An efficient monitoring concept with control charts for on-line sensors

M. Thomann, L. Rieger, S. Frommhold, H. Siegrist and W. Gujer
Swiss Federal Institute for Environmental Science (EAWAG) and Swiss Federal Institute of Technology (ETH), Überlandstr. 133, PO Box 611, CH-8600 Dübendorf, Switzerland (E-mail: thomann@eawag.ch; rieger@eawag.ch)

Abstract A monitoring concept for on-line sensors will be discussed which helps the WWTP staff to detect drift-, shift- and outlier effects as well as unsatisfactory calibration curves. The approach is based on the analysis of comparative measurements between the sensor and a reference method. It combines statistical analysis such as control charts and regression analysis with decision support rules. The combination of two different detection levels in the selected Shewhart control charts with additional criteria allows one to detect “out-of-control” situations early with an optimized measurement effort. Beside the statistical analysis the concept supports the operator with a graphical analysis to monitor the accuracy of on-line measurements efficiently. The widely applicable monitoring concept will be illustrated with examples for an ion-sensitive NH4+ and a MLSS-sensor.

Keywords Bias; control charts; data analysis; monitoring; on-line sensors; uncertainty

Introduction
In recent years the importance of online measurements on WWTPs has noticeably increased. The main applications are control strategies and continuous monitoring of the effluent quality standards. However, the new measurement technology is covering the inherent danger of non-identified bias, which can cause non-optimal control decisions or serious safety problems. Although there are nowadays some reliable sensors available for different kinds of nutrients, there is a need for concepts to deal with the quality assurance of these measurements in practical use (Olsson and Newell, 1998). It is common to control the on-line-sensors with grab samples (Baumann, 2001). In order to enhance their accuracy, the sensor signal should be analysed for drift and shift effects as well as gross errors (Figure 1). Control charts are powerful tools to solve this problem. They are widely used in analytical laboratories (Montgomery, 1996) but not very popular on WWTPs. Beside the above mentioned “out-of-control” situations the WWTP staff must detect insufficient calibration curves. There is a need for a concept, which combines these monitoring tasks and helps the operator to quantify the accuracy of the sensors. This information will be valuable for adequate control strategies and for the decision whether a correction of the measurement process is necessary. Hence, this paper is dealing with the following topics:
• Detection of bias in on-line measurements with a reference method and control charts.
• Decision rules for purposeful and efficient service actions and new calibration routines.
• Quantification of the accuracy of the on-line sensors.

Monitoring concept
The main goal of the concept is an in-control measuring process, which has to prove that the measurements are within a certain uncertainty range. The concept approach to detect out-of-control situations is based on the analysis of the differences between sensor values and corresponding grab samples measured with a reference method (ISO/CD, 2000). The
The monitoring phases consist of a warning phase, an alarm phase and a calibration analysis phase.

In the warning phase the comparative measurements between the sensor and the reference method will be investigated with a Shewhart control chart with appropriate out-of-control criteria. In contrast to the normally used Western Electric rules (Montgomery, 1996) the criteria of this warning phase are more stringent (Figure 3) in order to detect a potential out-of-control situation early. In case of a violation of the warning phase criteria, the cause for the failure will be sought, service actions (e.g. cleaning) performed and the comparative measurement will be immediately repeated.

If the warning situation remains after a service action and an immediate comparative measurement, a potential out-of-control situation is assumed and the alarm phase will be initiated. The goal of this phase is to validate the out-of-control hypothesis as early as possible with a low probability of a false alarm. A false alarm means that the alarm phase indicates an out-of-control situation although the process is still in-control. In order to speed up the detection of an out-of-control situation, the frequency of comparative measurements will be increased. The frequency remains high until an out-of-control situation is detected or no violation of the warning phase criteria occurs anymore (Figure 2). With the approach of two different control levels it is possible to increase the normal measurement frequency only in case of an indicated potential out-of-control situation.

