A. Spindler and J. Krampe. Quality control of wastewater treatment operational data by continuous mass balancing: dealing with missing measurements and delayed outputs

Quality control of wastewater treatment operational data by continuous mass balancing: Dealing with missing measurements and delayed outputs

A. Spindler, J. Krampe

Institute of Water Quality and Resource Management, Vienna University of Technology, Karlsplatz 13/226-1, 1040 Wien, Austria

E-Mail: *a.spi@iwag.tuwien.ac.at*

Abstract

Continuous mass balancing defines a new standard in data quality validation. Likewise relying on the principles of mass conservation it outperforms long term static mass balancing approaches because faults in data can be assigned to their time of occurrence. This research was carried out with practical application to routine operational data in mind and two major aspects are investigated to make this application feasible. Sludge concentrations of typically balanced components (COD, TN, TP) are not routinely measured in wastewater treatment plants. Therefore they need to be determined from alternative, more frequent measurements such as TSS. To provide the necessary statistical basis for such determination, monthly sludge sampling was found sufficient. Further, contrary to long term static mass balancing, the effects of delay between input and output loads must not be neglected in continuous mass balancing based on daily data. While a storage/release approach did not give the desired results, the consideration of hydraulic retention (first-order flow dynamics) fundamentally improved the performance of the proposed method.

Keywords

continuous mass balancing; data quality control; fault detection; statistical process control

INTRODUCTION

Two fundamental aspects of continuous data quality control by mass balancing of operational data are addressed in this work. One is the determination of the concentration of components of sludge flows by using alternative measurements, the other is the influence of storage and retention on short term balances. The aim is to provide a simple method for practical implementation of continuous data quality control.

Mass balancing is a means of gross error detection in measurement data and the fundamental idea behind data reconciliation. Relying strictly on the laws of mass conservation, mass balances must only be carried out for conservative components that can be measured in all input and output streams of a system. Pure elements are always conservative in wastewater treatment and one typical balanceable element is (total) phosphorus. Nitrogen balances are also possible, however when denitrification is involved, off-gas nitrogen is usually not measured. Another typically balanceable "component" is COD, which is basically a sum parameter for free electrons. Other commonly measured components are not conservative and therefore subject to reactions. Mass balancing based only on measuring the concentrations of such components is not generally possible. An example is TSS, because biomass grows converting dissolved organic material into particulate material. For appropriate subsystems (such as a dewatering unit when TSS is considered) the conservative property of such components might, however, be given. Water itself, expressed as flow Q, can also be balanced neglecting the influence of evaporation.

Common approaches to mass balancing require steady state data (Narasimhan, 2000). For highly dynamic wastewater treatment systems this is usually achieved by considering mean values over rather long time periods (at least two sludge ages, typically several months). In perfect steady state, the total input load of a component into a system is equal the total output load when no accumulation or release occurs. The value of mass balancing as the most important approach to redundant data quality control is widely agreed upon in literature (e.g. Barker and Dold, 1995; Nowak et al., 1999; Puig et al., 2008; Rieger et al., 2010; Villez et al., 2013).

Continuous mass balancing[1], contrary to the static approaches typically used in wastewater treatment, reveals the temporal behavior of the balancing error. It allows to distinguish unbalanced from well-balanced time periods in a data set or to continuously monitor the integrity of operational data. The CUSUM chart, a control chart based on a modified cumulative sum and first introduced by Page (1954), has been proven suitable for continuous balancing of flow data from wastewater treatment plants (WWTPs) by Spindler and Vanrolleghem (2012). In their study the variance of the vector of (daily) balancing errors was found to be an important indicator for good data quality. It also influences the applicable

---

11 The application of CUSUM charts for mass balancing was labeled "dynamic mass balancing" in a previous paper (Spindler and Vanrolleghem, 2012) to differentiate from the established approaches. However, as this approach does not actually target process dynamics, the naming was changed to "continuous mass balancing".

parameters (and therefore sensitivity) of a CUSUM chart. A high variance of the vector of balancing errors requires a higher sensitivity of the CUSUM chart in order to detect off-balance periods, which leads to slower detection and vice versa. See appendix A for a short introduction to CUSUM charts. The present paper investigates the application of CUSUM charts to general mass flow data from wastewater treatment with a focus on requirements regarding the handling of sludge loads.

In practice concentration measurements at WWTPs are usually conducted in flow proportional 24h composite samples. Daily loads are then calculated from the product of this average concentration and the cumulated flow of the respective day. Therefore, in this research continuous mass balancing is applied to daily loads. It follows, that measurements are preferably taken daily, without interruption. This requirement is commonly met for most flows and the concentrations of influent, effluent and reject water but hard to achieve for concentrations in primary sludge (PS), waste activated sludge (WAS) or digested sludge (DS). Measurement of typical balanceable sludge components (TP, TN, COD) is complicated because it requires thorough disintegration of the samples and small but representative sample volumes which are difficult to obtain. Therefore and because these data are circumstantial for daily plant operation, this type of measurement is usually not carried out in practice.

Due to the nature of wastewater treatment, sludge streams are part of virtually every balanceable subsystem of a WWTP. For operation and documentation they are usually characterized by volume and concentration of TSS (total suspended solids). Organic and inorganic constituents of sludge are measured as volatile and nonvolatile suspended solids (VSS and NVSS). TSS, VSS and NVSS are routine parameters and regularly measured on a daily basis. Grab samples are usually sufficient because sludge characteristics change only slowly. Only primary sludge is subject to faster fluctuations but thickened primary sludge can be analyzed instead or online TSS measurement is employed to determine an average value.

A common approach to quantify balanceable sludge components is their determination from TSS or VSS, assuming stable proportionality between the two factors. This is a rational approach, particularly for nitrogen and COD concentrations of WAS and DS, because nitrogen is a constituent and COD a property of the biomass which only the organic fraction of sludge is composed of. Phosphorus, on the other hand can also be chemically precipitated, thus becoming a constituent of the inorganic fraction of WAS and DS. Ekama (2009) includes an overview of literature values on COD and nitrogen concentrations of primary and activated sludge: COD/VSS ratios of activated sludge vary between 1.42 and 1.55, for primary sludge the range is even larger. Nitrogen and phosphorus are also often analyzed to determine nutrient levels for agricultural application. Their concentration in sludge depends heavily on the wastewater composition and treatment and ranges between less than 1% and 10% of TSS (Scharf et al., 1997). The temporal stability of the relations between balanceable sludge components and VSS or TSS within a single sludge is decisive for the reliability of this approach and determines the necessary measurement frequency. Both issues are addressed in this work.

As a second fundamental aspect the influence of delay on short term mass balances is investigated. “Delay” in this work is not meant in its strict meaning referring to flow through an idealized plug-flow reactor. It is rather used to describe the general effect of loads leaving a reactor distributed over a certain time span. For short term mass balances the precondition of steady state, as mentioned above, is not satisfied. Loads entering a reactor on one day do not necessarily leave it on the same day. This can be accounted for by the concept of storage (also: accumulation) and release. These occur when the input load to a balanced subsystem is unequal to the output load in a given time period. For example, the amount of sludge in an activated sludge unit (including clarifiers) depends on the organic influent load and waste sludge flow. When less waste activated sludge is withdrawn from the system, a higher COD, TN and TP load is stored with the sludge.

