Fault detection and isolation of sensors in aeration control systems

Bengt Carlsson, Jesús Zambrano

Division of System and Control, Department of Information Technology, Uppsala University, P.O. Box 337, 75105 Uppsala, Sweden.
(e-mail: bengt.calsson@it.uu.se; jesus.zambrano@it.uu.se).

Abstract: In this paper, we consider the problem of fault detection and isolation in the aeration system of an activated sludge process. The purpose is to detect and localize possible faults in dissolved oxygen and air flow sensors. The dissolved oxygen in each aerated zone is assumed to be controlled automatically. As the basis for a fault detection algorithm we use the ratio of air flow rates into different zones. The method is evaluated in two scenarios: using the Benchmark Simulation Model nº1 by Monte Carlo simulations, and using data from a wastewater treatment plant. The fault detection method shows good results for a correct and early fault detection and isolation.

Keywords: fault detection; isolation; aeration control; monitoring; incidence matrix.

Introduction

Fault detection (FD) and isolation is an active area of research due to the increasing complexity of industrial processes and growing demand for safety and reliability, the wastewater treatment plants (WWTPs) are not an exception. Besides the monitoring, the sensors are used for automatic control (eg. feedback and feedforward control) of plant performance. A necessary condition for a control system to work efficiently is that the sensor used in a control law is reliable. If this sensor gives a wrong value, too much resources (eg. energy for aeration) may be used or the treatment results may be poor (eg. high concentrations of ammonia in the effluent). The use of hardware redundancy, e.g. multiple sensors for the same variable, reduces the problem of a sensor fault, but is expensive and introduces complexity in the system.

Many different approaches have been suggested for FD applied to single or multiple variables in biological processes. Yoo et al. (2002) propose a modified PCA considering the importance of each transformed variable and not only the relative magnitude of the variance change. Baggiani and Marsili-Libelli (2009) show a dynamic PCA-based algorithm that can detect sensor failures in WWTPs. Corominas et al. (2011) present a comparison of different univariate FD methods (Shewhart, EWMA, and residuals EWMA) applied to the Benchmark Simulation Model No.1 long-term (BSM1_LT), where the sensor FD was studied in sensors under closed loop control. This problem poses special challenges, if a sensor signal is used in feedback control law, a fault in the sensor may not be visible from the sensor signal itself since the controller strive to keep the (possible faulty) sensor signal equal to the set point.

In this paper, the problem to detect and isolate sensor faults in the aeration system of an activated sludge process (ASP) is considered. In particular, faulty dissolved oxygen (DO) sensors under closed loop control. As the basis for the FD algorithm we use the ratio of air flow rates into different zones. The method is evaluated in two cases: using the Benchmark Simulation Model nº1 (BSM1) by Monte Carlo simulations, and using data from a WWTP.

The paper is organized as follows. First, two methods for detecting faults in DO sensors are outlined. Next, two case studies are described and the results from the methods are shown. Finally, discussions and conclusions are drawn.

Material and Methods

Consider an ASP with *N* aerated zones. The DO in each aerated zone is assumed to be controlled automatically. The airflow $\overline{q}_i(t)$ in every zone is known.

The Airflow Method (AM)

One method to detect faulty DO sensors in several aerated zones is by monitoring the airflow rate in every zone. In this method, a sensor fault in zone *i* is decided if:

$$\overline{q}_i(t) < a_i \quad or \quad \overline{q}_i(t) > b_i \; ; \quad \text{for } i = 1, 2, \dots, N \tag{1}$$

where a_i and b_i are the minimum and maximum bounds respectively, defined as:

$$a_i = \alpha_{min} \cdot \min\left\{\overline{q}_i(t)\Big|_{t\in A}\right\} \;\; ; \quad b_i = \alpha_{max} \cdot \max\left\{\overline{q}_i(t)\Big|_{t\in A}\right\} \tag{2}$$

where $\overline{q}_i(t)$ is a low-pass filtered value of the airflow rate into zone *i*. α_{min} and α_{max} are threshold factors used to define the lower and upper bounds, respectively. *A* is a set of data in non-faulty conditions.