![Monitoring approach for on-line sensors](image-url)

**Figure 2** Monitoring approach for on-line sensors. The bold arrow cycle shows the sequence of actions undertaken if the measuring process is assumed to be in-control.
When the warning phase does not show any violation of the criteria, one can presume that the process is not influenced by drift-, shift- or outlier effects. This is the precondition for the calibration analysis. The comparative measurements will be analysed with regression analysis in order to detect a potential insufficient calibration curve. If the tests do not indicate any calibration problem an in-control situation is assumed. The comparative measurements will be performed with the normal frequency. In case of a detected poor calibration curve, a new calibration of the sensor should be calculated.

**Preliminary tests**

*Test of the reference method.* It should be validated that the reference method is not influenced by systematic errors (Häck, 1999; ATV, 2000) with the following control experiments:

- Analysis of the calibration characteristics with standard solutions (ISO 1990).
- Standard addition experiment for the analysis of potential interferences (DIN 1998).
- Comparison of the investigated method with an alternative standard method (ATV 1997).

*Basic calibration on wastewater matrix.* Before the long-term behaviour of the sensor can be monitored under measuring conditions, the in-control situation must be validated with laboratory experiments (ISO/CD, 2000) and accuracy experiments under real operating conditions. For the comparability of different sensors the manufacturers mainly guarantee a certain precision performance under standard conditions (Baumann, 2001). The accuracy test of the sensor under real operating conditions has to be done by the WWTP staff. The following tests are proposed:

- Determination of the delay time of the sensor $T_{90}$ and the sample preparation (ATV 2000, ISO/CD 2000). This value is important for the comparability of sensor with a reference method.
- A standard addition (ISO, 1998) to determine the sensor’s trueness under measuring conditions.

**Drift, shift and outlier analysis**

The difference between the sensor and the reference value serves as a control variable in order to detect drift-, shift- or outlier effects (Figure 1).

$$D = C_{ref} - C_{meas}$$

Based on the preliminary tests for the accuracy of the reference method and the sensor, it will be assumed that the control variable is normally distributed with the mean $\mu = 0$ and the variance $\sigma^2$. The standard deviation of the differences $s_D$ should be estimated from a dataset of at least $n = 10 - 20$ samples (Cheeseman and Wilson, 1978; DIN, 1998).

$$s_D = \sqrt{\frac{n}{n-1} \sum_{i=1}^{n} (D_i - \overline{D})^2}$$

$$\overline{D} = \frac{1}{n} \sum_{i=1}^{n} D_i$$

With $s_D$ the upper and lower warning- and control limits (Montgomery, 1996) can be calculated and the Shewhart control charts with additional run rules for the warning and alarm phase can be constructed. The different out-of-control criteria for both phases will be shown in Figure 3.
Lower warning limit $LWL = -2 \cdot s_D$
Upper warning limit $UWL = +2 \cdot s_D$
Lower control limit $LCL = -3 \cdot s_D$
Upper control limit $UCL = +3 \cdot s_D$

Despite the fact that the combination of the Shewhart control chart with run rules (criteria b and c in the warning phase and criteria c and d in the alarm phase, Figure 3) enhances the sensitivity to detect drift and shift effects, the probability of a false alarm is also increased (Champ and Woodall, 1987; Montgomery, 1996). With a stochastic simulation the average number of measurements has been calculated until the charts indicate an alarm when the process is actually in-control. For the warning phase this so-called in-control average run length $ARL_0$ was 8, for the alarm phase it became 98. The low value for the warning phase does not have serious consequences because a mistakenly indicated warning situation leads to a service action and an immediately repeated comparative measurement. Assuming that the comparative experiments will normally be performed once a week (ATV, 2000) a false alarm situation occurs every second year on average.

