Monitoring fouling on dissolved oxygen sensors in WRRFs with active fault detection

Oscar Samuelsson[1,2*], Anders Björk[1], Jesús Zambrano[3] and Bengt Carlsson[2]

[1]IVL Swedish Environmental Research Institute, Process modelling & IT, Stockholm, Sweden

[2]Division of System and Control, Department of Information Technology, Uppsala University, Sweden

[3]School of Business, Society and Engineering, Mälardalen University, Västerås, Sweden

*Corresponding author: *oscar.samuelsson@ivl.se*

Abstract: In this study, the automatic air cleaning system for dissolved oxygen (DO) sensors was used to monitor the sensor's status. The response from an air cleaning impulse changed shape when artificial fouling was applied to the sensor. In addition, the shape of the IR changed when other relevant sensor faults were introduced. Two fault detection methods were used to interpret the change in impulse response, Time constant estimation and Gaussian process regression. Both methods were evaluated for their receiver operation characteristics and had high detection rate at a low false alarm rate in the conducted experiments. Both methods were sensitive enough to detect a fouling that resulted in a bias smaller than -0.2 mg/L. The results indicate that the suggested approach is promising and can increase the reliability of DO-measurements at WRRFs.

Keywords: active fault detection, monitoring, dissolved oxygen sensor, machine learning, fouling

INTRODUCTION

Accurate on-line dissolved oxygen (DO) measurements are central to control aeration in water resource recovery facilities (WRRFs). Despite this, more research has been focused on DO-control strategies, see e.g. (Amand *et al.* 2013), than on methods to detect inaccurate DO-measurements. One reason for the lack of fault detection (FD) methods might be that the common membrane electrochemical DO-sensor (MEC) is considered to be a robust and well-established measurement technology, introduced more than fifty years ago (Clark 1959). On the other hand, an alternative to MEC-sensors have been introduced that use an optical measurement technology (OPTs) (Demas *et al.* 1999), indicating that the strive for accurate DO-measurements is not completed.

Regardless of measurement technology, the sensor has to be clean to provide accurate measurements. This is a challenge within the filthy wastewater liquid that make sensors susceptible to biofilm fouling and blocking by fibrous waste. Thus, visual inspection of the sensor and cleaning are required on regular basis to guarantee accurate readings. A complement to the routine maintenance is to use FD-methods, which can to detect sensor disturbances ahead of a scheduled inspection.

So far, most FD-methods for monitoring DO-sensors have been based on existing on-line measurements that are correlated to the DO-measurement. Yoo et. al studied how principal component analysis could be used to detect precision degradation and complete sensor failure due to sludge clogging in a simulated SHARON process model (Yoo *et al.* 2008). Carlsson and Zambrano compared the ratio of airflows in consecutive aerated zones, to detect bias in DO-sensors

(Carlsson & Zambrano 2016). The latter method was effective to detect sudden changes in airflows and also slow drift in air flow ratios.

Unfortunately, we recognize three problems with the mentioned FD-methods. Firstly and most important, it is hard to use the existing FD-methods to distinguish sensor faults from process disturbances. Secondly, a fundamental problem with FD-methods based on correlated variables is the ambiguity in which variable that is faulty. A detected fault could equally well be a result of a fault in the correlated variable, rather than in the monitored one. Finally, a detected change in a correlated variable indicates that the process has already been disturbed before the faulty sensor could be corrected. Thus, a faulty sensor has to disturb the process, in order for the FD-method to be applicable.

One approach that could potentially relieve the three mentioned problems is active fault detection, as against traditional or passive fault detection. In active fault detection, an auxiliary signal is designed exclusively for fault detection and injected into the system (Esna Ashari *et al.* 2012). In this study, we used the impulse from an automatic air cleaning system of the DO-sensor as design signal, and monitored the impulse response (IR). A similar approach was suggested already in 1992 (Spanjers & Olsson 1992), where a changed time constant of the DO-sensor was shown to be a good indication of an artificially fouled DO-sensor. More recently, Andersson and Hallgren showed that the IR from an air-cleaning procedure changed, when the sensor was fouled by organic biofilm (Andersson & Hallgren 2015).

In this study we wanted to study the questions:

1) How is the IR changed due to fouling?
2) Can the IR change be characterized with enough precision so that a fouled sensor can be detected before a bias is introduced in the DO-measurement?