The Airflow Ratio Method (ARM)

The ARM calculates bounds on airflow ratios during normal (non-faulty) conditions and uses these bounds to detect sensor faults. In this method, a fault is decided if:

$$f_{i,j}(t) = \frac{\overline{q}_i(t)}{\overline{q}_j(t)} > \gamma_{i,j} \; ; \; \text{for } i = 1,2 \dots, N; \quad j = 1,2, \dots, N \;\; (i \neq j) \tag{3}$$

where $\gamma_{i,j} = \alpha_{i,j} f_{i,j}^{\max}$ is a threshold calculated in non-faulty conditions. $\alpha_{i,j}$ is a threshold factor. $f_{i,j}^{\max} = \max\left\{\overline{q}_i(t)/\overline{q}_j(t)\Big|_{t\in A}\right\}$. *A* is a set of data in non-faulty conditions. Given *N* zones, there are *N(N-1)* airflow ratios $f_{i,j}(t)$. Therefore, a fault is decided if any of these ratios is above its threshold value. Details concerning AM and ARM can be found in Carlsson et al. (2013) and Carlsson and Zambrano (2013).

Fault isolation

Since AM is based on the signal monitoring in every zone, the isolation of the fault in zone *i* is given when a sensor fault in zone *i* is decided. For ARM, the methodology to decide the isolation of a fault is different. For example, consider the case of a positive bias in a DO sensor of zone *i*. The airflow rate in this zone will tend to decrease, then all the $f_{j,i}$ ratios are likely to be greater than the correspondent $\gamma_{i,j}$. By applying a positive or negative bias to another zone gives different response of the ratios.

A basic way to classify these responds is by binary evaluation of every airflow ratio against the thresholds. This structure gives an *observed fault signature* $\varphi_{i,j}$. Applications of the observed fault signature can be found in Fagarasan and Iliescu (2008), which is used as a tool for isolating the fault source. Making use of this structure we get:

$$\varphi_{i,j}(t) = \begin{cases} 1 & \text{if } f_{i,j}(t) > \gamma_{i,j} \\ 0 & \text{otherwise} \end{cases} \text{; for a negative bias in the sensor of zone } i.$$

$$\varphi_{j,i}(t) = \begin{cases} 1 & \text{if } f_{j,i}(t) > \gamma_{j,i} \\ 0 & \text{otherwise} \end{cases} \text{; for a positive bias in the sensor of zone } i.$$

These observed fault signatures allows isolation of the faulty DO sensor. The observed fault signature for every scenario can be ordered in an *incidence matrix*.

Application of incidence matrix in biological processes has been used for isolation of multiple actuator and sensor faults in a waste water treatment process (Fragkoulis et al., 2011). Table 1 shows the incidence matrix for the case of *N*=3 zones in series.

Table 1. Incidence matrix for 3 zones in series

	$\varphi_{1,2}$	$\varphi_{1,3}$	$\varphi_{2,1}$	$\varphi_{2,3}$	$\varphi_{3,1}$	$\varphi_{3,2}$
negative bias in DO1	1	1	0	0	0	0
negative bias in DO2	0	0	1	1	0	0
negative bias in DO3	0	0	0	0	1	1
positive bias in DO1	0	0	1	0	1	0
positive bias in DO2	1	0	0	0	0	1
positive bias in DO3	0	1	0	1	0	0

Evaluation of the fault detection methods

In order to evaluate the performance of the algorithms in terms of the *percentage* of fault detections, fault isolations, and false alarms, the following indexes are used:

