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

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 (chemical oxygen demand, total nitrogen, total phosphate) are not routinely measured in wastewater treatment plants. Therefore they need to be determined from alternative, more frequent measurements such as total suspended solids. 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.

Key words | 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 chemical oxygen demand (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 total suspended solids (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 & Jordache 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 to 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 the literature (e.g. Barker & Dold 1995; Nowak et al. 1999; Puig et al. 2008; Rieger et al. 2010; Villez et al. 2013).

Continuous mass balancing, contrary to the static approaches typically used in wastewater treatment, reveals the temporal behavior of the balancing error. The application of CUSUM charts for mass balancing was labeled ‘dynamic mass balancing’ in a previous paper (Spindler & 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’. 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 & 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 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 (available online at http://www.iwaponline.com/wqrjc/050/056.pdf) 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 24 h 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 most commonly met for most flows and the concentrations of influent, effluent and reject water, but is hard to achieve for concentrations in primary sludge (PS), waste activated sludge (WAS) or digested sludge (DS). Measurement of typical balanceable sludge components (total phosphate (TP), total nitrogen (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.

Owing 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. Organic and inorganic constituents of sludge are measured as volatile and non-volatile 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 PS is subject to faster fluctuations but thickened PS 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 of which only the organic fraction of sludge is composed. 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 (AS): COD/VSS ratios of AS vary between 1.42 and 1.55, for PS 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 1 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 AS unit (including clarifiers) depends on the organic influent load and waste sludge flow. When less WAS 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 on which the balancing approach is based (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 1 day) such as methane or nitrogen gas production. For the effluent with a retention time of roughly 1 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 3 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. Plants 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) AS, PS and DS. On plant B, waste AS and PS are mixed.

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, VSS and 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:

\[
c_x = a_1 \cdot c_{SS} \quad (1)
\]

\[
c_x = a_1 \cdot c_{SS} + a_2 \quad (2)
\]

\[
c_x = a_1 \cdot c_{SS} + a_2 \cdot \sin(\omega t) + a_3 \cdot \cos(\omega t) + a_4 \cdot t + a_5 \quad (3)
\]

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 \(2\sigma_{rel}/\mu\), 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 (MCS). 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 ($\Sigma F_{in}$ and $\Sigma F_{out}$). 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. An example is given in the ‘Results and discussion’ 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 ($\Delta S$) is calculated for component loads (TN, TP, COD) contained in sludge (Equation (4))

$$\Delta S_i = V \cdot (x_{i-1} - x_i) \quad i = 1 \ldots n$$

$$\Delta S_i^+ = \max (0, \Delta S_i)$$

$$\Delta S_i^- = |\min (0, \Delta S_i)|$$

An increasing sludge concentration (storage, $\Delta S_i^+$) is counted as an additional input mass flow; a decrease in sludge concentration (release, $\Delta S_i^-$) 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 24 h 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 continuous stirred tank reactor (CSTR), the expected output’s concentration after a given time ($t$) is calculated as follows:

$$\frac{dx_{out}}{dt} = Q/V \cdot (x_{in} - x_{out})$$

(5)

with $\tau = V/Q$ (hydraulic retention of the balanced compound) integration yields

$$x_{out} = x_{in} - (x_{in} - x_0) \cdot \exp\left(-\frac{t}{\tau}\right)$$

(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 1 day ($t = 1$, index $i$) is calculated assuming a constant (mean) influent concentration and flow and a constant volume ($Q_{in} = Q_{out} = Q$):

$$\bar{x}_{out, expected,i} = x_{in,i} - \tau_i \cdot (x_{in,i} - x_{out, expected,i-1}) \cdot (1 - \exp\left(-\frac{1}{\tau_i}\right))$$

(7)
The expected output load is calculated from the initial concentration:

\[ F_{\text{out,expected},i} = Q_{\text{out},i} \cdot \bar{x}_{\text{out,expected},i} \]  

(8)

This expected output load, which is basically calculated from the measured input load (see Equation (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 DS, 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 \( x_{\text{out}} \), which is the way it was specified.

\[ \bar{x}_{\text{in,virtual},i} = (\bar{x}_{\text{in},i} \cdot Q_{\text{in},i} - \bar{x}_{\text{out,fast},i} \cdot Q_{\text{out,fast},i})/Q_{\text{out,slow},i} \]  

(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 Equation (7), the latter (\( \bar{x}_{\text{out,expected},i} \)) was chosen. This value has great influence on \( \bar{x}_{\text{out,expected},i} \). In fact, with long hydraulic retention, \( \bar{x}_{\text{out,expected},i} \) depends almost entirely on the initial concentration (It holds: \( \lim(\exp(-x), \ x \to 0) = 1-x \)). If measured values are used, the expected output concentration \( \bar{x}_{\text{out,expected},i} \) is heavily influenced by the measured output concentration \( \bar{x}_{\text{out},i-1} \). This leads to deterioration of the actual balance (where \( \bar{x}_{\text{out},i} \) is balanced against \( \bar{x}_{\text{out,expected},i} \)). Therefore, only the initial value \( \bar{x}_{\text{out},0} \) is taken from measurements, thereafter this value is taken from \( \bar{x}_{\text{out,expected},i-1} \) of the previous day. This way, all \( \bar{x}_{\text{out,expected}} \) are (almost) only calculated from the input which is a precondition for balancing against the measured values \( \bar{x}_{\text{out}} \).

