1601 © IWA Publishing 2008 Water Science & Technology—WST | 57.10 | 2008

Quality evaluation methods for wastewater treatment plant data

M. Thomann

ABSTRACT

Non identified systematic errors in data sets can cause severe problems inducing wrong decisions in function control, process modelling or planning of new treatment infrastructure. In this paper statistical methods are shown to identify systematic errors in full-scale WWTP data sets. With a redundant mass balance approach analyzing five different mass balances, systematic errors of about 10%–20% compared to the input fluxes can be identified at a 5%-significance level. A Shewhart control-chart approach to survey the data quality of on-line-sensors allows a statistical as well as a fast graphical analysis of the measurement process. A 19 month data set indicates that NO_3^-, PO_4^- and NH_4^- on-line analyzers in the filter effluent and MLSS sensors in the aeration tanks were not disturbed by any systematic error for 85–95% of the measuring time. The in-control-interval (±3·standard deviation) has a width of ±12–17% (NO_3—N), ±35–40% (PO_4—P), ±83% (NH_4—N) and ±12–15% (TS) of the measured reference value.

Key words | control charts, error propagation, sensors, statistical methods, uncertainty, wastewater treatment

M. Thomann
Holinger AG,
Galmsstrasse 4,
CH-4410 Liestal,
Switzerland
E-mail: *michael.thomann@holinger.com*

INTRODUCTION

Normally, a lot of measurements are generated on a WWTP. These data provide a basis for the operation of the plant, function control, process modelling, planning of future treatment facilities and economic evaluations. As a consequence, the quality of all these actions is directly dependent on the accuracy of the data. Without performing a data quality evaluation the information content of the used data is considerably restricted. The main goal of the data quality evaluation is to identify systematic and gross errors. Data quality evaluation procedures with full-scale WWTP data have to cope with several challenges. First of all, a large number of measurement parameters have to be analyzed. The effort to provide additional measurements for the data quality analysis should be as minimal as possible. Secondly, full scale WWTP data show different sampling frequencies of various parameters at several locations. Thus, the quality control methods should be able to check whole data sets. This paper shows statistical methods to analyze systematic errors in full-scale WWTP data sets. Beside a mass balance approach to analyze several mass fluxes, a control-chart method will be shown for the evaluation of the data quality of on-line measuring devices.

doi: 10.2166/wst.2008.151

EVALUATION OF MASS BALANCES

Methodology

Mass balance calculations for COD, total nitrogen and total phosphorus provide useful information about data accuracy of WWTP (e.g. Nowak *et al.* 1999). COD and nitrogen balances for an aeration tank system face the fact that redundant data sets are normally not available. Thus, further model assumptions are necessary to create redundant data sets and therefore contradictions in the balances can be used to evaluate systematic errors. Some authors propose an analysis of redundant mass balances by a least-square-approach with linear constraints in order to detect gross errors (Van der Heijden *et al.* 1994; Meijer *et al.*

2001). An alternative approach to detect systematic errors is to apply stochastic analysis of contradictions of mass balances (Coleman & Steele 1998). This approach is interesting for full scale WWTP data sets because it is not bound to any frequency distribution of the input- and output-loads. Furthermore the approach can easily be combined with Monte-Carlo-simulations in order to implement model assumptions in the statistical analysis (Thomann 2002b). For practical purposes, the proposed approach will be normally applied only to full redundant mass balances based on real measurements. The following equation reflects a mass balance for the substance C (Equation 1) over the system of a completely stirred reactor (Figure 1).

$$\frac{V \cdot dC}{dt} = Q \cdot (C_{in} - C) + r_C \cdot V \quad (1)$$

Assuming a steady-state-situation, which is a valuable assumption balancing a data set of a whole year, the accumulation term on the left side of Equation 1 can be neglected. In case a conservation equation is valuable for the substance C (e.g. phosphorus, nitrogen, COD, iron) the reaction term can be neglected as well (Equation 2).

$$0 = \overbrace{Q \cdot C_{in}}^{\text{input load}} - \overbrace{Q \cdot C}^{\text{output load}} \quad (2)$$

In real applications this difference is never zero. Because of random and systematic errors a redundant mass balance shows always a difference D (Equation 3).

$$\overbrace{D}^{\text{difference}} = \overbrace{Q \cdot C_{in}}^{\text{input load}} - \overbrace{Q \cdot C}^{\text{output load}} \quad (3)$$

To find out whether the difference D is caused by systematic errors one need to analyze its stochastic properties and describe its probability density function $f(D)$ (see for example Figure 2). Considering the input and

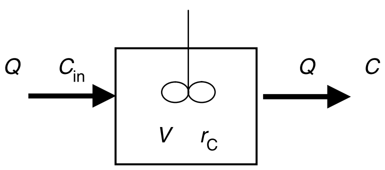

Figure 1 | System of a completely stirred reactor.