$$FD = \left(\frac{1}{M}\sum_{k=1}^{M} n_{FD}^{[k]}\right) \cdot 100; \quad \text{where} \quad n_{FD}^{[k]} = \begin{cases} 1 & \text{if } t_{\text{fault}} < t_d < t_{\text{end}} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$FI = \left(\frac{1}{M}\sum_{k=1}^{M} n_{FI}^{[k]}\right) \cdot 100; \quad \text{where} \quad n_{FI}^{[k]} = \begin{cases} 1 & \text{if } t_d < t_i < t_{\text{end}} \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$FA = \left(\frac{1}{M}\sum_{k=1}^{M} n_{FA}^{[k]}\right) \cdot 100; \quad \text{where} \quad n_{FA}^{[k]} = \begin{cases} 1 & \text{if } t_d < t_{\text{fault}} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $\mathrm{t_{fault}}$, $\mathrm{t_d}$, $\mathrm{t_i}$, and $\mathrm{t_{end}}$ are the time of the true fault occurrence, the time of the fault detection, the time of the fault isolation, and the total evaluation time, respectively. FD is the percentage of fault detections; FI is the percentage of fault isolations and FA is the percentage of false alarms. *M* refers to the total number of simulation runs, and *k* refers to the k^{th} simulation.

The delay involved in the fault detection and isolation is taken into account. For AM the time of the fault detection and the time of the fault isolation are the same $(t_d = t_i)$. When ARM is used, the decision of fault detection and fault isolation are made separately. Therefore, fault detection index (I_{FD}) and fault isolation index (I_{FI}) have been defined as follows:

$$I_{FD} = \left(1 - \frac{1}{Mt_{\text{end}}}\sum_{k=1}^{M} t_{FD}^{[k]}\right) \cdot 100; \quad \text{where } t_{FD}^{[k]} = \begin{cases} \Delta t_d & \text{for correct FD} \\ t_{\text{end}} & \text{otherwise} \end{cases} \tag{7}$$

$$I_{FI} = \left(1 - \frac{1}{M}\sum_{k=1}^{M} \frac{t_{FI}^{[k]}}{(t_{\text{end}} - t_{\text{fault}})}\right) \cdot 100; \; \text{where } t_{FI}^{[k]} = \begin{cases} \Delta t_i & \text{for correct FI} \\ t_{\text{end}} - t_{\text{fault}} & \text{otherwise} \end{cases} \tag{8}$$

where $\Delta t_d = t_d - t_{\text{fault}}$ is the fault detection delay, and $\Delta t_i = t_i - t_{\text{fault}}$ is the fault isolation delay.

Case study 1: Synthetic data from BSM1

The BSM1 (Copp, 2002) was selected as the simulation platform. The BSM1 includes model, plant layout (pre-denitrification plant with five activated sludge zones in series, two anoxic and three aerobic), control systems and a benchmark procedure. The system was simulated using the MATLAB/Simulink® platform. Three different dynamic influent events were considered: dry, rain and storm. The simulation time was extended from 14 to 21 days in order have some more days for the FD evaluations. The DO feedback PI-control in zone 5 was also applied to zone 3 and 4. Every control loop has a set-point of 2 mgO_2/l. The nitrate feedback PI-controller in the second anoxic compartment is kept as by default. We use the notation DOj for the DO sensor in zone *j*.

The implementation of the faults was assessed based on the approach developed by Corominas et al. (2011) and previous work given by Rosen et al. (2008). The DO sensors belong to class A, with a response time of 1 min in non-faulty conditions, a measurement range of 0–10 mgO_2/l and a noise standard deviation of 0.25 mgO_2/l. Currently, no air flow model is defined in BSM1. For simplicity, K_La was selected as the monitored variable for the FD methods.

This case study was evaluated via MonteCarlo simulations, which includes the following steps:

Step 1: A set of dynamic simulations are executed for the three different influent events in non-faulty conditions. In this step, the thresholds are calculated.

Step 2: An influent event and a faulty DO sensor are selected.

Step 3: BSM1 is simulated for 150 days in order to reach steady state conditions.

Step 4: A dynamic simulation is performed during 21 days. The time of the fault event is generated via MonteCarlo method, in which the domain of the possible fault time is defined (in this case, the fault may occur between day 7 and 14). The fault time is generated randomly from a probability distribution over the domain.

Step 5: If the k^{th} run is lower than *M* (total number of simulations), go to Step 3. Otherwise, go to Step 2.