As it turned out that this storage/release approach is highly sensitive to measurement errors, another concept to account for delayed outputs was investigated. Hydraulic retention (or first-order flow dynamics) can be used to calculate the effluent concentration from a (perfectly mixed) tank, depending on the influent concentration. Here, a constant tank volume was assumed which is typical in wastewater treatment. The assumption of a constant influent flow is derived from the frequency of the measurements the balancing approach is based on (1/d). In continuous balancing based on daily loads the effect of hydraulic retention can be neglected only for streams with very short retention times (less than one day) such as methane or nitrogen gas production. For the effluent with a retention time of roughly one day neglecting this delay is also allowed because it contains only a small proportion of the daily input load and has little influence on the balance.

METHODS

Regression analysis

To investigate the determination of COD, TN and TP from different fractions of suspended solids (SS), three different data sets were used. Data set A contains weekly (at least) routine data from a large Austrian WWTP and covers a time span of almost three and a half years. Data set B stems from a pilot scale anaerobic digestion stationed at another large Austrian WWTP. Sludge concentrations were measured during 47 consecutive weeks. Data set C contains values from sludge samples that were analyzed supplementary to routine operational data in order to achieve balanceability of yet another Austrian WWTP. These samples were taken 21 times over a period of 24 weeks. Plant A and C are subject to strong influence from industries, mainly chemical, accounting for up to 50% of the organic load. Concentrations were measured in the (waste) activated sludge (AS), primary sludge (PS) and digested sludge (DS). On plant B, waste activated and primary sludge are mixed (AS&PS).

Simple and multiple linear regression with and without intercept are applied to determine concentrations of COD, TN and TP from SS. Different SS fractions are considered, namely total (TSS), volatile (VSS) and nonvolatile (NVSS) suspended solids. For consideration of temporal behavior, the inclusion of trend and seasonality is compared to simple linear dependency from SS. The investigated and here reported regression models are of the following types:

$$c_x = a_1 \bullet c_{SS} \qquad \text{eq. (1)}$$

$$c_x = a_1 \bullet c_{SS} + a_2 \qquad \text{eq. (2)}$$

$$c_x = a_1 \bullet c_{SS} + a_2 \bullet \sin(\omega t) + a_3 \bullet \cos(\omega t) + a_4 \bullet t + a_5 \qquad \text{eq. (3)}$$

For evaluation of significance of the regression three different parameters are used: the coefficient of determination (R^2, calculated as explained variance), Akaike's Information criterion (AIC, for balancing model fit and complexity, accounting for the number of model parameters) and the relative two standard deviation range around the mean ($2\sigma_{res}/\mu$, containing about 95% of the measured values).

The large number of data points in data set A also allows for evaluation of lower measurement frequencies by Monte Carlo simulation. This was done by investigating the probability of only slightly deteriorated results (an increase of the relative two standard deviation range of not more than 10%) when determining the regression models from only monthly or quarterly (instead of weekly) measured data.

Continuous balancing under the influence of delay

In the second part of this work some exemplary balances are calculated for plant C based on the adequate determination of sludge concentrations. The balancing error *e* for a chosen subsystem is calculated from the difference between the sum of all input loads and the sum of all output loads (ΣF_{in} and ΣF_{out}). This error can be related to the total input load, giving the relative balancing error e_{rel}. The determination of balancing equations for large and complex plants can be facilitated using an automated approach (Spindler, 2014). For continuous balancing, it is the vector of daily balancing errors that needs to be calculated instead of an overall mean balancing error. This error vector is then analyzed using CUSUM charts (see below). An example is given in the results section.

When wastewater treatment balances are calculated on a daily basis, the delay between input and output loads has to be considered. Two different approaches to account for this delay are investigated, i.e. the concept of storage and release and the concept of hydraulic retention. For better comparison of these different approaches each continuous balance will be calculated three times: one directly (without delay), one including storage and release (based on the SS concentration in the reactor) and one under consideration of hydraulic retention.

Storage (ΔS) is calculated for component loads (TN, TP, COD) contained in sludge (eq. 4).

$$\Delta S_i = V \bullet (x_i - x_{i-1}) \qquad i = 1 \ldots n \qquad \text{eq. (4)}$$

$$\Delta S_i^+ = max(0, \Delta S_i) \qquad \text{eq. (4a)}$$

$$\Delta S_i^- = |min(0, \Delta S_i)| \qquad \text{eq. (4b)}$$

An increasing sludge concentration (storage, ΔS_i^+) is counted as an additional output mass flow; a decrease in sludge concentration (release, ΔS_i^-) is counted as an additional input mass flow (see results). This way, storage and release loads are regarded as physical streams which makes interpretation (e.g. of the magnitude of average storage and release) more intuitive. It also facilitates the automatic determination of balancing equations according to Spindler (2014).

Note that for a correct determination of daily storage, a component's concentration would actually have to be known exactly at the beginning of each 24h composite sampling cycle. This is not always the case in practice. For sludge, for example, grab samples are commonly used and representativeness for the corresponding composite sample has to be assumed.

Because the storage/release approach did not give the desired results (see below), another approach to account for a delayed output load was investigated. The effect of hydraulic retention is taken into consideration by calculating an "expected output mass flow" from the initial concentration of a component (x_0) in the reactor, its influent concentration (x_{in}, assumed constant), the flow rate (Q) and the reactor volume (V). The expected output mass flow can then be balanced against the measured output. Assuming an ideal CSTR the expected output's concentration after a given time (t) is calculated as follows:

$$\frac{d x_{out}}{dt} = Q/V \bullet (x_{in} - x_{out}) \qquad \text{eq. (5)}$$

With τ = V/Q (hydraulic retention of the balanced compound) integration yields

$$x_{\text{out}} = x_{\text{in}} - (x_{\text{in}} - x_0) \bullet \exp\left(\frac{-t}{\tau}\right) \quad \text{eq. (6)}$$

Equation (5) describes the hydraulic transport through an ideal CSTR. Obviously, this is a purely hydraulic model and reactions must not be regarded. Mass balancing is based on the laws of mass conservation (of a component). Reactions only alter the distribution of a component between different output paths, they do not change its total sum.

For the calculation of the daily error vector, the expected mean output concentration for one day ($t=1$, index i) is calculated assuming a constant (mean) influent concentration and flow and a constant volume ($Q_{in}=Q_{out}=Q$):

$$\dot{x}_{\text{out,expected,i}} = \dot{x}_{\text{in,i}} - \tau_{\text{i}} \bullet (\dot{x}_{\text{in,i}} - \dot{x}_{\text{out,expected,i-1}}) \bullet \left(1 - \exp\left(\frac{-1}{\tau_{\text{i}}}\right)\right) \quad \text{eq. (7)}$$

The expected output load is calculated from the expected mean output concentration.

$$F_{\text{out,expected,i}} = \dot{Q}_{\text{out,i}} \bullet \dot{x}_{\text{out,expected,i}} \quad \text{eq. (8)}$$

This expected output load, which is basically calculated from the measured input load (see eq. 7), is then balanced against the measured output load. An example is given in the results.