We made detailed repeated experiments with artificial fouling at full-scale conditions for two sensor types, MEC-sensors and OPTs. We further studied the possibility to detect changes in IRs at small bias (<0.2 mg/L). We used receiver operating characteristics (ROC), see e.g. (Fawcett 2006), to visualize the performance of Time constant estimation as in (Andersson & Hallgren 2015) and Gaussian process regression (Rasmussen & Williams 2005).

MATERIALS AND METHODS

Artificial faults experiments

All artificial fouling experiments were conducted at full-scale conditions in Henriksdals WRRF during summer conditions with a wastewater temperature between 18 and 19 °C. Experiments were conducted at four positions in the activated sludge process. The main purpose was to obtain measurements of IRs at different DO-levels (0 – 4.5 mg/L). Therefore, both aerated and unaerated zones were considered (zones 4 – 6). A few measurements were also made in the return sludge channel to study whether a change in suspended solids concentration would have an impact on the IR. The suspended solids concentration was between 7500 and 9500 mg/L in the return sludge channel, in contrast to 2500 – 3000 mg/L in zones 4 – 6.

Artificial fouling procedure

In this study, to goal was to apply a fouling substance which was:

1) simple to repeatedly apply and remove without damaging the sensor
2) similar to organic biofilm, or at least exhibiting similar effect with a small bias
3) sustainable towards an impulse from the automatic air cleaning procedure

A wide variety of fouling agents were considered and a mixture of ball-bearing grease and floating grease from the pre-sedimentation fulfilled the above criteria. The second criteria was hard to verify, but both MEC-sensors and OPTs got a negative bias when fouled with the grease mixture, which was the same effect from organic fouling noticed in full-scale by (Andersson & Hallgren 2015).

One fouling procedure consisted of the following:

- Calibration/verification of accurate measurements of the test sensors by comparison with the reference sensors
- Measure repeated IRs from clean test sensors
- Manual fouling of the test sensors with grease mixture and repeated impulse measurements at fouled conditions
- Manual cleaning of test sensor and verifying its agreement with reference sensors

Sensor set-up and data collection

Five DO-sensors (Cerlic O2X DUO) were connected to a data acquisition system with hardware and software (LabVIEW) from National Instruments) and data were stored in a PostgreSQL database in the same laptop computer. Each DO-sensor could be switched between MEC and OPT by changing the top of the sensor. Two of the DO-sensors, one OPT and one MEC, were used to study IRs (test sensors), two as references (MEC-sensors), and the last one was used as back-up sensor in case of a failing test or reference sensor (MEC). The reference sensors were joined together about 1.5 m away from the test sensors to avoid interference from the air impulses. All sensors were mounted at rods according to the manufacturer's instruction at a slight angle (5-30°) and at 0.5 m depth. All membranes were replaced with new ones and calibrated in the beginning of the experiments. The length of an air cleaning impulse was set to 15s at 2 bar, which was enough to obtain a clear IR, even for high DO levels (4 mg/L).

Data pre-processing

Data were sampled with 8 Hz but down-sampled during the data pre-processing to 1 Hz to resemble full-scale conditions. First, the data were low-pass filtered (anti-alias filter) and afterwards down-sampled to 1 Hz. One reference DO-concentration was estimated from the two reference sensors using their weighted least squares (WLS) estimate. The bias for both test sensors was calculated as the difference between the test sensor and WLS-estimate for the time between IRs.

Fault detection methods

Two FD-methods were used to interpret the IRs, Time constant estimation (TCE) and Gaussian process regression with sequential Monte Carlo (GPR), described in (Svensson. *et al.* 2015).

The time constant of an IR was defined as the time it takes to reach 63 percent of the peak amplitude. The initial DO-level was estimated as the mean DO-level five seconds before an impulse. Ten normal (non-faulty) profiles were randomly selected to train both methods. The standard deviation for the time constant was calculated on the training data and used as setting for alarm threshold. Thus using the threshold 3 stds would result in a more restrictive detection with fewer alarms, compared to e.g. 1 stds. We used number of standard deviations instead of e.g. a fixed time constant to incorporate the magnitude of the normal variation for a non-faulty sensor in the FD-method.