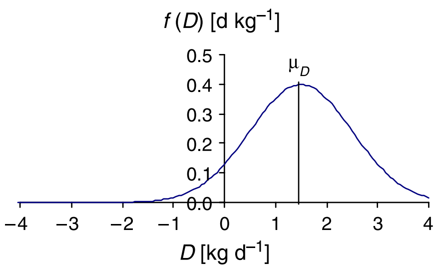

Figure 2 | Example of the probability density function $f(D)$ of the difference D.

output load as independent random variables X_{in} and X_{out}, the expected value of the difference μ_D can be calculated (Equation 4).

$$\mu_D = \mu_{in} - \mu_{out} \quad (4)$$

Assuming normally distributed independent random variables X_{in} and X_{out} the best estimates of the expected values is the arithmetic average (Equation 5 and 6). Therefore the difference D can be calculated (Equation 7). Applying the Gaussian error propagation law to the mass balance Equation (Equation 3), the variance of the difference σ_D^2 can be calculated (Equation 8).

$$\bar{x}_{in} = \frac{1}{n_{in}} \sum_{i=1}^{n_{in}} x_{in,i} \quad (5)$$

$$\bar{x}_{out} = \frac{1}{n_{out}} \sum_{i=1}^{n_{out}} x_{out,i} \quad (6)$$

$$D = \bar{x}_{in} - \bar{x}_{out} \quad (7)$$

$$\sigma_D^2 = \sigma_{\bar{x}_{in}}^2 + \sigma_{\bar{x}_{out}}^2 \quad (8)$$

With a stochastic simulation the probability density function $f(D)$ can also be calculated for not normally distributed input and output loads. Time series and expert knowledge can be implemented as well (see Thomann 2002b). Because the density function $f(D)$ is not always normally distributed, a test approach is provided which is valuable for any density function $f(D)$. The null hypothesis H_0 and the alternative hypothesis H_A are shown in Equation 9 and 10.

$$H_0 : \mu_D = 0 \quad (9)$$

$$H_A : \mu_D \neq 0 \quad (10)$$

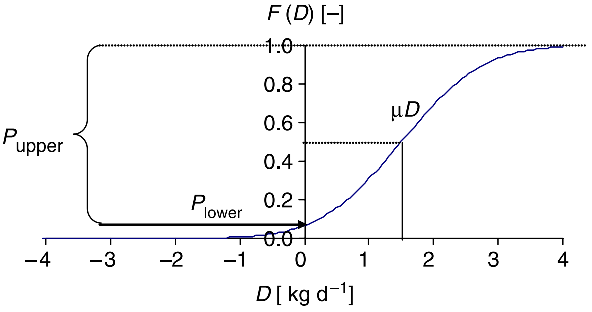

Figure 3 | Cumulative distribution function of the difference D.

The criteria for a rejection of the null hypothesis H_0 can be formulated (considering a certain level of significance α) calculating the probabilities P_{lower} and P_{upper} (Equation 11 and 12 see Figure 3). The null hypothesis H_0 will be rejected if one of the criteria in Equation 13 or Equation 14 is fulfilled.

$$P_{lower} = P\langle D \le 0 \rangle = \int_{-\infty}^{0} f\langle D \rangle \cdot \mathrm{d}D \tag{11}$$

$$P_{upper} = P\langle D \ge 0 \rangle = 1 - P_{lower} \tag{12}$$

$$\text{Rejection of } H_0 \text{ if} : P_{lower} \le \frac{\alpha}{2} \tag{13}$$

$$\text{or} : P_{upper} \le \frac{\alpha}{2} \tag{14}$$

Plant description - WWTP Ergolz 1

The WWTP Ergolz 1 (Figure 4) is a municipal sewage plant built for 30,000 population equivalents. It consists of a physical treatment, biological nitrogen removal (nitrification and denitrification), simultaneous chemical phosphorus precipitation and a final sand filtration. The sludge is treated with an anaerobic mesophilic digester. Applying the shown approach to full-scale WWTP data, five different mass balances can be evaluated (Figure 4). Detailed measurements show that the filter backwash water can be neglected for phosphorus and iron balances (<1% of the sum of input fluxes to the WWTP). The chosen analytical effort for the analysis of the mass balances is shown in Table 1.