In case two output paths exist, retention needs to be considered for the slow path only (usually related to the sludge). For example, methane is produced almost instantly from the organic input load in an anaerobic digester. The delay between input and gas production (fast output path) can be neglected when dealing with daily mean data. The digested sludge, however, has a rather long retention time and delay has to be accounted for. This is achieved by calculating a virtual input concentration discounting the fast output load from the actual input load. In this way, equation (5) has to be solved only for one x_{out}, which is the way it was specified.

$$\dot{x}_{\text{in,virtual,i}} = (\dot{x}_{\text{in,i}} \bullet \dot{Q}_{\text{in,i}} - \dot{x}_{\text{out,fast,i}} \bullet \dot{Q}_{\text{out,fast,i}}) / \dot{Q}_{\text{out,slow,i}} \quad \text{eq. (9)}$$

One important question remains: How should the initial concentration in the tank be chosen? It could either be the measured or the previously predicted concentration. In eq. (7), the latter ($x_{out,expected,i-1}$) was chosen. This value has great influence on $x_{out,expected,i}$. In fact, with long hydraulic retention, $x_{out,expected,i}$ depends almost entirely on the initial concentration (It holds: $lim(exp(-x), x \rightarrow 0) = 1-x$). If measured values are used, the expected output concentration $x_{out,expected,i}$ is heavily influenced by the measured output concentration $x_{out,i-1}$. This leads to deterioration of the actual balance (where $x_{out,i}$ is balanced against $x_{out,expected,i}$). Therefore, only the initial value $x_{out,0}$ is taken from measurements, thereafter this value is taken from $x_{out,expected,i-1}$ of the previous day. This way, all $x_{out,expected}$ are (almost) only calculated from the input which is a precondition for balancing against the measured values x_{out}.

CUSUM charts were calculated according to Spindler and Vanrolleghem (2012, see the appendix for an introduction). In this previous work the method was found to reliably detect even small deviations of the balancing error from the expected zero mean in the case of systematic measurement errors. The CUSUM parameters have to be chosen carefully. Once the choice of an average in control run length ARL_0 is made, the control limit h depends only on the reference value k. It was calculated using the spc package (Knoth, 2009) for R (R Core Team, 2013). When the CUSUM chart exceeds the control limit h, it signals a significant deviation from the expected value (0), i.e. an off-balance situation. For ARL_0, the classical value 370 (Montgomery, 2009) was chosen. Small reference values k lead to higher sensitivity (smaller optimally detectable error $\Delta\mu_{opt}$) at the cost of slower detection (increasing ARL). Practice has shown that a good choice of k gives $\Delta\mu_{opt}$ within 10%-20% of the input load. As the variance of the error vector becomes larger, k is chosen smaller (but not below 0.2) to facilitate detection. Error vectors with a small variance are a good indicator of high data quality themselves. In these cases k can be chosen higher to avoid signals at minor disturbances.

RESULTS & DISCUSSION

The first part of this work was concerned with the determination of sludge component concentrations (TP, TN, COD) from frequent alternative measurements, namely fractions of suspended solids (SS). For application of continuous balancing based on CUSUM charts, daily values for these components are required, a precondition usually not met in practice. The determination of sludge components *from* fractions of SS does *not* mean performing balances of SS (in the second part of the results section) which is not generally possible.

Regression analysis for determination of non-measured concentrations

P, N and COD where determined from fractions of SS for three different WWTPs (A,B,C). Data were collected weekly for plants B and C and at least weekly for plant A. Results for the three regression models (eq. 1-3) are given in Table 1. The third regression model also takes into account possible temporal behavior (trend and seasonality) of the variables. The best available model is indicated by "++". In most cases, this is the model including seasonality. If a simpler model reaches comparable significance, this is indicated by "+". Significance is given by the coefficient of determination R^2 and the relative two standard deviation range around the mean. AIC was also calculated but did not give any additional evidence and is therefore not shown in Table 1.

VSS turned out to be the best choice of a SS fraction for the determination of COD. For determination of TN and TP, other fractions give slightly better results in some cases but VSS always remains a good alternative for determination of TN and in most cases for TP, too. Only for the determination of TP in digested sludge (DS) of plant C the volatile fraction alone is not a suitable parameter. In some cases the best results are achieved by assuming VSS and NVSS to be independent, i.e. not constrained by TSS.

Data set A reveals poorer overall regression quality than data sets B and C. It should be kept in mind, however, that this data set covers a time span of almost three and a half years and external influences on sludge characteristics during this period are quite likely. Still, 95% of the residuals lie within ±15% to ±25% of the mean concentration for data set A with the exception of TN and TP values for primary sludge (PS).

Data set B, covering almost one year and analyzed in the laboratory of the authors' home institution, yields coefficients of determination between 0.69 and 0.95. The residuals lie mostly within ±6% to ±13% of the mean concentration. Only for TP determination in mixed sludge (AS&PS) this interval is ±19% of the mean. Data set C, covering only 24 weeks and also analyzed in the authors' home institution, gives similar results. Coefficients of determination lie between 0.60 and 0.96 with one exception (0.43 for TP in PS). The range of residuals is mostly within ±5% to ±9% of the mean concentrations. Again, exceptions occur only for determination of TN and TP in PS.

The determination of COD gives mostly acceptable results (residuals range ±25% or lower), with simple linear regression models being sufficient. In two cases the intercept must not be neglected. Only for the activated sludge (AS) of plant C the temporal behavior requires consideration, too. For determination of TN and TP from data sets B and C acceptable results are achieved in AS and DS. For plant A, the poorer quality of regression models is attributed to the higher number of data as stated above. For the PS however, meaningful regression seems harder to achieve, especially for TP but also for TN.

Table 1. Results from the regression analysis for determination of COD, TN and TP from SS fractions. **"++"** best result (along with R^2 and $2\sigma_{res}/\mu$); **"+"** close to best results but less

parameters; **"(…)"** alternative SS fraction for similar accuracy; AIC not shown; AS…activated sludge, PS…primary sludge, DS…digested sludge.

variable	sludge	Plant	n	$a_1 \cdot c_{SS}$	$a_1 \cdot c_{SS}+a_2$	$a_1 \cdot c_{SS}+a_2 \cdot \sin(\omega t)+a_3 \cdot \cos(\omega t)+a_4 \cdot t+a_5$	suitable SS fraction	R^2	$2\sigma_{res}/\mu$
COD	AS	A	175	+	++		VSS	0,59	17%
COD	AS&PS	B	47	+		++	VSS	0,95	8%
COD	AS	C	21			++	VSS	0,6	7%
COD	PS	A	188	++			VSS	0,82	25%
COD	PS	C	21	+	++		VSS	0,96	5%
COD	DS	A	367		+	++	VSS	0,43	17%
COD	DS	B	47	++			VSS	0,69	13%
COD	DS	C	21	++			VSS	0,94	5%
TN	AS	A	177			++	VSS	0,47	23%
TN	AS&PS	B	47	+		++	TSS (VSS)	0,89	11%
TN	AS	C	21			++	VSS (TSS)	0,67	9%
TN	PS	A	185	++			VSS	0,67	35%
TN	PS	C	21	+		++	VSS (& NVSS)	0,6	31%
TN	DS	A	365			++	VSS	0,52	16%
TN	DS	B	47			++	TSS (VSS)	0,87	6%
TN	DS	C	21		+	++	VSS (& NVSS)	0,72	8%
TP	AS	A	177			++	VSS	0,34	23%
TP	AS&PS	B	47	+		++	TSS (VSS)	0,69	19%
TP	AS	C	21	+		++	TSS (VSS&NVSS)	0,87	6%
TP	PS	A	189			++	VSS	0,49	53%
TP	PS	C	21			++	VSS (& NVSS, TSS)	0,43	41%
TP	DS	A	369		+	++	VSS	0,53	15%
TP	DS	B	47		+	++	NVSS (TSS,VSS)	0,83	7%
TP	DS	C	21		+	++	NVSS (& VSS, TSS)	0,95	5%

It is important to notice that this assessment is purely statistical. Therefore, extrapolation of results into different ranges of SS concentrations (e.g. from AS to thickened AS) or time periods is not reliable. The regression can be applied to determine concentrations of less frequently measured sludge components from more frequently (preferably every day) measured fractions of SS. An obvious deterministic relation exists only for direct proportionality between COD (as well as TN) concentrations in sludge and VSS. But although such a relation seems reasonable for these sludge components, counterexamples (mainly for TN) are found in the results.