The validation of the chosen limits (Figure 3) should be carefully investigated for the proposed Shewhart control charts. The problem arises because the experiment for the calculation of the standard deviation of the differences $s_D$ is mostly performed at one level of the sensors working range. In order to test the validation of this estimate for the whole working range, the homogeneity of variances should be tested with an additional experiment and a F-test. After the application of the concept, it is suggested that the limits
be recalculated after every twentieth comparative measurement if no out-of-control situation has been detected since the last calculation.

**Calibration analysis**

With the proposed adapted Shewhart control charts drift-, shift-, and outlier effects (Figure 1) are analysed. On the other hand it could happen that a poor calibration curve does not lead to an out-of-control situation in the alarm phase (situations 1 and 2 in Figure 4). As a criterion to assess a potential calibration error the sensor value should be plotted against the measurement of the reference method (Figure 4). If no false calibration curve exists the expected sensor value for a corresponding reference value should vary randomly around a straight line with slope $\beta = 1$ and offset $\alpha = 0$. To analyse this relationship the following linear regression model will be applied to the comparative measurement dataset.

$$y = \beta \cdot x + \alpha + \xi$$

With the application of this model it will be assumed that the reference measurements are not influenced by systematic errors and that their random uncertainty is small compared with those of the sensor. After the minimal number of five datapoints (ISO 1990) since the last basic calibration it should be investigated if there is a significant relationship between the sensor value and the reference measurement. Hence, the null hypothesis $H_0: \beta = 0$ should be tested against the alternative hypothesis $H_A: \beta \neq 0$ with a t-test on the 5% significance level (Draper and Smith, 1981).

In case of a significant relationship, the estimated parameters $\hat{\beta}$ and $\hat{\alpha}$ and their standard errors $se(\hat{\beta})$ and $se(\hat{\alpha})$ will be analysed with the following t-tests (Draper and Smith, 1981). On the 5%-significance level the following hypotheses for $\beta$ and $\alpha$ should be tested (Table 1).

If an out-of-control situation has been detected in the alarm phase, the question would arise, which cause has triggered the alarm. Although the solution of this problem has not yet been implemented in the concept, the calibration analysis can be a helpful tool to decide whether a calibration problem has caused the alarm as will be shown in the example for the MLSS-sensor.

**Results and discussion**

**Monitoring of an ion-sensitive NH$_4^+$-sensor**

At the WWTP Morgental (Switzerland) an ion-sensitive NH$_4^+$-sensor (Nadler Chemische Analysentechnik, Zuzwil, Switzerland) in the primary effluent has been monitored over a period of four months. (NH$_4^+$+NH$_3$)-N in the grab samples has been analysed with an indophenol blue method and a spectrophotometer (CADAS 30, LCK 303, Dr. Lange).

![Figure 4](image-url) **Figure 4** Out-of-control situations due to a poor calibration curve
Ammonium (NH4+–N) was calculated from the membrane-filtered (0.45 µm) grab samples with pH and temperature. A Merck NH4+-reference solution with 1,000 ± 2 mg NH4+ L–1 has been used for standards and standard additions. All dilutions have been performed with nanopure water.

Preliminary tests. The used indophenol blue method has been compared with a Flow Injection Analysis method, FIA (ASIA, Ismatec AG, Glattbrugg, Switzerland). Eight standard solutions and standard addition samples have been analysed simultaneously with both methods (Table 2).

The precision of both methods is comparable (Table 2). Although the indophenol blue method seems to be less precise than the FIA method, the recovery rate of 97.5% for ammonia validates the assumption that there are no significant interferences for the investigated wastewater matrix. A comparison of the measured values of both methods for the standard addition samples did not show any significant differences. Because the methods are based on different analytical principles, it could be concluded that there are no constant systematic errors. The indophenol blue method can be used as reference method with sufficient certainty. The ion-sensitive NH4+-sensor (Rieger et al., 2002) has been investigated (Table 3) with a calibration experiment (ISO/CD, 2000) and a standard addition (DIN, 1998).