The iron concentration in the influent is the only parameter which is not measured. Measurements on several municipal WWTP's (Thomann 2002*b*) showed that iron concentrations in the influent can accurately be modelled with an uniform probability density function with mean of 1 mg_{Fe}/l and standard deviation of 0.5 mg_{Fe}/l.

RESULTS AND DISCUSSION

The phosphorus mass balances over the whole WWTP and aeration tank (Table 2 and Figure 5) show highly significant differences from zero considering a significance level $\alpha = 5\%$.

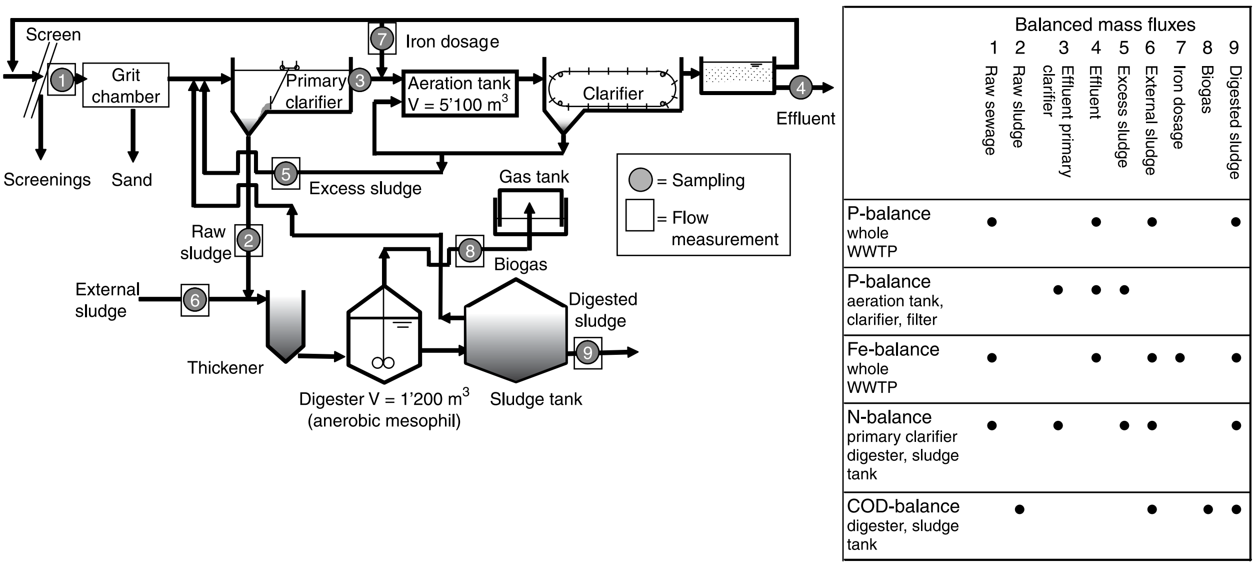

Figure 4 | WWTP Ergolz 1 with analyzed mass balances.

Table 1 | Analytical effort 2004 until 2006 at WWTP Ergolz 1 [samples per year]

	Analyzed mass fluxes								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Q	Daily	Daily			Daily	No external sludge	Daily	Daily	Daily
TS		52			Daily				Daily
P	40		71	72	11				8
N	35		35	35	8				8
COD	13	12		6	8				5
DOC/TOC	6			71					
CO_2								50	
Fe	Estimated (see above)			4			Every filling		4