The GPR applied to WRRFs has been described in (Samuelsson *et al.* In press) and can in short be described as a non-linear regression method that performs a smooth interpolation of training data. The GPR also produces a data dependent uncertainty interval (prediction interval) were both the predictions and the prediction interval are based on the assumption that data can be described as a Gaussian process. The prediction interval was used as threshold to raise an alarm if a test IR was outside the expected prediction interval. The number of standard deviations in the prediction interval was used as threshold.

An important part of the GPR is to select a proper kernel (covariance function). Here we used the standard squared exponential kernel since we expected the IR to be smooth. The initial DO-level before the IR was included as a second variable since we expected that it would influence the IRs. The GPR was included in this study since a data driven flexible regression method should be a good method to recognize differences in the IR's pattern.

Performance evaluation

In this study we used Receiver operating characteristics (ROC) to visualize the performance of the two FD-methods to distinguish fouled IRs from normal ones. The ROC is a graph of the false alarm rate (x-axis) versus the detection rate (y-axis), see e.g. (Fawcett 2006). For FD-methods that require a parameter to be set, the ROC enables an objective comparison of the FD-method´s performance, regardless of parameter settings.

A straight line is commonly included in a ROC graph, which indicates the same probability to detect a fault, as to make a false alarm detection. Here we denote this as the coin toss detector (CTD), which indicates the lower performance bound that any applied FD-method should outcompete. In opposite to the CTD, the optimal detector should have zero false alarm probability and detection probability, much like a step in the ROC-graph.

RESULTS AND DISCUSSION

The experiments resulted in 13 fouling procedures of varying length (from 1 hour to over one week's measurements). All data were pre-processed and the IRs (IRs) were extracted from the raw data and grouped according to their status (normal or fouled), where each IR was 180s long. This resulted in 515 normal IRs and 2628 IRs for fouled conditions for the membrane electrochemical (MEC)-sensors. Examples of typical IRs can be seen in Figure 1.

For the optical sensor (OPT), the sensor was damaged during the experimental start and was replaced in the end of the experimental period. This resulted in a reduced group of data for the OPT-sensors (21 normal and 2182 faulty IRs), therefore the main results presented here are based on MEC-sensors rather than on OPTs.

A majority of the fouling procedures for MEC-sensors resulted in the desired bias magnitude of 0 – 0.2 mg/L. For two of the initial fouling procedures, the average bias was larger than desired, due to excess of grease and large variations in the reference sensor's DO-level. These two data sets were not considered in the results.

In addition to the artificial fouling fault, a number of accidental faults occurred during the experimental period. Fortunately, we could store and study the impact of these faults on the IRs as an extension of the original scope of the study.

Variations in IRs at normal (non-faulty) conditions

Clear IRs were obtained for the MEC-sensors and a typical set of IRs at non-faulty conditions can be seen in Figure 1. The rise magnitude, i.e. the difference between initial and maximum DO-level during one IR, had a large variation (Figure 1). This variation was partly a consequence of a varied initial DO-level (Figure 2), but mainly illustrates that even for clean sensors, the shape of the IR had significant variations. Indeed, the variations increased towards the end of the IR. The increase is natural since the DO-level in the activated sludge basin (ASP) change during the time of the IR. We should therefore, in general, expect that the initial DO-level differ from the DO-level after an IR.

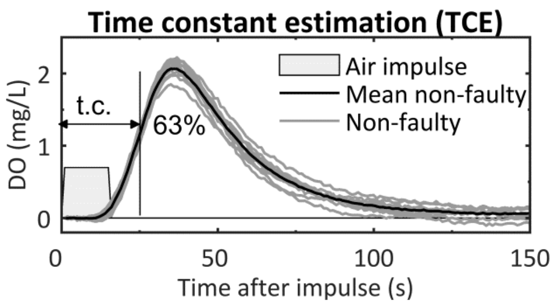

Figure 1. Ten impulse responses (IRs) from a non-faulty MEC-sensor (solid grey line) and their mean value (solid black line). All IRs were normalized by subtracting the intial DO-level from its IR. The time constant is indicated as the time from the start of the cleaning procedure (grey box, air impulse), to reach 63 percent of the peak value.