The precision is satisfactory because the coefficient of variation was smaller than the recommended 3% (ATV, 2000). The trueness under measuring conditions is equally

### Table 1  T-tests for the analysis of the calibration curve

<table>
<thead>
<tr>
<th>Null hypothesis $H_0$</th>
<th>Alternative hypothesis $H_A$</th>
<th>Test statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>slope $\beta = 1$</td>
<td>$\beta \neq 1$</td>
<td>$t = \frac{\hat{\beta} - 1}{s(\hat{\beta})}$</td>
</tr>
<tr>
<td>offset $\alpha = 0$</td>
<td>$\alpha \neq 0$</td>
<td>$t = \frac{\hat{\alpha}}{s(\hat{\alpha})}$</td>
</tr>
</tbody>
</table>

### Table 2  Results of the test experiments for the later used reference method

<table>
<thead>
<tr>
<th></th>
<th>Dr. Lange (LCK 303)</th>
<th>FIA-Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td>Coefficient of variation $V_{\alpha} [%]$</td>
<td>1.49</td>
</tr>
<tr>
<td>Trueness</td>
<td>Recovery rate of NH4+-N [%]</td>
<td>97.5</td>
</tr>
<tr>
<td></td>
<td>Standard addition with 0.45 µm filtered wastewater (DIN, 1998)</td>
<td>(+/- 7.5)</td>
</tr>
</tbody>
</table>

### Table 3  Results of the laboratory calibration experiments for an ion-sensitive NH4+-sensor

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Response time</td>
<td>[sec]</td>
<td>&lt; 5</td>
<td></td>
</tr>
<tr>
<td>Linearity (tested range)</td>
<td>[mgN L–1]</td>
<td>4.66 – 32.61</td>
<td></td>
</tr>
<tr>
<td>Coefficient of variation $V_{\alpha}$</td>
<td>[%]</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>Repeatability (at lower end of working range)</td>
<td>[mgN L–1]</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>[mgN L–1]</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>Trueness (at lower end of working range)</td>
<td>[mgN L–1]</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td>[mgN L–1]</td>
<td>-0.1</td>
<td></td>
</tr>
<tr>
<td>Recovery rate for NH4+-N</td>
<td>[%]</td>
<td>98.7</td>
<td>(+/- 2.0)</td>
</tr>
<tr>
<td>(with 95%-confidence interval)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
acceptable because the recovery rate in a standard addition experiment with primary clarified wastewater was 98.7%.

The standard deviation of the differences $s_D = 0.62 \text{ mgN L}^{-1}$ ($\bar{D} = 19.1 \text{ mgN L}^{-1}$) has been calculated with fourteen comparative measurements. The homogeneity of the variances has been checked for a higher level ($\bar{D} = 44.3 \text{ mgN L}^{-1}$, $s_D = 1.01 \text{ mgN L}^{-1}$, $n = 5$). The F-test did not show any significant differences between the variances (5% significance level). Therefore, the limits for the control charts calculated with $s_D = 0.62 \text{ mgN L}^{-1}$ are assumed to also be valid for higher concentrations up to 44.3 mgN L$^{-1}$.

**Drift-, shift- and outlier analysis.** Monitoring the sensor (Figure 5), an out-of-control situation was detected early. After a new calibration the alarm situation was reached again. A close investigation of the sensor showed that the reference electrode had been affected by iron precipitation products which caused the drift effect. A change of the measuring position leads to measurements without any disturbance (12 June–10 Aug). Warning situations occurred twice, but after a cleaning and a repetition of the comparative measurement, no out-of-control situation could be affirmed. The third warning situation (13 Aug) lead after repeated measurements to an early detection of an out-of-control situation (15 Aug). After a change of the sensor membrane the process was in-control again.