The analysis of the three years shows that the results are reproducible. The differences of balances a and b are all positive and their calculated standard deviations are all in the same range. Because the iron influent flux of the WWTP is small and the measurement of the iron dosage is mostly accurate it is possible to identify with the iron balance systematic errors in the digested sludge flux. In this case study the iron balances do not show significant differences on the 5%-significance level in 2004 and 2006 (see balance c in Table 2 and Figure 5). Thus, the hypothesis seemed reasonable that the systematic error in the phosphorus balances is caused by the influent discharge measurement. It could be shown with a control experiment filling an empty tank (data not shown) that the flowmeters in the influent of the WWTP overestimates the discharges for about 19% on average. This systematic error explains the main part of the significant differences in the phosphorus balances in Figure 5. Other highly significant differences from zero have been identified in the COD balances of the years 2004, 2005 and 2006 over the digester and the sludge tank (balance e in Table 2 and Figure 5). An analysis of the causes in 2004 showed that the raw sludge COD flux was biased due to gross errors in the COD analysis. After changing the raw sludge sampling and analytical procedure the COD balances in 2005 and 2006 showed much smaller but still significant differences (22% of influent COD flux). In 2007 the gas flowmeters will be analyzed in detail in order to search the cause for the significant difference. Furthermore, it could be calculated with the standard deviations of the differences in Table 2 which systematic error value can be identified with the shown mass balance approach and the chosen analytical effort (see Table 1). Systematic errors between 10% and 20% compared to the sum of the input fluxes of the mass balances can be detected with the shown approach considering a 5%-significance level.

CONTROL-CHART ANALYSIS

Methodology

Automatic concentration measurement devices can be disturbed by drift, shift and outliers effects. These out-of-control situations occur because of systematic or gross errors. The main goal of the shown control-chart concept (for details see Thomann *et al.* 2002*a*) is to prove that the on-line-measurements are not disturbed by one of the above described errors. In case the sensor is in the so called in-control situation only random effects cause differences between sensor and control measurements. On-line-sensors are normally controlled with grab samples measured with a reference method which can be assumed to be free of bias (ISO 2003). For testing this assumption one can perform different analytical experiments which are described in detail by Thomann (2002*b*). The difference D between sensor and reference value serves as a control variable in order to detect drift-, shift- or outlier effects (Equation 15).

$$D = C_{ref} - C_{meas} \quad (15)$$

Assuring that the sensor is properly calibrated and the reference is free of bias we can assume that the control variable D is normally distributed with the mean $\mu = 0$ and the variance σ^2. The standard deviation s_D (Equation 16)

Table 2 | Results of mass balances for WWTP Ergolz 1 (Arithmetic averages of the fluxes ± standard deviation of the arithmetic average) The difference *D* is *t* calculated according Equation 7. The filter back washing flux can be neglected because it is <1% of the sum of input fluxes in balance a, b and c. Standard deviation of the difference is calculated with the gaussian error propagation law according to Equation 8 taking all input and output fluxes of the balance into account

			1 Raw sewage	**2** Raw sludge	**3** Effl. primary clarifier	**4** Effluent	**5** Excess sludge	**6** External sludge	**7** Iron dosage	**8** Biogas	**9** Digested sludge	Difference *D*
a)	P-balance	2004	68.5 ± 3.1			−4.6 ± 0.2		0			−40.6 ± 2.3	23.3 ± 3.9
	[kg_P d^{-1}]	2005	71.2 ± 2.9			−5.3 ± 0.2		0			−36.7 ± 2.5	29.2 ± 3.8
	Whole WTTP	2006	74.3 ± 3.0			−5.5 ± 0.2		0			−42.7 ± 3.7	26.1 ± 4.8
b)	P-balance	2004			45.2 ± 1.3	−4.6 ± 0.2	−33.5 ± 1.3					7.1 ± 1.8
	[kg_P d^{-1}]	2005			47.6 ± 1.2	−5.3 ± 0.3	−30.5 ± 1.2					11.8 ± 1.8
	Aeration tank, clarifier, sand filter	2006			47.7 ± 1.4	−5.5 ± 0.2	−35.5 ± 2.8					6.7 ± 3.1
c)	Fe-balance	2004	14.0 ± 4.1			−1.1 ± 0.3		0	63.6 ± 3.2		−66.5 ± 5.0	10.0 ± 7.2
	[kg_{Fe} d^{-1}]	2005	14.0 ± 4.2			−1.1 ± 0.2		0	83.0 ± 4.2		−68.6 ± 6.4	27.3 ± 8.6
	Whole WWTP	2006	16.4 ± 4.7			−0.9 ± 0.2		0	50.1 ± 2.5		−64.6 ± 12.0	1.0 ± 13.2
d)	N-balance	2004	409.7 ± 21.5		−401.1 ± 20.9		75.4 ± 4.3	0			−55.4 ± 5.1	28.6 ± 30.7
	[kg_N d^{-1}]	2005	441.9 ± 46.3		−424.7 ± 27.8		66.7 ± 6.7	0			−55.7 ± 5.2	28.2 ± 54.7
	Primary clarifier, digester, sludge tank	2006	375.7 ± 16.4		−348.4 ± 16.3		81.1 ± 7.2	0			−72.8 ± 6.3	35.6 ± 25.0
e)	COD-	2004		4,900 ± 245				0		−1,453 ± 23	−1,207 ± 102	2240 ± 266
	balance	2005		3,155 ± 179				0		−1,397 ± 32	−1,018 ± 67	740 ± 194
	[kg_{COD} d^{-1}] digester, sludge tank	2006		3,327 ± 197				0		−1,643 ± 38	−941 ± 72	743 ± 213