Some regressions are obviously less reliable. This regards mainly TP and TN in PS. The reason for this remains not totally clear. It probably has to deal mainly with the high variability of primary sludge composition. The third example of the following results section (Continuous balancing) could be an indication that continuous balancing might not be as successful when component concentrations in sludges are not reliably determined.

The required minimum measurement frequency for sludge components (along with fractions of SS) was analyzed by Monte Carlo simulation (MCS). It reveals that for data set A similar regression results as in Table 1 can be achieved when the regression is based on monthly data instead of weekly measurements. The probability for the residuals' two standard deviation range to increase by more than 10% above its original value is below 3% in all cases (data not shown). MCS was based on the best available model for each sludge and concentration, in most cases including seasonality. When only quarterly data is simulated, these results cannot be reproduced. Only data that is not influenced by seasonality can be reliably determined from measurements at this low frequency.

Continuous balancing

Following the determination of sludge components from daily measured SS fractions, three different continuous balances were calculated for plant C. Those are the NVSS and COD balances of the anaerobic digester and the total phosphorus balance of the combination of primary clarifier and activated sludge tank (including secondary clarifier). Performing and NVSS balance for the anaerobic digester is in line with the requirement of conservative components as precipitation is negligible. Each balance was calculated three times:

(I) Without consideration of storage and retention
(II) With storage based on daily SS-fluctuations
(III) With hydraulic retention

A calculation example is given for the COD balance of the digester (for data see appendix B):

Daily input loads (calculated from flow and concentration):

$$\sum F_{\mathrm{in,i}} = F^{\mathrm{COD}}_{\mathrm{Co,i}} + F^{\mathrm{COD}}_{\mathrm{PS,i}} + F^{\mathrm{COD}}_{\mathrm{WAS,i}}$$

Daily output loads:

$$\sum F_{out,i} = F^{COD}_{DS,i} + F^{COD}_{gas,i}$$

The error vector without consideration of storage and retention follows as:

(I) $$e_{\mathrm{rel,i}} = \left(\sum F_{\mathrm{in,i}} - \sum F_{\mathrm{out,i}}\right) / \acute{F}_{\mathrm{in}}$$

Storage and release are easily integrated into (I) as additional loads:

(II) $$e_{\mathrm{rel,i}} = \left(\sum F_{\mathrm{in,i}} + \Delta S^{-}_{\mathrm{i}} - \sum F_{\mathrm{out,i}} - \Delta S^{+}_{\mathrm{i}}\right) / \acute{F}_{\mathrm{in}}$$

For consideration of hydraulic retention, the two output paths have to be considered separately. Methane is produced from input COD practically without delay (fast output path). Hydraulic retention occurs for the digested sludge (slow output path). The virtual input concentration is therefore calculated from the difference between input load and the fast output load:

$$x_{in,virtual,i} = (\sum F_{in,i} - F^{COD}_{gas,i}) / Q_{DS,i}$$

The expected output load results from the virtual input concentration (the digester volume for calculation of τ_i is 8000 m³):

$$F_{out,expected,i} = [x_{in,virtual,i} - \tau_i \bullet (\acute{x}_{in,virtual,i} - \acute{x}_{out,i-1}) \bullet (1 - \exp - \frac{1}{\tau_i})] \bullet Q_{DS,i}$$

Finally, the error vector under consideration of hydraulic retention is:

(III) $e_{rel,i} = (F_{out,expected,i} - F^{COD}_{DS,i}) / \acute{F}_{out,expected}$

Results are given in figures 1-3. The figures include the relative error vector (dark points left side) and the relative input and output loads (grey lines left side). On the right side, the CUSUM charts are depicted; reference value *k* and control limit *h* along with the optimally detectable error ($\Delta\mu_{opt}$) and the average run length ($ARL_{\Delta\mu}$) are given. The CUSUM chart signals (dots turning from grey to black) when the control limit is exceeded either on the positive or on the negative side.

The first example is the NVSS balance of the anaerobic digester. The hydraulic retention time is very high at 47 days. The relative standard deviation of the error vector in case (I) is 0.34 and the CUSUM chart signals two off-balance periods, once between days 50-60 and then an almost constant systematic error (linear slope) starting after day 90. The consideration of storage, case (II), leads to a much higher relative standard deviation of 1.24. The average storage load is around ±700 kg/d, more than 1/3 of the influent and effluent load. Because of the high standard deviation of the error vector the CUSUM parameters were chosen for maximum sensitivity. Still, the optimally detectable error is very high at 47% of the mean influent load and the average run length (*ARL*) for this error is at 42 days. The CUSUM chart does not signal in case (II). In case (III), considering retention of the input NVSS load leads to a very low relative standard deviation of only 0.06. Accordingly, CUSUM parameters can be chosen less sensitive which results in an optimally detectable error of 8.5% and an *ARL* of only 5 days. The CUSUM chart shows a long period of stability until day 120 after which the system goes out of balance, in the same way as in case (I). Because of the low standard deviation of the error vector, this is even visible, though not as clearly, from the balancing error plot itself.