CUSUM charts were calculated according to Spindler & Vanrolleghem (2012) (see the Appendix for an introduction, online at http://www.iwaponline.com/wqrjc/050/056.pdf). 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 2015). 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 \( \Delta \mu_{\text{opt}} \)) at the cost of slower detection (increasing ARL). Practice has shown that a good choice of \( k \) gives \( \Delta \mu_{\text{opt}} \) 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 AND 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 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 were 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 (Equations (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 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 $\pm 15$–$\pm 25\%$ of the mean concentration for data set A with the exception of TN and TP values for PS.

Data set B, covering almost 1 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 $\pm 6$–$\pm 13\%$ of the mean concentration.

### Table 1: Results from the regression analysis for determination of COD, TN and TP from SS fractions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sludge</th>
<th>Plant</th>
<th>$n$</th>
<th>$a_1c_{SS}$</th>
<th>$a_1c_{SS} + a_2$</th>
<th>$a_1c_{SS} + a_2\sin(\omega t) + a_2\cos(\omega t) + a_3t + a_5$</th>
<th>Suitable SS fraction</th>
<th>$R^2$</th>
<th>$2\sigma_{res}/\mu$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COD</td>
<td>AS</td>
<td>A</td>
<td>175</td>
<td>+</td>
<td>++</td>
<td>$a_1c_{SS} + a_2\sin(\omega t)$</td>
<td>VSS</td>
<td>0.59</td>
<td>17</td>
</tr>
<tr>
<td>COD</td>
<td>AS&amp;PS</td>
<td>B</td>
<td>47</td>
<td>+</td>
<td>++</td>
<td>$a_1c_{SS} + a_2\cos(\omega t)$</td>
<td>VSS</td>
<td>0.95</td>
<td>8</td>
</tr>
<tr>
<td>COD</td>
<td>AS</td>
<td>C</td>
<td>21</td>
<td>+</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS</td>
<td>0.6</td>
<td>7</td>
</tr>
<tr>
<td>COD</td>
<td>PS</td>
<td>A</td>
<td>188</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS} + a_2\sin(\omega t)$</td>
<td>VSS</td>
<td>0.82</td>
<td>25</td>
</tr>
<tr>
<td>COD</td>
<td>PS</td>
<td>C</td>
<td>21</td>
<td>+</td>
<td>++</td>
<td>$a_1c_{SS} + a_2\cos(\omega t)$</td>
<td>VSS</td>
<td>0.96</td>
<td>5</td>
</tr>
<tr>
<td>COD</td>
<td>DS</td>
<td>A</td>
<td>367</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS} + a_2\sin(\omega t)$</td>
<td>VSS</td>
<td>0.43</td>
<td>17</td>
</tr>
<tr>
<td>COD</td>
<td>DS</td>
<td>B</td>
<td>47</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS} + a_2\cos(\omega t)$</td>
<td>VSS</td>
<td>0.69</td>
<td>13</td>
</tr>
<tr>
<td>COD</td>
<td>DS</td>
<td>C</td>
<td>21</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS} + a_2\sin(\omega t)$</td>
<td>VSS</td>
<td>0.94</td>
<td>5</td>
</tr>
<tr>
<td>TN</td>
<td>AS</td>
<td>A</td>
<td>177</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS</td>
<td>0.47</td>
<td>23</td>
</tr>
<tr>
<td>TN</td>
<td>AS&amp;PS</td>
<td>B</td>
<td>47</td>
<td>+</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>TSS (VSS)</td>
<td>0.89</td>
<td>11</td>
</tr>
<tr>
<td>TN</td>
<td>AS</td>
<td>C</td>
<td>21</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS (TSS)</td>
<td>0.67</td>
<td>9</td>
</tr>
<tr>
<td>TN</td>
<td>PS</td>
<td>A</td>
<td>185</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS</td>
<td>0.67</td>
<td>35</td>
</tr>
<tr>
<td>TN</td>
<td>PS</td>
<td>C</td>
<td>21</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS (&amp; NVSS)</td>
<td>0.6</td>
<td>31</td>
</tr>
<tr>
<td>TN</td>
<td>DS</td>
<td>A</td>
<td>365</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS</td>
<td>0.52</td>
<td>16</td>
</tr>
<tr>
<td>TN</td>
<td>DS</td>
<td>B</td>
<td>47</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>TSS (VSS)</td>
<td>0.87</td>
<td>6</td>
</tr>
<tr>
<td>TN</td>
<td>DS</td>
<td>C</td>
<td>21</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS (&amp; NVSS)</td>
<td>0.72</td>
<td>8</td>
</tr>
<tr>
<td>TP</td>
<td>AS</td>
<td>A</td>
<td>177</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS</td>
<td>0.34</td>
<td>23</td>
</tr>
<tr>
<td>TP</td>
<td>AS&amp;PS</td>
<td>B</td>
<td>47</td>
<td>+</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>TSS (VSS)</td>
<td>0.69</td>
<td>19</td>
</tr>
<tr>
<td>TP</td>
<td>AS</td>
<td>C</td>
<td>21</td>
<td>+</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>TSS (VSS&amp;NVSS)</td>
<td>0.87</td>
<td>6</td>
</tr>
<tr>
<td>TP</td>
<td>PS</td>
<td>A</td>
<td>189</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS</td>
<td>0.49</td>
<td>53</td>
</tr>
<tr>
<td>TP</td>
<td>PS</td>
<td>C</td>
<td>21</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS (&amp; NVSS, TSS)</td>
<td>0.43</td>
<td>41</td>
</tr>
<tr>
<td>TP</td>
<td>DS</td>
<td>A</td>
<td>369</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>VSS</td>
<td>0.53</td>
<td>15</td>
</tr>
<tr>
<td>TP</td>
<td>DS</td>
<td>B</td>
<td>47</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>NVSS (TSS,VSS)</td>
<td>0.83</td>
<td>7</td>
</tr>
<tr>
<td>TP</td>
<td>DS</td>
<td>C</td>
<td>21</td>
<td>++</td>
<td>++</td>
<td>$a_1c_{SS}$</td>
<td>NVSS (&amp; VSS, TSS)</td>
<td>0.95</td>
<td>5</td>
</tr>
</tbody>
</table>