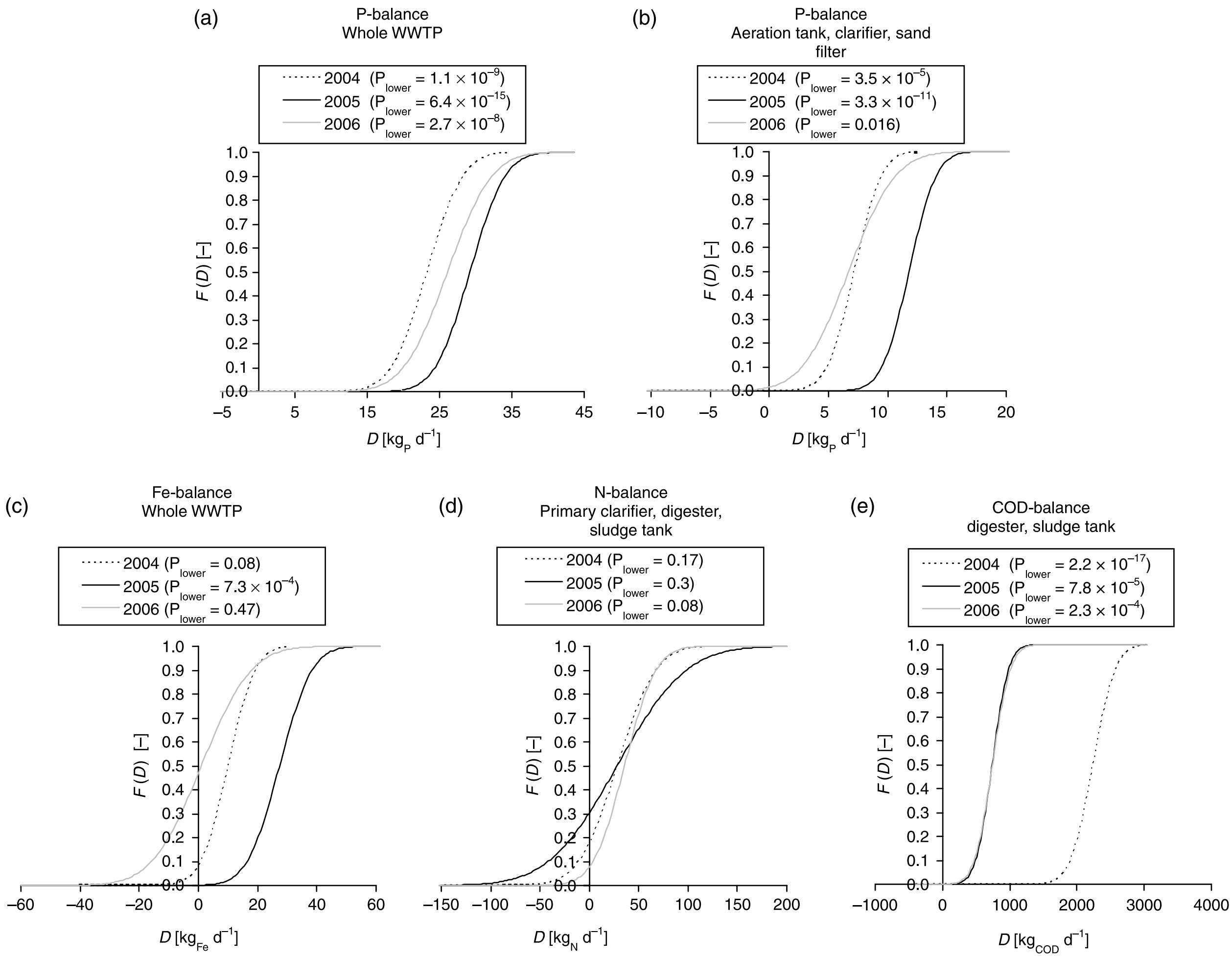

Figure 5 | Cumulative distribution function of the difference *D* for five evaluated mass balances.

should be estimated from a data set of at least $n = 10$ samples.

$$s_D = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \bar{D})^2}{n-1}} \tag{16}$$

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i \tag{17}$$

With s_D the upper and lower warning- and control-limits (Montgomery 1996) can be calculated and the Shewhart control-charts for the warning and alarm phase can be constructed. The different out-of-control criteria for both phases will be shown in Figure 6. The warning criteria (left side in Figure 6) are chosen to have an early warning criterion. If such a warning occurs the control measuring effort will be enhanced by the operator. The two stage control chart concept allows to optimize the measuring effort of WWTP operators.

Results and discussion - WWTP Zürich Werdhölzli