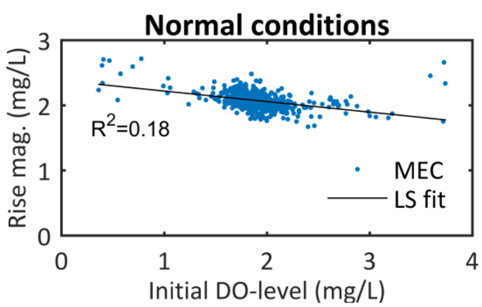

Figure 2. Correlation between the rise magnitude and initial DO-level for clean MEC-sensors (blue dots). The straight line least squares fit (solid black line) indicate a weak correlation between the initial DO-level and rise magnitude.

Detection of fouled membrane electrochemical sensors

The performances of GPR and TCE in detecting fouled MEC-sensors were evaluated with a ROC graph (Figure 3). The evaluation was based on the fouled data group that was obtained by measuring at different zones in the ASP.

Both FD-methods were significantly better than the Coin toss detector (CTD), and they had high detection rates (>0.9) at a low false alarm rate (<0.03). The GPR was slightly better than TCE since it reached maximum detection rate at a lower false alarm rate, as compared to TCE. The reason for GPRs slightly better performance is not obvious. One explanation could be that GPR compensated for varying initial DO-levels (Figure 4). However, as already noticed in Figure 2, the effect from varying initial DO-level was minor. We argue that it is more likely that the GPR successfully detected other changes in the IR, than changes in the time constant. As shown in Figure 5, the IR changed in many ways due to different faults, and there might be other indications than the time constant that distinguish a faulty IR from a normal one.

As indicated by the text arrows in Figure 3, a good trade-off between detection and false alarm rate was obtained for large values of the standard deviation setting (stds) for both methods. Recall that the stds was calculated from ten normal training IRs that should reflect the normal variation. Thus, a large stds indicate that the method allow a large normal variation before a detection. Both methods allowed large stdss but at the same time had a high detection rate. This indicates that the difference between normal IRs and fouled IRs was large. Still, the difference was obtained at small or no bias at all.

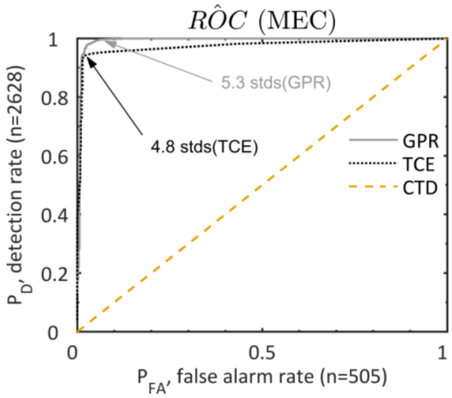

Figure 3. Estimated reciever operation characteristics (ROC) for Gaussian process regression (GPR)(solid grey line), Time constant estimation (TCE)(dotted black line), and Coin toss detector (CTD)(dashed yellow line) when applied to data from MEC-sensors. The MEC test data contained 505 clean (normal) and 2628 fouled (faulty) IRs. Arrows indicate number the standard devation setting (stds) in GPR and TCE to obtain the indicated false alarm and detection rates.

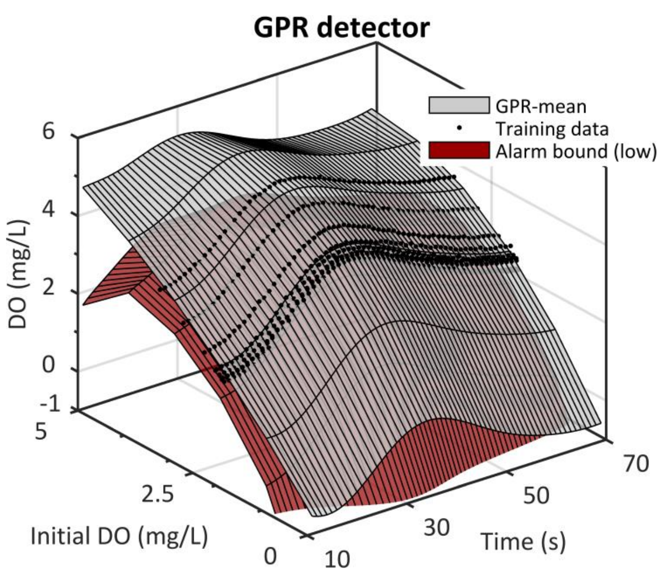

Figure 4. Illustration of Gaussian process regression detection method (GPR) that use ten training IRs (black dots) to predict a mean value (grey mesh) and alarm bounds (red mesh) for a new test IR. The upper alarm bound is not shown for illustrative reasons.