Case study 2: Real data from a WWTP

The FD methods were tested in a full-scale WWTP located in Stockholm, Sweden. In this case, the setup consisted on three aerated zones of one treatment line. The DO in each aerated zone was automatically controlled. The valve position was used as the monitored variable for the FD methods. The data was collected using a sampling time of 6 minutes. The experiment included the detection of bias and noise type faults. For bias fault, a value of -0.5mgO_2/l was added to the DO sensor value. Regarding the noise fault, it was generated by stopping the periodical cleaning of the sensor.

By default, the threshold values for AR (a_i ; b_i) and for ARM ($\gamma_{i,j}$) were calculated with $\alpha = 1.1$ (10% over the maximum values given in normal conditions). M = 20 total of simulation runs were executed for every influent event and faulty zone.

Results and discussions

Case study 1: BSM1

As an example, a positive bias of 1mgO_2/l was applied to the DO sensor in zone 5 (DO5). Figure 1 shows the airflow ratio profiles in non-faulty and faulty conditions.

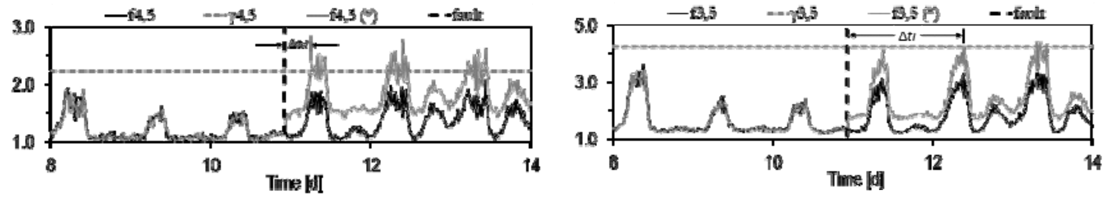

Figure 1. Airflow ratio profiles in normal conditions (fi,j in black), and when a bias of 1mgO_2/l is applied in DO5 (fi,j(*), in gray). Thresholds (dashed gray) and fault occurrence (dashed black).

Using the incidence matrix given in Table 1, the observed fault signature for this example is showed in Table 2.

Table 2. Observed fault signature for a positive bias in DO5

	$\varphi_{3,4}$	$\varphi_{3,5}$	$\varphi_{4,3}$	$\varphi_{4,5}$	$\varphi_{5,3}$	$\varphi_{5,4}$
positive bias in DO5	0	1	0	1	0	0

Note that $f_{4,5}$ is the first ratio above the threshold (giving $\varphi_{4,5} = 1$), therefore a sensor fault is decided. The second ratio above the threshold is $f_{3,5}$ (giving $\varphi_{3,5} = 1$). Hence, a fault in DO5 is decided. Furthermore, note the delay in the FD (Δt_d) and the delay in the fault isolation (Δt_i).

The same faulty condition was evaluated by MonteCarlo simulations, considering different locations of the faulty sensor (zone 3, 4 and 5), different influents (dry, rain and storm) and a fault occurrence generated randomly between day 7 and 14. Results are shown in Table 3. Note that AM and ARM give similar results in terms of the amount of fault detection, isolations and false alarms. However, taking into account the time needed for isolation, ARM shows better results. From this simulation study, ARM presents promising results for a correct and early fault detection and isolation.

Table 3. Summary of results of case study 1 for bias-type fault.