The WWTP Zürich-Werdhölzli is the largest municipal WWTP in Switzerland (capacity 670,000 population equivalents). In the biological nitrogen removal step (activated sludge system with nitrification/denitrification and chemical phosphorus removal) the data quality of about 73 on-line and in-line-sensors is surveyed continuosly. Since January 2006 the software GEKO (2005) has been used to perform the two step Shewhart control-chart analysis (Figure 7).

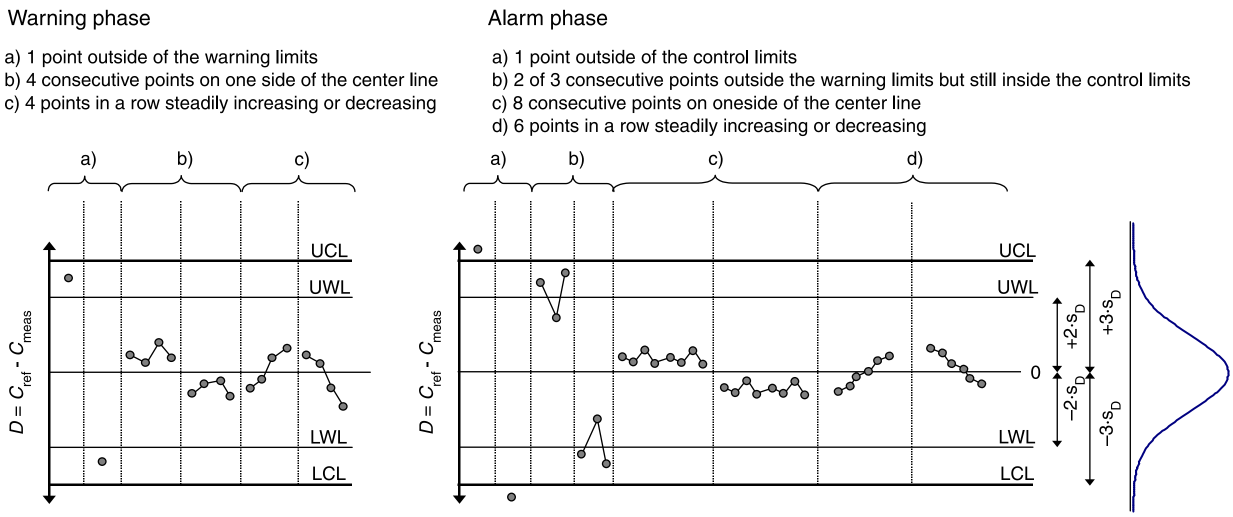

Figure 6 | Shewhart control charts with less stringent out-of-control criteria for the warning phase and normally used criteria (Montgomery 1996) for the alarm phase.

In Table 3 the summarized results of 19 months performance of five on-line sensors in the filter effluent and of two MLSS-in-line-sensors in the aeration tanks are shown. Considering the number of out-of-control events in Table 3 it could be concluded that all the analyzed sensors were in 85–95% of the measuring time not disturbed by any systematic or gross error.