Diagnosis of sensor faults

During the experiments, several faults accidently occurred that also may happen during everyday plant operations. The mean of the faulty IRs are visualized in Figure 5.

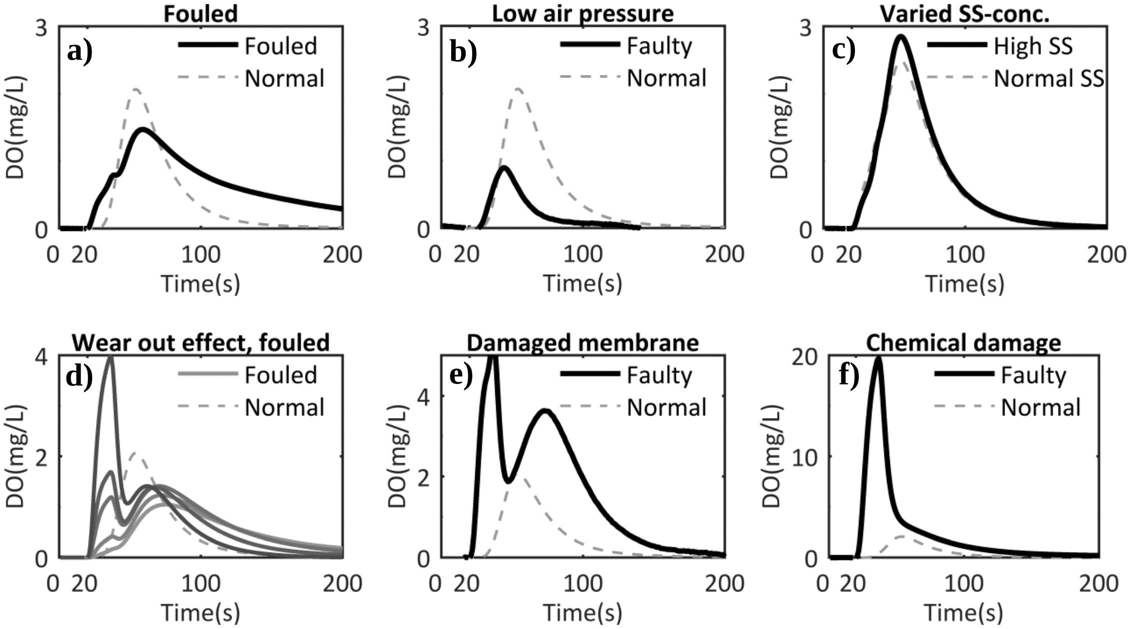

Figure 5. a – f) Mean IRs for various faults/ measurement conditions in the MEC-sensors. The air cleaning impulse was between time 20 – 35s. **a)** Fouled membrane (grease mixture). Mean IR from all fouled test data (solid line, n=2628). **b)** Low air pressure during the automatic air-cleaning procedure (n=9). **c)** The effect of increasing the SS-level from 2500 mg/L (normal SS, n=9) to 8500 mg/L (high SS, n=10). **d)** The gradual wear out effect by repeated IRs on a fouled MEC-sensor. An increased wear out of the membrane is indicated by darker grey (n=2180). **e)** A perforated membrane during manual cleaning (n=2). **f)** The effect of using harsh cleaning liquid to remove excess grease (n=98).

Fouling the MEC with grease mixture resulted in a dampened IR with an extended time to regain the original DO-level (Figure 5a). The small tweak during the impulse rise was not a typical pattern, but was merely an effect of combining IRs with two different shapes to the mean IR (data not shown). One set of the IRs displayed a “double peak behaviour” similar to Figure 5d), whereas one set had a straight increase with dampened IRs similar to Figure 5b).

In the end of the experimental period, a long term test for 11 days was conducted, were the MEC-sensor was fouled and subject to repeated cleaning events about 200 times per day. This was far more than the recommended amount of cleaning procedures, which resulted in a gradual change of the IR due to wearing out the membrane (Figure 5d). It is interesting that the shape of the IR changed from a dampened IR to exhibit an increasingly pronounced “double peak behaviour”. A double peak was also seen for the damaged membrane (perforated during manual cleaning) (Figure 5e).