‘+’ 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.
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–±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 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.

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. However, 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 refers mainly to TP and TN in PS. The reason for this remains unclear. It probably has to deal mainly with the high variability of PS 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 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 are simulated, these results cannot be reproduced. Only data that are 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. These are the NVSS and COD balances of the anaerobic digester and the total phosphorus balance of the combination of primary clarifier and AS tank (including secondary clarifier). Performance 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, available online at http://www.iwaponline.com/wqrjc/050/056.pdf).

Daily input loads (calculated from flow and concentration):

\[ \sum F_{in,j} = F_{COD, j} + F_{PS, j} + F_{WAS, j} \]

Daily output loads:

\[ \sum F_{out,j} = F_{DS, j} + F_{gas, j} \]

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

(I) \[ \epsilon_{rel,j} = \frac{(\sum F_{in,j} - \sum F_{out,j})}{F_{in}} \]

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

(II) \[ \epsilon_{rel,j} = \frac{(\sum F_{in,j} + \Delta S_{j} - \sum F_{out,j} - \Delta S_{j})}{F_{in}} \]

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 DS (slow output path). The virtual input concentration is therefore calculated from the difference between input load and the fast output load

\[ x_{\text{in, virtual}, i} = \left( \sum F_{\text{in}, i} - F_{\text{COD gas}, i} \right) / Q_{\text{DS}, i} \]

The expected output load results from the virtual input concentration (the digester volume for calculation of \( \tau_i \) is 8,000 m³)

\[ F_{\text{out, expected}, i} = \left( x_{\text{in, virtual}, i} - \tau_i \cdot (x_{\text{in, virtual}, i} - x_{\text{out, i-1}}) \right) \left( 1 - \exp \left( -\frac{1}{\tau_i} \right) \right) \cdot Q_{\text{DS}, i} \]

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

\[ (\text{III}) \ e_{\text{rel}, i} = \left( F_{\text{out, expected}, i} - F_{\text{COD DS}, i} \right) / F_{\text{out, expected}} \]

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 (gray 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 (\( \Delta \mu_{\text{opt}} \)) and the average run length (ARL\(_{1\%} \)) are given. The CUSUM chart signals (dots turning from gray 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 (HRT) is very high at 47 days. The relative standard deviation of the error vector in

![Figure 1](https://iwaponline.com/wqrj/article-pdf/50/3/228/379338/wqrjc0500228.pdf)

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 gray lines (normalized to mean 1). The CUSUM charts (right) signal an off-balance situation (indicated by color changing from gray to black), when the upper or lower graphs exceed their control limit \( h \).
case (I) is 0.34 and the CUSUM chart signals two off-balance periods, once between days 50 and 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 one-third 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 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 sensitively 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.

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 the
For days 70–162 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.

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. Owing 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 (WAS). The PS and the effluent together thus constitute the fast output paths with hydraulic retention of around 1 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 and 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 150.

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 24 h 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 AS 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 extent 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, non-linear 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. The authors would welcome 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 are included in the Appendix (online at http://www.iwaponline.com/wqrjc/050/056.pdf).

**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 are 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


First received 12 December 2014; accepted in revised form 9 January 2015. Available online 12 February 2015