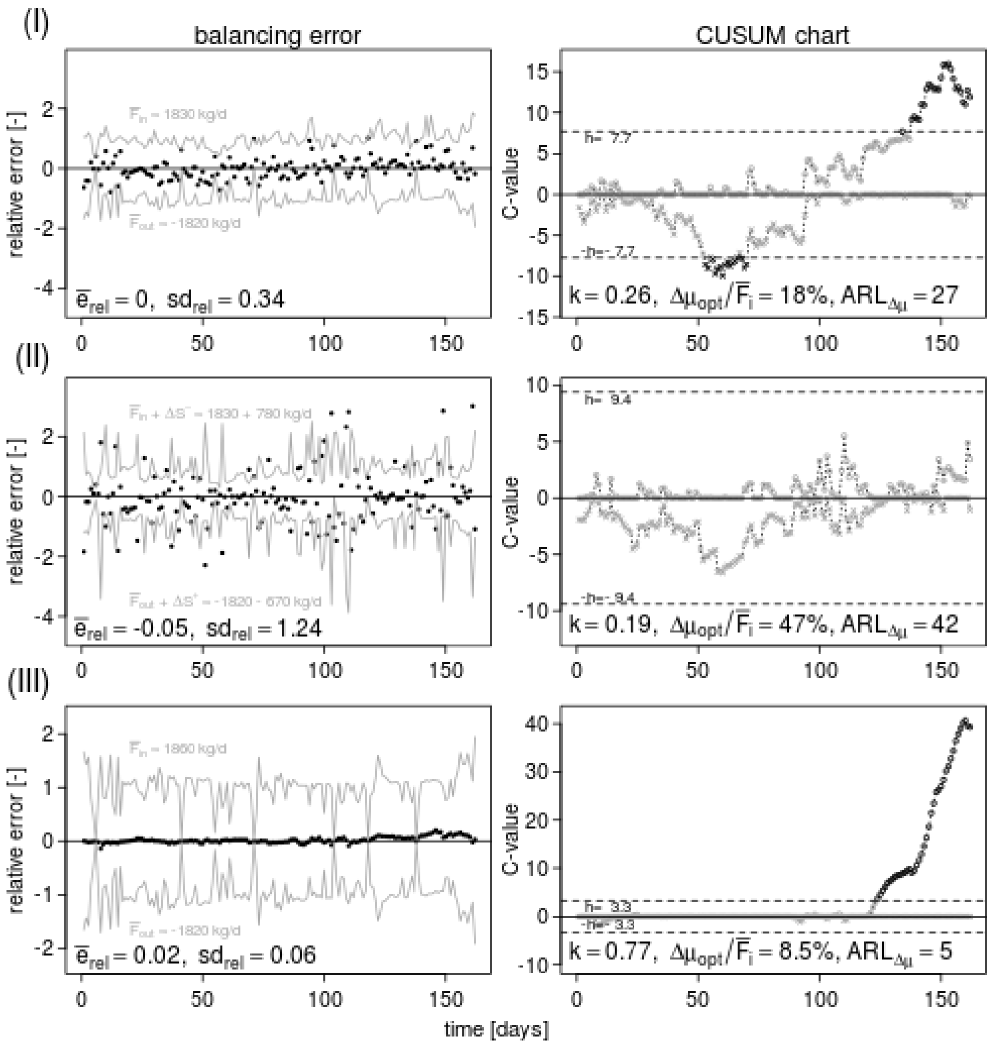

Figure 1. Error vector (left) and two-sided CUSUM chart (right) for the anaerobic digester NVSS balance. (I) without consideration of storage and retention, (II) with storage based on daily SS-fluctuations, (III) with hydraulic retention. Along with the error vector (left, black dots) the total input and output loads are given as grey lines (normalized to mean 1). The CUSUM charts (right) signal an off-balance situation (indicated by color changing from grey to black), when the upper or lower graphs exceed their control limit *h*.

The second example is again for the anaerobic digester, this time considering COD which has two output streams (methane gas and sludge) contrary to NVSS in the first example (only sludge). In cases (I) and (II) (the balance without consideration of delay and the balance considering storage), do not give a (clear) signal. The system seems well balanced. Again, the relative standard deviation of the error vector is higher in case (II) than in case (I). However, when retention is taken into account (III), the analysis changes. The relative standard deviation of the error vector drops again to a low value (0.10) allowing for reliable detection of even small errors. The CUSUM chart signals a constant error starting from around day 70. When calculated only for the first 70 days, the mean balance error is 0.2% (not shown in figure). For days 70 to 162 it jumps to 16% (not shown), indicating a systematic error in (at least) one of the input or output loads. It was verified in a separate balance (not shown) that this error is not in the flow. Anyway, a flow error would influence both the NVSS balance and the COD balance in the same direction, which is not the case. With COD in sludges (PS, WAS, DS) being calculated from VSS, the error could lie in TSS measurement, however, the two charts (NVSS and COD) start signaling at different times, indicating (an)other source(es) of error. For the COD balance, this could well be in the COD concentration of the co-substrate as this value was interpolated from very few measurements.

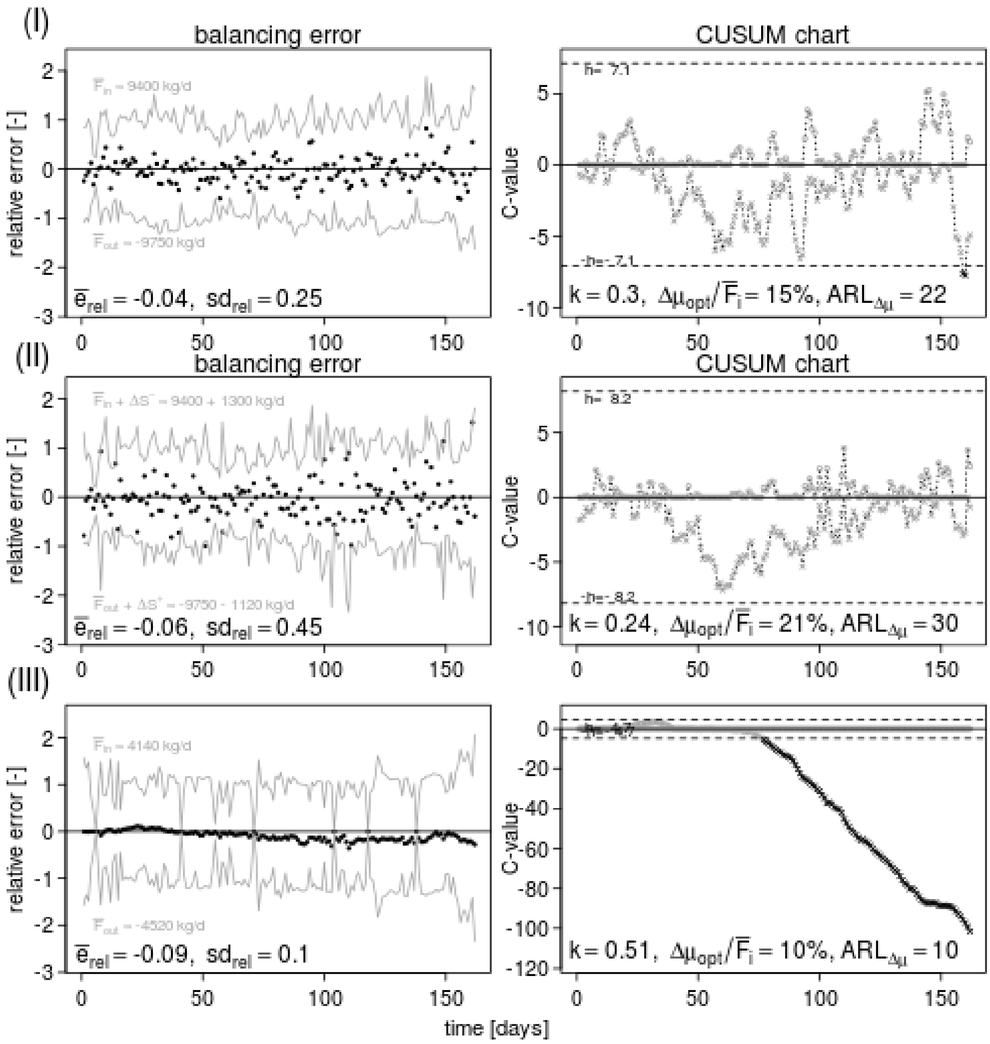

Figure 2. Error vector (left) and two-sided CUSUM chart (right) for the anaerobic digester COD balance. (I) without consideration of storage and retention, (II) with storage based on daily SS-fluctuations, (III) with hydraulic retention. See figure 1 for a detailed explanation.

The third example is the phosphorus balance around the combination of the primary clarifier and the aeration tank (including secondary clarifiers). Just like the second example it was based on a regression model for the determination of sludge loads. In example three, however, there is one component (TP in PS) for which the regression model did not fit the data very well. Due to the low load (around 50% of design capacity) sludge retention time (SRT) is long at this stage (33 days). The SRT determines the hydraulic retention of the slow output path (waste activated sludge). The primary sludge and the effluent together thus constitute the fast output paths with hydraulic retention of around one day. The relative standard deviation of the error vector is again higher in case (II) than in case (I) and does not allow for enough sensitivity of the CUSUM chart to detect off-balance periods. Considering retention (case III), the standard deviation improves slightly compared to the direct balance but remains higher than in the previous two examples. This may be connected to the lower quality of the regression model for TP in PS. The CUSUM chart leads to a very different interpretation. While the most stable time period in case (I) is between days 30-85, this changes to days 85-130 when retention is accounted for. Both charts give a second signal on the negative side following a sudden drop after day 130.