Influent	**Zone**	**Method**	**Δt_d [d]**	**Δt_i [d]**	**FD [%]**	**FI [%]**	**FA [%]**	**I_{FD} [%]**	**I_{FI} [%]**
Dry	3	AM	1.75 +/- 1.67		100	100	0	91.7	83.4
		ARM	0.03 +/- 0.03	0.18 +/- 0.17	100	100	0	99.9	98.3
	4	AM	0.54 +/- 0.52		100	100	0	97.5	96.9
		ARM	0.07 +/- 0.09	0.46 +/- 0.24	100	100	0	99.7	95.6
	5	AM	0.61 +/- 0.81		100	100	0	97.1	94.2
		ARM	0.21 +/- 0.21	0.32 +/- 0.27	100	100	0	99.0	97.0
Rain	3	AM	0.59 +/- 0.45		100	100	0	97.2	94.3
		ARM	0.03 +/- 0.05	0.07 +/- 0.1	100	100	0	99.9	99.3
	4	AM	0.82 +/- 0.81		90	100	10	93.3	85.4
		ARM	0.03 +/- 0.02	1.93 +/- 1.62	95	100	5	97.5	78.3
	5	AM	1.56 +/- 1.45		100	100	0	92.6	85.2
		ARM	0.62 /- 0.77	1.55 +/- 1.45	100	100	0	97.1	85.2
Storm	3	AM	0.58 +/- 0.58		100	75	0	97.3	64.3
		ARM	0.05 +/- 0.06	0.11 +/- 0.1	100	100	0	99.8	99.0
	4	AM	0.69 +/- 0.63		100	75	0	96.7	63.4
		ARM	0.07 +/- 0.09	0.78 +/- 0.60	100	90	0	99.7	93.4
	5	AM	1.47 +/- 1.44		100	100	0	93.0	86.0
		ARM	0.45 +/- 0.43	1.01 +/- 0.98	100	100	0	97.9	90.4

Note that for the case of 1 mgO_2/l of bias, similar results for AM and ARM in terms of the amount of fault detection *FD* and isolation *FI* are obtained. However, taking into account the instant of isolation, the index I_{FI} shows better results for ARM algorithm. This reflects that, in general, the time needed for fault isolation is lower in ARM than in the AM algorithm.

The effect of different levels of bias was also studied. For that, a range of bias from +/-0.25 to +/- 1 mgO_2/l was considered. The non-faulty condition (0 mgO_2/l) was included in the evaluation. Figure 2 shows the percentage of fault detections (FD) and false alarms (FA) for this evaluation.