Based on the in-control data pairs the standard deviations of the differences s_D (Table 3) have been recalculated. The in-control-interval has a width of ± 12–17% (NO_3—N), ± 35–40% (PO_4—P), ± 83% ($NH_4^+ - N$) and ± 12–15% (TS) of the measured reference value. The high relative value for NH_4^+—N is due to the small $NH_4^+ - N$-concentration in the filter effluent ($C_{ref,\varnothing} = 0.18\, mg_N/l$). However, the width of the in-control-interval of $\pm 0.15\, mg_N/l$ ($= \pm 3{\cdot}S_D$) is narrow enough to accurately control the effluent quality standard of $2\, mg_N/l$.

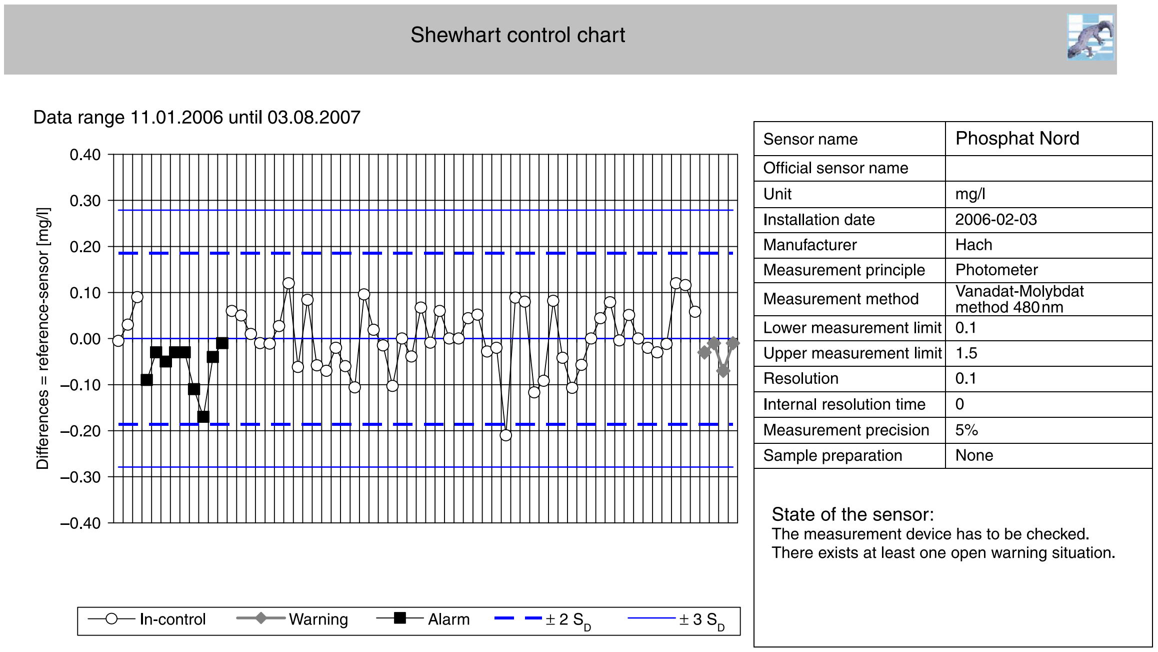

Figure 7 | Example of a Shewhart control-chart for $PO_4^- - P$, effluent filter north, WWTP Zürich-Werdhölzli, out-of-control situations are marked with bold squares.

Table 3 | Results for 7 monitored sensors (WWTP Zürich-Werdhölzli Jan 06–Aug 07)