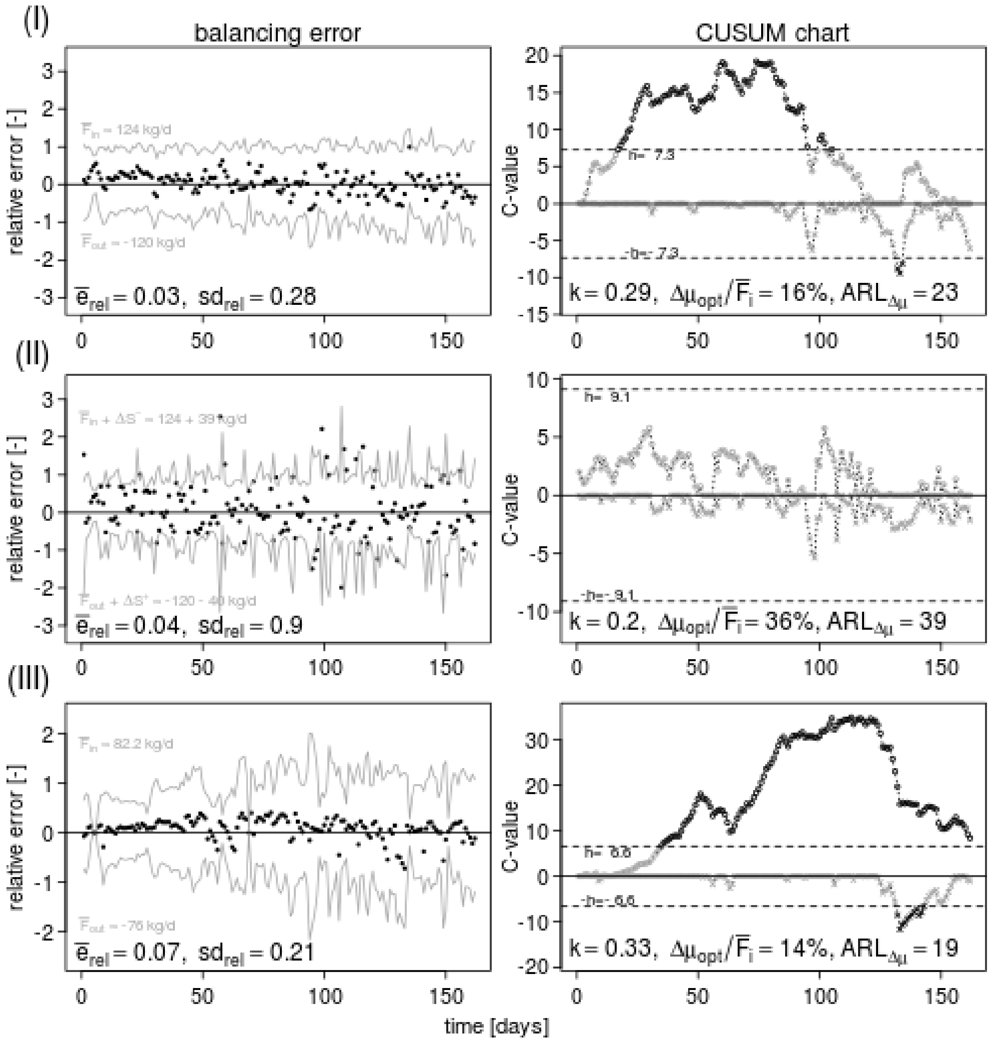

Figure 3. Error vector (left) and two-sided CUSUM chart (right) for the PC/AST TP balance. (I) without consideration of storage and retention, (II) with storage based on daily SS-fluctuations, (III) with hydraulic retention. See figure 1 for a detailed explanation.

The results emphasize that flow dynamics must not be neglected in continuous balances. Under consideration of retention, the variability of the error vector is smaller than without.

Small error vector variability indicates similar trends of input and output loads, a sign of little noise in data. This leads to much higher sensitivity of the CUSUM chart and strengthens confidence in its (off-balance) signals. Hydraulic retention can be calculated sufficiently under assumption of an ideal CSTR and based on daily flow values. In cases where the hydraulic flow through reactors is better described by a plug flow, the methodology can be adopted accordingly. The calculation of storage from daily fluctuations in SS concentrations appears to be not feasible as it leads to an increased variance of the error vector. There are some reasons that might explain this observation. First, the method relies on daily SS concentration measurements which are not very accurate with random errors of around ±10% to be assumed. This has a great effect especially for reactors with long HRT as the stored mass is much larger than daily input and output mass flows. Secondly, storage and release are calculated from differentials (actual and previous day), the integration of which is known to amplify noise. Filtering might reduce this effect but could also lead to deletion of information contained in data. As a third aspect, SS concentrations should actually be known at a fixed time corresponding to 24h composite sampling to accurately calculate the stored amount of sludge but in practice only grab samples are available. A simple simulation study (results not shown) revealed a considerable influence of using the correct sampling time for the calculation of the stored sludge amounts (which is another source of error). For activated sludge systems, measurement of SS concentrations is also subject to large random errors as sludge can be temporarily stored in the clarifiers. All these influences increase the random error of the calculated storage and therefore lead to larger balancing error variability. The hydraulic retention approach on the other hand, depends on a measurement only for the generation of a starting value and after that determines the effect of delay from the retention model. The choice of the starting value for the concentration in the balanced reactor is of relatively little influence. In case of a systematic measurement error for this measurement (which is also the measurement for the slow response output path) a signal of the CUSUM chart will soon occur. If only the starting value was chosen wrong and the following values are free of systematic errors, the CUSUM chart might signal initially but would soon turn back towards zero.

This work, as it is presented here, omits to a large extend its connection to data reconciliation as known and widely applied in process engineering. Some readers might draw the conclusion that these results might have been reached more efficiently by direct application of existing methods for dynamic, nonlinear data reconciliation. There are a number of reasons for this omission. First, wastewater treatment is very different from the majority of process engineering applications in the way that the influent to the system is the main disturbance rather than a controlled variable. Secondly, in data reconciliation (as the name implies) the correction of measurements is the main focus, with gross error detection as a prerequisite or a byproduct. In practical wastewater treatment applications it is, however, sufficient to become aware of faults in data, possibly along with a conclusion as to which measurement is corrupted. The CUSUM chart offers a very descriptive and easily implementable way to enable operators to draw their own conclusions about the state of their measurements. And as a third aspect, the methods of data reconciliation have not yet been proven to be applicable to operational data from wastewater treatment. With delight the authors would see a process engineer taking on the challenge to improve gross error detection in wastewater treatment data. For this reason, the data used in the second example is included in the appendix.

CONCLUSIONS

Continuous mass balancing requires the consideration of the temporal delay between input and output mass flows to correctly determine the quality of operational data. Neglecting this delay is likely to yield erroneous interpretations. While the calculation of storage and release (calculated from fluctuations in SS concentrations) does not seem feasible as it leads to an increased variability of the error vector, hydraulic retention does adequately account for this effect. For the future it would be desirable to investigate further into the correctness of off-balance signals given by CUSUM charts. Because this is often complicated with real data, the application of the Benchmark Simulation model might be appropriate for this task.