We have no clear explanation for the double peaks although they seem to be present at both normal and faulty data, but more frequently were the membrane is mechanically damaged.
A decrease in air pressure for the automatic air cleaning system resulted in a dampened peak (Figure 5b), which was reasonable.
The measurements in the return sludge channel with high SS-level only indicated a slight increase in peak height, which was at the same order of magnitude as the normal variation (c.f. Figure 1 and 5c). This indicates that the SS-level had none or minor impact on the IR.

The high increase of the IR from cleaning the MEC with harsh cleaning liquid (Figure 5f) indicate that the membrane became more sensitive towards the IR, potentially since the diffusion of oxygen molecules was facilitated.

We should note that both the double peak behaviour (Figure 5 d) and the extreme peaks (Figure 5f) were also present in the data in the study by Andersson and Hallgren (unpublished data). The deviating patterns should be further studied and not be dismissed as outliers since they may be an early indication of a faulty membrane. Since a double peak significantly deviate from a normal IR, we would expect a smaller normal variation if they could be avoided during training, resulting in an increased detection performance.

CONCLUSION

The impulse response from the automatic air cleaning procedure was a good indication of the DO-sensor's status and was used to detect a fouled membrane electrochemical DO-sensor at a bias smaller than 0.2 mg/L. In addition, the shape of the IR clearly changed when other relevant faults were introduced which further makes the suggested approach promising for fault diagnosis as well. The results suggest that active fault detection is a promising methodology that can increase the reliability of measurements at WRRFs.

ACKNOWLEDGMENTS

The research leading to these results has received funding from Käppala Association, Syvab and Stockholm Water, Foundation for IVL Swedish Environmental Research Institute and the Swedish Water and Wastewater Association, which is gratefully acknowledged. Jesus Zambrano has received funding from the Knowledge Foundation (20140168) under the project More Sense, ABB, Mälarenergi, Flexiclean and Structor Miljöteknik.". We gratefully acknowledge support during the experiments by Simon Robertsson. We received most valuable comments for setting up the experimental plan by Erik Lindblom and Henri Spanjers. In addition, we would like to acknowledge Sofia Andersson and Fredrik Hallgren that provided data and experiences from previous experiments.

REFERENCES

Amand L., Olsson G. and Carlsson B. (2013). Aeration control - a review. *Water Sci Technol* **67**(11), 2374-98.

Andersson S. and Hallgren F. (2015). Sensor fault detection methods applied on dissolved oxygen sensors at a full scale WWTP. In: *9th Watermatex Congress*, Gold Coast, Australia.

Carlsson B. and Zambrano J. (2016). Fault detection and isolation of sensors in aeration control systems. *Water Science and Technology* **73**(3), 648-53.

Clark L. C. (1959). Electrochemical device for chemical analysis. In, U.S.

Demas J. N., DeGraff B. A. and Coleman P. B. (1999). Oxygen sensors based on luminescence quenching. *Analytical Chemistry* **71**(23), 793A-800A.

Esna Ashari A., Nikoukhah R. and Campbell S. L. (2012). Effects of feedback on active fault detection. *Automatica* **48**(5), 866-72.

Fawcett T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters* **27**(8), 861-74.

Rasmussen C. E. and Williams C. K. I. (2005). *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, Cambridge, Massachusetts

London, England.

Samuelsson O., Björk A., Zambrano J. and Carlsson B. (In press). Gaussian process regression for monitoring and fault detection of wastewater treatment processes. *Water Science & Technology*.

Spanjers H. and Olsson G. (1992). Modelling of the dissolved oxygen probe response in the improvement of the performance of a continous respirationmeter. *Wat. Res.* **26**(7), 945-54.

Svensson. A., Dahlin. J. and Schön. T. B. (2015). Marginalizing Gaussian process hyperparameters using sequential Monte Carlo. In: *6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Cancun, Mexico.

Yoo C. K., Villez K., Van Hulle S. W. H. and Vanrolleghem P. A. (2008). Enhanced process monitoring for wastewater treatment systems. *Environmetrics* **19**(6), 602-17.