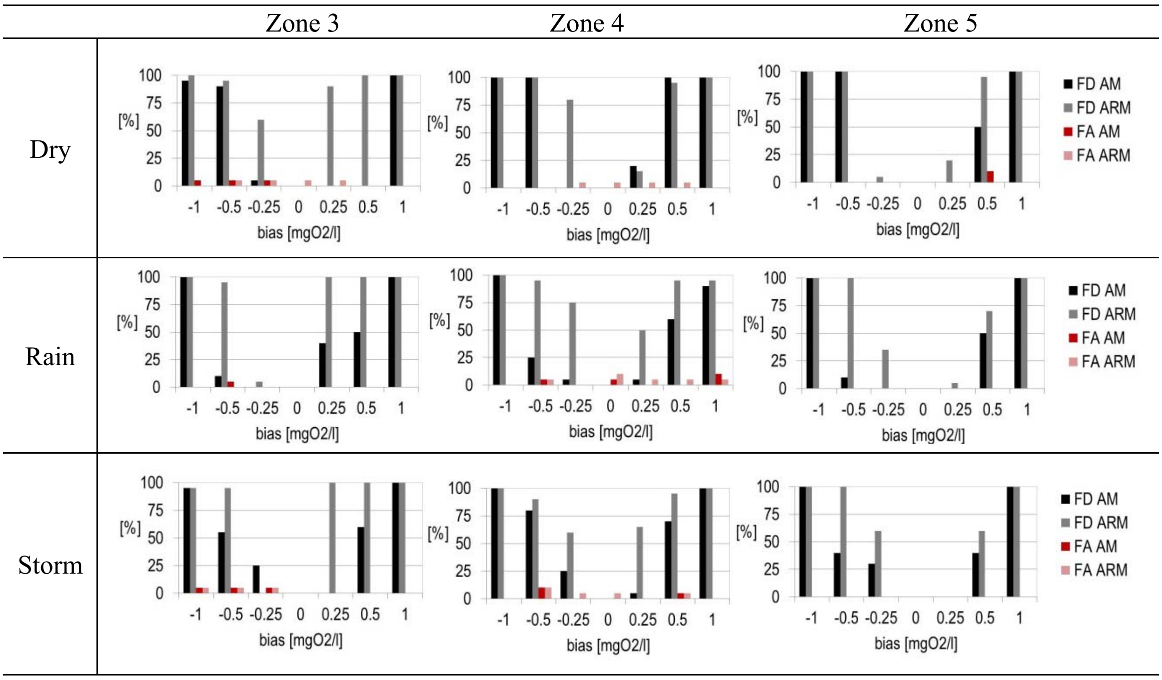

Figure 2. Percentage of fault detections (FD) and false alarms (FA) for AM and ARM methods for different bias faults.

The sensitivity analysis for the bias given in Figure 2 shows that the ARM has a higher rate of fault detections compared to the AM algorithm. As expected, both methods show similar performance for high level of bias (+/- 1). However, this performance decreases when the bias is in the range of - 0.25 to + 0.25, which is the range of the noise standard deviation defined for the sensor modeling in non-faulty conditions.

Case study 2: Full-scale WWTP

Figure 3 shows an example of the results obtained in the WWTP. In this case, a negative bias of 0.5mgO_2/l was applied to the DO sensor in the second zone at t_{fault}=44.43d, and the valve position is monitored before and after the fault occurrence. In the plots, the thresholds are shown with grey dashed lines, and the fault occurrence and fault detection are shown with blue and red dashed lines, respectively. V_i refers to the valve position in zone i.

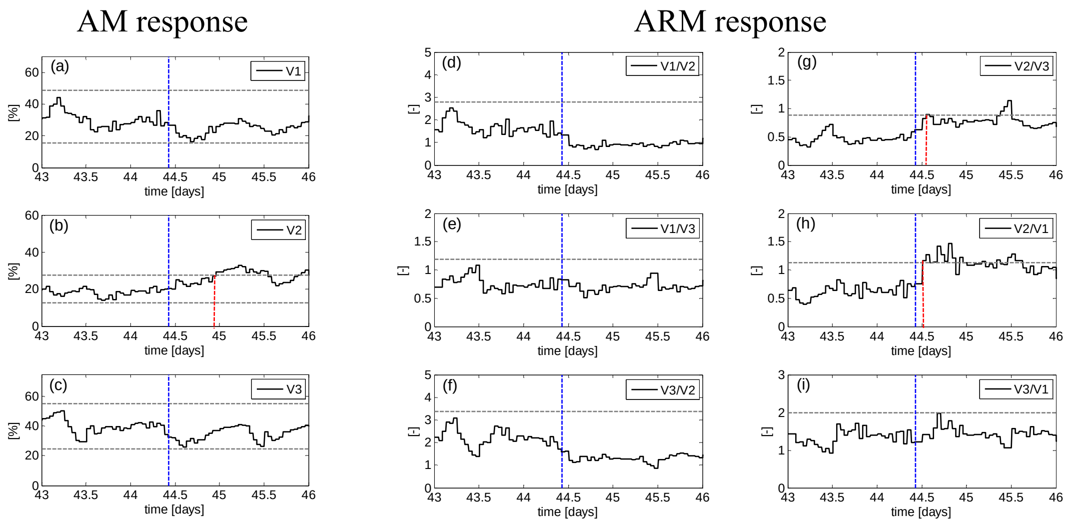

Figure 3. FD responses. (a), (b) and (c): AM response. (d) to (i): ARM response. Thresholds (dashed grey), true fault occurrence (dashed blue), fault detection (dashed red). V_i is valve position in zone i.

Observe that for AM, it is the valve position in the second aerated zone (V2) who crosses the threshold (giving t_d=44.94d). For ARM, it is first the ratio V2/V1 (giving t_d=44.52d) and then V2/V3 (giving t_i=44.56d) and who cross the thresholds. Hence, as indicated in Table 1, by using these two ratios fault isolation in the second zone can be decided. Note that ARM gives fault isolation earlier than AM.