The determination of COD, TN and TP from SS fractions is possible in most cases. Purely statistical analysis, in most cases also considering time dependency, yields the best results. Therefore, special care has to be taken when these models are applied; extrapolation beyond the underlying range of time and SS concentrations is not advisable. For long term data, multiple determination is likely to be more appropriate than determination of one single parameter set. Further investigation into this question might be useful. It was found that monthly grab samples are sufficient for the determination of sludge concentrations of COD, TN and TP along with TSS and VSS.

Through this study, the practical applicability of continuous mass balancing has been proven. For a successful outcome of any data evaluation effort including mass balancing, WWTP operators need to be encouraged to ensure balanceability of their measured operational data. This is best achieved by practically calculating those balances that contain the most important measurements but can also be facilitated by redundancy evaluation. In most cases, additional external measurements of sludge components and the corresponding, more frequent, on-site TSS and VSS measurements will be required.

Continuous mass balancing, mastering the insufficiencies of static balances, has the potential to become a standard for data quality verification not only in practice but also in future pilot or technical scale scientific research within the field of wastewater treatment.

REFERENCES

Barker, P. S. and Dold, P. L. (1995) COD and nitrogen mass balances in activated-sludge systems. *Water Research*, 29(2), 633–643.

Ekama, G. A. (2009) Using bioprocess stoichiometry to build a plant-wide mass balance based steady-state WWTP model. *Water Research*, 43, 2101–2120.

Knoth, S. (2009). *spc: Statistical Process Control.* R package version 0.3. http://CRAN.R-project.org/package=spc

Montgomery, D. (2009) *Introduction to Statistical Quality Control*, Wiley, Hoboken N.J.

Narasimhan, S. and Jordache, C. (2000) *Data Reconciliation & Gross Error Detection: An Intelligent Use of Process Data*, Gulf Professional Publishing, Houston.

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

Page, E. S. (1954) Continuous inspection schemes. *Biometrika*, 41(1-2), 100-115.

Puig, S., van Loosdrecht, M. C. M., Colprim, J., and Meijer, S. C. . (2008) Data evaluation of full-scale wastewater treatment plants by mass balance. *Water Research*, 42, 4645–4655.

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. Version 2.15.3, Vienna, Austria, R Foundation for Statistical Computing. [online] http://www.R-project.org/

Rieger, L., Takács, I., Villez, K., Siegrist, H., Lessard, P., Vanrolleghem, P. A. and Comeau, Y. (2010) Data reconciliation for wastewater treatment plant simulation studies—Planning for high-quality data and typical sources of errors. *Water Environment Research*, 82(5), 426–433.

Scharf, S., Schneider, M. and Zethner, G. (1997) Zur Situation der Verwertung und Entsorgung des kommunalen Klärschlammes in Österreich. *Monographien Band 095.* Bundesministerium für Umwelt, Jugend und Familie, Wien.

Spindler, A. and Vanrolleghem, P. A. (2012) Dynamic mass balancing for wastewater treatment data quality control using CUSUM charts. *Water Science and Technology*, 65(12), 2148–2153.

Spindler, A. (2014) Structural redundancy of data from wastewater treatment systems. Determination of individual balance equations. *Water Research*, 57, 193–201.

Villez, K., Corominas, L. and Vanrolleghem, P. A. (2013). Sensor fault detection and diagnosis based on bilinear mass balances in wastewater treatment systems. *Proceedings of the 11th IWA Conference on Instrumentation Control and Automation (ICA2013)*, Narbonne, Sept. 18-20, 2013, Appeared on USB-stick (IWA-12744).

APPENDIX A – The CUSUM chart (from Spindler and Vanrolleghem, 2012)

CUSUM charts are used widely in statistical process control to detect small changes (e.g. shifts or drifts) in the mean μ (the target value) of a monitored process variable. Small in this context means changes of less than one standard deviation.

CUSUM charts are designed to detect one-sided changes (increase or decrease) of the monitored variable X. For the two-sided case (increase and decrease), one upper (positive) and one lower (negative) CUSUM chart have to be combined. For convenience, data is normalized to zero mean and standard deviation one. The CUSUM is a modified cumulative sum of a process variable X, consecutively adding up the values x_t, $t=1,\ldots,n$ where n is the length of vector X. The two modifications are:

i. The upper (positive) CUSUM may not drop below zero, the lower (negative) CUSUM may not rise above zero.

ii. A smoothing parameter (reference value k) restricts the sensitivity of the method by constantly drawing the CUSUM series towards the target value (zero for normalized data).

The two-sided CUSUM for normalized data may be defined as:

$$C_t^+ = max\left(0, C_{t-1}^+ - k + x_t\right)$$
$$C_t^- = min\left(0, C_{t-1}^- + k + x_t\right) \quad \text{with} \quad C_0 = 0 \qquad (3)$$

The CUSUM series signals an undesired shift $\Delta\mu$ of the process mean by exceeding a chosen control limit ($+h$ or $-h$). Thus, the reference value k and the control limit h are the two parameters which determine the behavior of the CUSUM chart. The optimal value of k is $\Delta\mu/2$, half the size of the shift to detect. The control limit h may then be chosen according to the desired average run length ARL_0 of the CUSUM series.

The average run length ARL_0 is the average number of time steps (i.e. data points) after which the CUSUM series will give a signal even though the true shift of the mean is zero (false alarm). Indeed, due to the probabilistic nature of the data (random errors), a long enough CUSUM series will eventually exceed any control limit. This corresponds to the type I error (false positive) in statistical tests. Therefore, a compromise has to be made. In the past, ARL_0

was chosen as 370 which is equivalent to a 3σ control limit on a Shewart control chart.

When k and h have been chosen, the average run length $ARL_{\Delta\mu}$ (for detection of a true shift $\Delta\mu$ of the mean) can be calculated. $ARL_{\Delta\mu}$ increases with decreasing values of k (when h is adjusted to keep a constant ARL_0) and therefore with smaller shifts $\Delta\mu$. In statistical process control a fast response, i.e. low $ARL_{\Delta\mu}$ is desirable.

APPENDIX B – Sludge loads for the COD balance of the anaerobic digester (kg/d)