Due to the measurement principle of the ion-sensitive electrode a periodic calibration interval was necessary (14 days on average). A recalculation of the standard deviation of the differences $s_D = 0.65 \text{ mgN L}^{-1}$ with the in-control data (12 June–10 Aug) showed that the chosen limits for the control charts can be used for further monitoring. Based on this in-control period a trueness of 0.12 mgN L$^{-1}$ under field conditions (working range 4.5–24.2 mgN L$^{-1}$) has been calculated with the mean of the differences (ISO/CD, 2000). The example shows the advantage of the two different control levels. The out-of-control situations could be detected early. An increased measuring effort was only required after a warning situation occurred.

**Monitoring of a MLSS-sensor**

At the WWTP Thunersee (Switzerland) the performance of a MLSS-Sensor (Cosmos D-CO-1202, Züllig AG, Rheineck, Switzerland) in the aerated section of an activated sludge tank was analysed by the WWTP staff with gravimetrical analysis and paper filters over a period of four months. Based on this large dataset of routine measurements, the shown concept will be tested as a hypothetical example because the concept has not yet been applied on this WWTP. The goal of the example is to illustrate the additional information, which will be gained from the calibration analysis when an out-of-control situation occurs.

Before the control period the sensor had been calibrated with a two-point in-situ-calibration. To set up the limits for the control charts, the standard deviation of the differences between the sensor signal and the gravimetrical analysis was calculated with thirteen comparative data pairs (Figure 6).

After the fourth control measurement a warning situation occurred due to four consecutive positive signs. Despite the inspection of the sensor and the repeated measurement the warning situation remained until day 25 when an out-of-control situation was detected due to exceeding the upper control limit. The staff had to decide if it was caused by a drift effect or a poor calibration. The calibration analysis could be helpful. Because the range of the sensor value was small (e.g. 2–3.5 g L$^{-1}$) before the out-of-control detection, no significant relationship between the sensor and the gravimetrical analysis could be identified at that time. To solve this problem and to assess the calibration characteristics, an in-situ calibration experiment had to be performed by changing the sludge mass in the analysed aeration tank. The performed calibration analysis detected a poor calibration curve. With
Figure 5  Application of the monitoring concept to an ion-sensitive NH$_4^+$-sensor
Figure 6 Hypothetical monitoring example with real data of a MLSS-sensor in an activated sludge tank.
the parallel recorded raw signals of the sensor it was shown that a nonlinear calibration curve had to be applied. After the application of the fitted second order calibration curve no further out-of-control situation occurred. The example shows how the calibration analysis can be helpful to search for the cause of a detected out-of-control situation. Furthermore, the importance of the sensors raw signal for a new calibration should be emphasized.

**Conclusions**

The concept allows detection of drift-, shift- and outliers as well as poor calibration of the sensor. With the new approach of Shewhart control charts with two different levels, it is possible to detect an out-of-control situation early. An increased control frequency will only be required in case of a detected warning situation. An early detection and optimization of the measuring effort is obtained.

Because the control chart technique allows the statistical as well as a fast graphical analysis of the measurement process it is a powerful tool for data analysis. The implementation of the concept in a software environment with a database is under construction. This would enhance the efficiency of data analysis and guarantee that relevant information such as comparative measurements, raw signals of the sensor, service actions or calibration characteristics will be immediately available for the operator.

The approach helps the WWTP staff not only to detect out-of-control situations, it also gives some decision support for appropriate service actions (e.g. calibration analysis).

The concept is widely applicable for measurements on WWTPs where a reference method exists, which is not influenced by systematic errors. Beside the demonstrated examples it can be applied to all kinds of on-line analyzers, sensors or sampling devices.

Applying the concept the staff gain information about the accuracy of the on-line measurements. It is also valuable information for control strategies and for uncertainty calculation in simulation studies. Additionally, it helps to assess the profitability of complex control concepts.

**Acknowledgements**

The authors would like to thank U. Achermann, P. Glaus, F. Haas, B. Strupler, K. Wiget, J. Hohl from WWTP Morgental, A. Baumann and J. Hohler from WWTP Thunersee, J. Nadler, M. Böhler (EAWAG) and K. Rottermann (EAWAG) for their generous support.

**References**