Since AM and ARM are FD methods based on the monitoring of the manipulated variable, this indirect way of monitoring makes the delay in the fault detection depend on the moment at which the fault signal occurs. For example, in the case of ARM, if there is a fault in DO*i* sensor, the delay in the fault detection will be shorter if the correspondent $f_{i,j}$ ratio is close to its maximum values. Similarly, the delay will be longer if the correspondent $f_{i,j}$ ratio is close to its minimum values in the fault moment. The same analysis applies for the airflow monitoring in the AM algorithm.

Conclusions

This study shows the performance of a simple fault detection algorithm, ARM, developed in order to detect and isolate faults in the dissolved oxygen sensor during closed loop control. The new approach is compared with AM, which monitor the aeration in every zone.

The method assumes that the DO sensors are in closed loop control. The method can be used for an arbitrary number of zones in series. Three aerated zones are used as case study, applying the method to the BSM1. Naturally, AM and ARM can be used for fault detection and isolation, although simulation and experiments showed that ARM gives better performance. In this respect, the definition of indexes for fault detection and fault isolation allowed quantifying and comparing the methods, taking into account not only the number of detection but also the time delay involved.

The method can be extended in several ways including:

- In many plants there are a number of parallel lines, each with a number of aerated zones. A natural extension of ARM for this case is to also compute the ratios between zones in different lines.
- If there is a significant time delay in the air flows into different zones, this delay may be adjusted by calculating the ratio in Equation (3) as:

$$f_{i,j}(t) > \frac{\bar{q}_i(t)}{\bar{q}_j(t-\tau)} \quad (9)$$

where τ is an estimate of the time delay between zone i and j. This may improve the performance of the detection. Note, however, that in order to calculate Equation (9) a time delay is unavoidable. An interesting topic for further research is to compare Equation (3) with (9) for systems with significant hydraulic delays.

Acknowledgments

The authors acknowledge funding support under the European Union's Seventh Framework Programme managed by the Research Executive Agency (REA) http://ec.europa.eu/research/rea (FP7/2007_2013), Grant Agreement N.315145 (Diamond). The MATLAB implementation of the BSM1 model has been provided by Lund University, courtesy of Dr. Ulf Jeppsson.

References

Baggiani, F. and Marsili-Libelli, S. (2009). Real-time fault detection and isolation in biological wastewater treatment plants. Water Science and Technology, **60**(11), 2949–2961.

Carlsson, B. and Zambrano, J. (2013). Fault detection of sensors in aeration control systems – the airflow ratio method. Technical report available on www.it.uu.se/research/publications.

Carlsson, B., Zambrano, J., Dahllöf, K., Isaksson, E. and Josefsson, G. (2013). Fault detection of sensors in aeration control systems – the airflow ratio method. In Proceedings of the 11th IWA conference on instrumentation control and automation, Narbonne, France, September 18-20.

Copp, J. B. (2002). The COST Simulation Benchmark-Description and Simulator Manual. Office for Official Publications of the European Communities, Luxembourg.

Corominas, L., Villez, K., Aguado, D., Rieger, L., Rosén, C. and Vanrolleghem, P. (2011). Performance evaluation of fault detection methods for wastewater treatment processes. Biotechnology and Bioengineering **108**(2), 333–344.

Fagarasan, I. and Iliescu, S. St. (2008). Parity Equations for Fault Detection and Isolation. In Proceedings of the IEEE International Conference on Automation, Quality and Testing, Robotics, **1**, 99–103.

Fragkoulis, D., Roux, G. and Dahhou B. (2011). Detection, isolation and identification of multiple actuator and sensor faults in nonlinear dynamic systems: Application to a waste water treatment process. Applied Mathematical Modelling, **35**, 522–543.

Rosen, C., Rieger, L., Jeppson, U. and Vanrolleghem, P. (2008). Adding realism to simulated sensors and actuators. Water Science and Technology, **57**(3), 337–344.

Yoo, C.K., Choi, S.W. and Lee, I. (2002). Disturbance detection and isolation in the activated sludge process. Water Science and Technology, **45**(4–5), 217–226.