day (i)	F^{COD}_{Co}	F^{COD}_{PS}	F^{COD}_{WAS}	F^{COD}_{DS}	F^{COD}_{gas}	ΔS^{COD}	Q_{DS} [m³/d]	day (i)	F^{COD}_{Co}	F^{COD}_{PS}	F^{COD}_{WAS}	F^{COD}_{DS}	F^{COD}_{gas}	ΔS^{COD}	Q_{DS} [m³/d]
1	435	5274	2258	6509	3699	6091	249	82	1930	2963	3563	5855	4341	698	226
2	735	3680	3701	5582	3719	-483	214	83	1939	2229	3217	4849	4439	698	187
3	1035	4236	4360	6142	3940	-483	236	84	2134	3607	3628	3690	4432	698	142
4	1035	5038	1940	3138	3704	-483	121	85	2746	2827	3241	5421	5099	698	207
5	435	2183	23	1838	3704	-483	71	86	2952	3875	4831	4624	5245	5593	184
6	435	1670	18	0	3348	-483	0	87	2823	2979	4649	5228	6305	210	208
7	1215	4043	1554	3432	3320	-483	133	88	1645	4038	3374	5092	7697	210	202
8	3007	4970	2998	6466	4151	-9606	239	89	1031	1813	3081	4979	6202	-5140	190
9	3028	3592	3584	3328	4646	3096	125	90	1043	2546	3604	5215	5459	-5123	191
10	1391	5883	4100	3092	4113	6133	120	91	1733	2213	4292	5239	5353	227	192
11	735	4128	3385	4098	3850	54	159	92	1731	2288	4662	5249	5590	227	192
12	735	2642	3099	6219	3915	54	242	93	1902	2404	5970	5060	5291	5543	192
13	585	3063	2602	2670	3853	54	104	94	848	3015	8949	3257	4471	465	123
14	735	3625	2591	1952	3680	-6016	73	95	602	4860	8731	4898	3937	5746	192
15	2334	3871	3116	6058	4056	6124	235	96	1134	2662	4985	4622	4071	-4815	174
16	2921	3871	3416	1840	4226	376	71	97	1283	2421	3845	5105	4131	465	192
17	3234	3713	2800	4309	4157	1383	168	98	1859	2933	5163	4919	5023	5692	192
18	3808	3868	2433	4001	4476	1378	156	99	2481	1825	2338	5226	5552	-7366	192
19	4134	2670	2507	4236	4396	1374	166	100	3102	3243	3125	2708	5937	-7384	94
20	3409	3394	2569	4014	4278	1369	158	101	2223	7051	3395	5344	6553	5608	192
21	4185	3962	2631	4256	4516	1365	168	102	1795	2845	5152	5255	6344	2996	192
22	4125	4760	2527	4426	5541	1360	176	103	1816	3900	3785	5371	6079	-12462	179
23	3482	3551	2137	3617	6743	6323	149	104	1838	1804	3373	0	6123	6860	0
24	2793	4411	1915	3339	6416	58	138	105	1859	3881	2597	3633	5948	6822	132
25	2741	4420	2004	3351	6300	58	138	106	2481	3004	5335	5248	6092	2957	194
26	2690	2978	2100	2951	6064	-5896	117	107	2802	6253	2700	3402	5994	-393	126
27	2638	3817	2178	3682	6252	60	146	108	2073	5497	2574	3521	6710	4739	136
28	2886	5875	2676	4692	6403	60	186	109	1945	3471	3713	5359	6105	-10671	192
29	2985	4250	2769	4237	6571	3034	171	110	1966	2256	3786	5834	5857	-13313	191
30	2933	8074	3356	4816	6562	-2838	191	111	1988	2137	3358	4760	5876	7320	165
31	2731	4793	3714	3677	6568	136	145	112	2009	4648	3864	5224	5672	4769	188
32	2680	3201	3789	5401	7016	136	214	113	2781	4991	3727	4011	5539	-3011	142
33	2778	2601	3774	4853	7111	136	192	114	3402	2424	5084	3609	5549	-415	128
34	2577	4054	3751	2860	7047	136	113	115	3269	3433	3206	4325	5551	2913	156
35	2375	4604	4112	3415	6259	-3414	132	116	2959	2127	3688	5231	5667	2902	192
36	2323	4384	2967	5042	6910	-2226	192	117	1926	1407	4321	4611	6020	296	169
37	2422	3435	3128	4690	6993	6048	185	118	1944	1848	5087	15	5904	296	1
38	2070	5998	3539	4990	6918	-2750	192	119	1775	3994	6651	4151	5463	296	152
39	2018	4036	2948	4347	7425	1379	167	120	1714	3826	3982	4321	4596	2874	161
40	1967	3269	2341	4363	7180	3702	170	121	1765	4186	5865	5565	4206	393	207
41	1915	2692	2270	0	6467	777	0	122	1613	3512	8298	7538	4978	393	279
42	1863	3765	3356	3603	5785	777	139	123	1912	5012	7704	6769	5907	393	250
43	3012	5388	3720	4982	6061	777	192	124	1611	2280	6059	5755	6346	393	213
44	2960	4386	3573	4974	5907	777	191	125	1609	2667	4631	5707	6149	393	210
45	2773	3943	3792	4892	5755	2955	191	126	1757	3641	4861	5498	5749	393	202
46	3036	5139	3989	4347	5904	-2747	167	127	1906	4386	4716	5222	5742	393	192
47	2400	3131	3967	3517	5942	2951	137	128	2955	4465	5117	5235	5966	393	192
48	2063	3469	1539	4958	5816	-1605	191	129	3103	4452	5597	5453	6015	-4361	192
49	2026	2922	3743	5091	5606	-3878	191	130	2801	3597	4986	5470	5906	673	192
50	2289	4915	4000	4953	5176	105	186	131	1900	3421	5067	5171	5855	673	181
51	2102	2272	3068	4803	4658	8624	191	132	1899	2406	3550	3728	5958	673	130
52	2666	2797	2563	4891	4854	3182	198	133	1560	3415	5307	5719	5605	-4237	192
53	2629	2749	2183	4959	5180	-2484	196	134	1559	5555	6997	5739	5629	698	192
54	2742	1958	2133	3440	5042	349	136	135	1221	2786	1921	4647	5530	698	155
55	1297	2427	1952	695	4895	349	27	136	1034	3395	2820	3614	5072	4912	125
56	1744	3886	539	3126	4841	349	123	137	1297	2396	4801	5249	5090	-4804	175
57	1085	1995	1186	4921	4859	-2450	190	138	997	2698	4346	6	5328	2481	0
58	2121	3135	2640	2634	4262	8523	108	139	846	2560	4090	3632	5258	2479	126
59	1934	4082	2494	3288	4588	126	134	140	1410	5082	3199	5501	4730	52	190
60	524	2512	2187	4699	3956	126	192	141	2195	4983	3562	5516	4462	52	191
61	374	3293	2279	1398	4041	126	57	142	1895	12083	3745	5679	4241	52	196
62	524	3665	2371	3012	3917	126	123	143	1444	4868	3777	5428	4738	2305	191
63	795	3622	2920	4706	4100	126	192	144	843	10727	4829	5452	4532	-119	192
64	2200	5476	3552	4446	4570	126	181	145	843	5691	4914	5375	5031	2311	193
65	1494	4096	4266	4448	4535	126	181	146	991	2746	4683	5110	5077	4748	191
66	1325	4674	4853	4759	4640	48	194	147	840	3575	3682	5324	4962	-4981	191
67	374	5990	3365	4240	4303	48	172	148	1589	5412	2974	5091	4661	2320	187
68	374	3219	2635	4645	4282	48	189	149	1889	3752	3999	5114	4731	-12371	171
69	374	2409	2437	4891	4525	-5502	191	150	1289	5721	4536	3746	4516	4699	130
70	652	3707	2396	3660	4570	50	143	151	1888	6746	3703	3736	4429	4712	135
71	994	3150	3453	8	4512	50	0	152	1320	5010	3136	5468	4676	-4095	191
72	851	3303	2469	2320	4354	50	91	153	869	3504	3087	4048	4555	-1161	141
73	1308	4749	2937	6980	4447	2772	278	154	1018	3542	2994	5537	4530	-181	193
74	865	2076	3007	4722	4673	2772	192	155	1467	2812	3295	6337	6678	-181	221
75	971	6671	1852	4203	4746	0	171	156	1317	4920	3733	8073	7569	-181	281
76	980	2582	2976	4959	5216	0	201	157	866	7126	4142	7005	6125	-906	245
77	1193	2900	3217	4458	4275	-5545	174	158	1015	5176	4108	6993	6213	-906	246
78	1354	4494	2352	4328	3734	0	169	159	865	5063	4027	8024	6177	-906	283
79	1065	3989	4134	3203	3505	0	125	160	1014	6002	4227	6895	5345	-906	244
80	3520	4642	4274	5587	3657	698	217	161	1013	10235	4775	5843	4969	-11096	193
81	2096	4511	4632	5168	4414	698	200	162	863	9189	5210	9691	5531	4155	334