Sensor location	Measured Sensor type	Parameter	Reference method		$c_{Ref,\varnothing}$	S_D	$3 \cdot S_D$ $c_{Ref,\varnothing}$	*n*	Out-of-control events
Effluent filter north	$NO_3^- - N$	UV-absorption (Nitratax, Dr. Lange)	Dimethylphenol method, Dr. Lange cuvette tests	(mg l^{-1})	6.87	0.39	17%	57	5(9%)
Effluent filter south					7.29	0.30	12%	97	10(10%)
Effluent filter north	$PO_4^- - P$	Vanadat-Molibdat-method (Hach)	Molybdän-blue-method, Dr. Lange cuvette tests	(mg l^{-1})	0.67	0.09	40%	66	9(14%)
Effluent filter south					0.68	0.08	35%	87	5(6%)
Effluent filter north	$NH_4^+ - N$	On-line-NH_4^+ sensor (Staiger Mohilo)	Indophenolblue method, Dr. Lange cuvette tests	(mg l^{-1})	0.18	0.05	83%	97	5(5%)
Aeration tank Bin 13	TS	MLSS-in-line-sensor (Staiger Mohilo)	Gravimetrical analysis with paper filters	(g l^{-1})	3.96	0.16	12%	49	4(8%)
Aeration tank Bin 14					3.91	0.19	15%	33	3(9%)

CONCLUSIONS

- Not identified systematic errors in data sets can cause problems inducing wrong decisions in function control, process modelling and planning of new treatment infrastructure.
- Although existing data sets of municipal WWTP would enable the identification of systematic errors with statistical methods, this analysis is seldom done.
- With the presented statistical mass balance approach five different redundant mass balances are analyzed to identify systematic errors which constitute about 10% to 20% of the amount of the input fluxes, considering a 5%-significance level.
- A two step Shewhart control-chart analysis allows to check the data quality of sensors with a statistical and a fast graphical analysis. The combination of the control chart analysis with a database gives the operator the possibility to survey efficiently all relevant information concerning sensor accuracy and maintenance.
- A 19 month data set of the WWTP Zürich-Werdhölzli indicated that NO_{3-}, PO_{4-} and NH_4-on-line-analyzers in the filter effluent and MLSS-sensors in the aeration tanks were in 85–95% of the measuring time not disturbed by any systematic or gross error. The in-control-interval ($\pm 3\cdot$standard deviation) has a width of ± 12–17% (NO_3—N), ± 27–35% (PO_4—P), ± 83% (NH_4—N) and ± 12–15% (TS) of the measured reference value.

ACKNOWLEDGEMENTS

The author would like to thank G. Koch, M. Ermuth and WWTP Ergolz 1 (AIB Kanton Basel-Landschaft), WWTP Zürich-Werdhölzli, AWEL Zürich, W.Gujer, L. Rieger, S. Burger, J. Herzig and D. Bauer for their generous support.

REFERENCES

Coleman, H. W. & Steele, W. G. 1998 *Experimentation and Uncertainty Analysis for Engineers*, 2nd edition. John Wiley and Sons, New York, ISBN 0-471-12146-0.

GEKO 2005 Gerätekontrolle v1.0, aqua-system ag, CH-8400 Winterthur, www.aquasystem.ch [in German].

ISO 2003 *Water quality - On-line sensors/Analysing equipment for water - Specification and performance tests*, ISO 15839:2003.

Meijer, S. C. F., van Loosdrecht, M. C. M. V. & Heijnen, J. J. 2001 Metabolic modelling of full-scale biological nitrogen and phosphorus removing WWTP's. *Water Res.* **35**(11), 2711–2723.

Montgomery, D. C. 1996 *Introduction to Statistical Quality Control*, 3rd edition. Wiley, New York.

Nowak, O., Franz, A., Svardal, K., Müller, V. & Kühn, V. 1999 Parameter estimation for activated sludge models with the help of mass balances. *Water Sci. Technol.* **39**(4), 113–120.

Thomann, M., Rieger, L., Frommhold, S., Siegrist, H. & Gujer, W. 2002*a* An efficient monitoring concept with control charts for on-line sensors. *Water Sci. Technol.* **46**(4–5), 107–116.

Thomann, M. 2002*b* Datenkontrolle von Abwasserreinigungsanlagen mit Massenbilanzen, Experimenten und statistischen Methoden. Ph.D. thesis ETH Zürich, Schriftenreihe des Instituts für Hydromechanik und Wasserwirtschaft (IHW), http://e-collection.ethbib.ethz.ch/ecol-pool/diss/fulltext/eth14824.pdf [in German].

Van der Heijden, R. T. J. M., Romein, B., Heijnen, J. J., Hellinga, C. & Luyben, K. C. A. M. 1994 Linear constraint relation in biochemical reaction systems: III. Sequential application of data reconciliation for sensitive detection of systematic errors. *Biotechnol. Bioeng.* **44**, 781–791.