type of fault should reflect the benefits and costs associated with implementing fault detection strategies.

As a last note, one may consider that in practice an alarm in an on-line application eventually leads to a control action aimed at compensating the problem, either by operator-based or automatic adjustment of plant operation (i.e., fault-tolerant control). In both cases, the monitoring system is expected to lead to more robust plant performance. Economic evaluation of the resulting closed loop plant performance may then be easier and more relevant to practice than the evaluation of the monitoring system on its own.

Conclusions

A first case study using BSM1_LT for evaluating monitoring performance in wastewater treatment systems (focusing on sensor faults) has been presented. A practical index for monitoring performance has been developed as a combined effort of the IWA Task Group on Benchmarking Control Strategies and experts on monitoring methods. The results obtained have proven that the proposed index is a valid tool to screen fault detection methods and to pinpoint their limitations.

The results from the method comparison show better performance to detect a sensor measurement shift for adaptive methods (residuals of EWMA) and when monitoring the actuator signals in the control loop (e.g., airflow). The applicability of the methods is limited if they do not account for changing process behavior. Therefore, more complex methods should be developed and tested that account for real-life process dynamics (e.g., temperature, inhibition, changing loads, etc.). In this study, it has been demonstrated that adapting the limits of the EWMA control charts according to temperature changes improved the performance of the methods. Further research will be conducted to test improved methods (multivariate and adaptive) and to assess costs and benefits of positive alarm detection, false alarms, and speed.

It is expected that this work encourages engineers and researchers to develop and test their own fault-detection methods using the BSM1_LT platform together with the proposed evaluation index.

This research is supported by the Canada Research Chair in Water Quality Modeling and a NSERC Special Research Opportunities grant as part of the Canadian contribution to the European Union 6th framework project NEPTUNE. Lluís Corominas benefits from the postdoctoral fellowship ''Beatriu de Pinós'' of the Government of Catalonia. The authors would like to thank Ulf Jeppsson for his contribution to the development of the BSM1_LT platform and the evaluation index.

Glossary

Estimated state	State of a sensor/actuator obtained from a fault detection system.
False acceptance	The fault detection system does not detect a fault when there is a fault occurring.
False alarm	The fault detection system detects a fault when there is no fault occurring.
Fault	An undesired deviation of at least one characteristic property or parameter of the system from the acceptable/usual/standard condition.
Fault detection	Determination of the faults present in a system and the time of detection.
Faulty operation	State of a sensor/actuator that indicates that a fault is affecting the sensor/actuator functioning.
Monitoring	The continuous real-time task of determining the conditions of a physical system, by recording information, recognizing and indicating anomalies in the behavior.
Normal operation	State of a sensor/actuator that indicates correct functioning.
Penalty	The disadvantage or painful consequences of a condition action or inaction.
True state	Known state of a sensor/actuator.

References

Aguado D, Rosen C. 2008. Multivariate statistical monitoring of continuous wastewater treatment plants. Eng Appl Artif Intell 21:1080–1091.

Aguado D, Ferrer A, Ferrer J, Seco A. 2007. Multivariate SPC of a sequencing batch reactor for wastewater treatment. Chemometr Intell Lab Syst 85:82–93.

BSM. 2009. International Water Association (IWA) Task Group on Benchmarking of Control Strategies for Wastewater Treatment Plants (BSM). http://www.benchmarkwwtp.org.

Genovesi A, Harmand J, Steyer JP. 1999. A fuzzy logic based diagnosis system for the on-line supervision of an anaerobic digester pilot-plant. Biochem Eng J 3(3):171–183.

Lee DS, Vanrolleghem PA. 2003. Monitoring of a sequencing batch reactor using adaptive multiblock principal component analysis. Biotechnol Bioeng 82:489–497.

Lee DS, Park JM, Vanrolleghem PA. 2005. Adaptive multiscale principal component analysis for on-line monitoring of a sequencing batch reactor. J Biotechnol 116:195–210.

Lennox J, Rosen C. 2002. Adaptive multiscale principal components analysis for online monitoring of wastewater treatment. Water Sci Technol 45(4–5):227–235.

Montgomery DC. 2009. Introduction to statistical quality control. 6th edn. New York: John Wiley & Sons.

Rieger L, Alex J, Winkler S, Boehler M, Thomann M, Siegrist H. 2003. Progress in sensor technology—Progress in process control? Part I: Sensor property investigation and classification. Water Sci Technol 47(2):103–112.

Rieger L, Alex J, Gujer W, Siegrist H. 2006. Modelling of aeration systems at wastewater treatment plants. Water Sci Technol 53(4–5):439–447.

Rosen C, Lennox JA. 2001. Multivariate and multiscale monitoring of wastewater treatment operation. Water Res 35:3402–3410.

Rosen C, Jeppsson U, Vanrolleghem PA. 2004. Towards a common benchmark for long-term process control and monitoring performance evaluation. Water Sci Technol 50(11):41–49.

Rosen C, Rieger L, Jeppsson U, Vanrolleghem P. 2008a. Adding realism to simulated sensors and actuators. Water Sci Technol 57(3):337–344.

Rosen C, Aguado D, Comas J, Alex J, Copp JB, Gernaey KV, Jeppsson U, Pons M-N, Steyer J-P, Vanrolleghem PA. 2008b. Inhibition and toxicity modelling with the long-term control benchmark model (BSM1_LT) framework. In: Proceedings IWA World Water Congress 2008. Vienna, Austria, September 7–12, 2008 (on CD-ROM).

Schraa O, Tole B, Copp JB. 2006. Fault detection for control of wastewater treatment plants. Water Sci Technol 53(4–5):375–382.

Spiegel MR, Liu J. 1999. Schaum's mathematical handbook of formulas and tables. 2nd edn. United States of America: McGraw-Hill.

Venkatasubramanian V, Rengaswamy R, Yin K, Kavuri SN. 2003a. A review of process fault detection and diagnosis. Part I: Quantitative model-based methods. Comput Chem Eng 27:293–311.

Venkatasubramanian V, Rengaswamy R, Kavuri SN. 2003b. A review of process fault detection and diagnosis. Part II: Qualitative models and search strategies. Comput Chem Eng 27:313–326.

Venkatasubramanian V, Rengaswamy R, Kavuri SN. 2003c. A review of process fault detection and diagnosis. Part III: Process history based methods. Comput Chem Eng 27:327–346.

Villez K, Ruiz M, Sin G, Colomer J, Rosen C, Vanrolleghem PA. 2008. Combining multiway principal component analysis and clustering for efficient data mining of historical data sets of SBR processes. Water Sci Technol 57(10):1659–1666.

Yoo CK, Villez K, Lee I-B, Rosen C, Vanrolleghem PA. 2006. Multi-model statistical process monitoring and diagnosis of a sequencing batch reactor. Biotechnol Bioeng 96:687–701.