Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Structural redundancy of data from wastewater treatment systems. Determination of individual balance equations

A. Spindler

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

ARTICLE INFO

Article history:
Received 23 September 2013
Received in revised form
4 March 2014
Accepted 5 March 2014
Available online 27 March 2014

Keywords:
Data validation
Gross error detection
Mass balancing
Observability
Redundancy

ABSTRACT

Although data reconciliation is intensely applied in process engineering, almost none of its powerful methods are employed for validation of operational data from wastewater treatment plants. This is partly due to some prerequisites that are difficult to meet including steady state, known variances of process variables and absence of gross errors. However, an algorithm can be derived from the classical approaches to data reconciliation that allows to find a comprehensive set of equations describing redundancy in the data when measured and unmeasured variables (flows and concentrations) are defined. This is a precondition for methods of data validation based on individual mass balances such as CUSUM charts. The procedure can also be applied to verify the necessity of existing or additional measurements with respect to the improvement of the data's redundancy. Results are given for a large wastewater treatment plant. The introduction aims at establishing a link between methods known from data reconciliation in process engineering and their application in wastewater treatment.

1. Introduction

This work discusses a fundamental approach to the validation of operational data from wastewater treatment plants through mass balancing. Historic records of plant data reflect the performance of a treatment plant and are regularly exploited for monitoring, benchmarking and simulation, to adjust control strategies and to plan for process redesign or plant extension. However, poor quality of historic data records is the main obstacle for these tasks. This has been agreed upon widely in literature (e.g. Rieger et al., 2010; Puig et al., 2008; Meijer et al., 2002; Barker and Dold, 1995) as well as different IWA workshops on this question (e.g. Mont Sainte-Anne 2010, Budapest 2011).

The type of operational data typically used for these tasks are daily flow volumes and concentrations measured in 24 h-composite samples (where flow-proportionality is required for matching balances, especially in flows with strongly varying concentrations such as the influent). Higher frequency sensor data is more relevant in automated process control and therefore not of primary interest here. However, sensor readings are usually adjusted to the less frequent but more reliable laboratory measurements. Therefore, the validation of operational data from composite samples is also of considerable relevance for plant control.

E-mail address: a.spi@iwag.tuwien.ac.at.

http://dx.doi.org/10.1016/j.watres.2014.03.042
0043-1354/

Spindler and Vanrolleghem (2012) showed that the application of CUSUM charts is a suitable approach to continuous mass balancing[1] and detects off-balance periods more reliably than mass balances based on long term averages of data. Continuous mass balancing following this method requires individual balance equations which describe redundancy of the measured data.

This work will provide a procedure for the computational determination of the complete set of possible redundancy equations (also: balance equations) for a given plant layout. This aim is different from, but closely related to the principles and objectives of *data reconciliation*. With mass balancing as the key to data reconciliation and gross error detection, there appears to exist a gap between development and application of methods used in process engineering and wastewater treatment. Therefore a very short overview and comparison of the developments in both fields is given in the following parts of the introduction. After the presentation of the proposed method results will be given for its application to a large and complex wastewater treatment plant.

1.1. Data reconciliation in process engineering

Data reconciliation has developed mainly in the field of (chemical) process engineering. It allows to improve the measured values of process variables such as flows and concentrations based on the laws of conservation. Data reconciliation requires *redundancy* of the measured variables which means that they can also be calculated from other measured variables.

A vast amount of literature exists. Research began some 50 years ago when the concept of data reconciliation was introduced by Kuehn and Davidson (1961). Further research developed initially in two lines – the *topology oriented approach* first presented by Václavek (1969; Václavek and Loučka, 1976) and the *equation oriented approach*, represented among others by Crowe (1986; Crowe et al., 1983). Some of the most recent progress in the field has been achieved by Kelly (e.g. 1998; 2004). Four comprehensive books have been written (Madron and Veverka, 1992; Narasimhan and Jordache, 2000; Romagnoli and Sánchez, 2000; Bagajewicz, 2010). Good overviews about research development are also provided in Crowe (1996) and Ponzoni et al. (1999).

A basic step in data reconciliation is the classification of the process variables. A process variable can either be directly *measured* (observed) or *unmeasured*. Unmeasured refers to variables that could be measured (at least theoretically) but are not for some reason. A process variable is *observable*, if it can be calculated from a subset of other measured variables. Measured observable process variables are called *redundant*. Crowe (1989) also classifies *barely observable* (unmeasured) variables which require at least one non-redundant measured variable to be calculated. *Structural* redundancy refers only to the theoretical calculability of a measured variable while *practical* redundancy also considers numerical and statistical accuracy of this calculation. The following short example is given to illustrate the difference between structural and practical redundancy.

The volume of dewatered sludge is negligible compared to influent and effluent of a wastewater treatment plant. For structural redundancy of the overall flow it would, however, still be required to be measured. Obviously the amount of dewatered sludge cannot be reconciled from this balance as the propagation of errors would pose a very high uncertainty on this calculation. On the other hand, in- and effluent would still be practically balanceable without the amount of dewatered sludge being measured.

1.2. Data validation in wastewater treatment

So far the concept of data reconciliation has received little attention in wastewater treatment. This becomes obvious in the terminology. The term *mass balance* is prevalent, possibly inspired by the work of Nowak (1994, 1999). Rieger et al. (2010) actually refer to the *order of redundancy* as “overlapping balances”. It reveals the practitioner’s perspective where the individual mass balances receive higher attention than the reconciliation of the entire data set. This will be discussed further in the following section.

Literature in wastewater treatment focuses mainly on sensor fault detection and so far hardly regards redundancy of measurements. Until recently wastewater related literature cited only two works from the field of data reconciliation in process engineering (Meijer et al., 2002; Puig et al., 2008; Schraa et al., 2006).

Van der Heijden et al. (1994) adapt research from the field of chemical process engineering and apply it to elemental mass balances in fermentation processes. Following works in the field of wastewater treatment (Meijer et al., 2002; Puig et al., 2008) apply the methods of Van der Heijden et al. (1994) thus re-adapting them back into process oriented applications where they originally stem from. Meijer et al. (2002) stress the importance of validation of operational data for use in simulation studies. Puig et al. (2008) point out that the dynamic nature of wastewater treatment makes mass balancing difficult. Both works rely exclusively on the method developed by Van der Heijden et al. (1994) which was implemented in the software Macrobal (Hellinga, 1992). However, when applying data reconciliation to elemental mass balances (Macrobal’s purpose) the composition of substances is exactly known (fixed) which is not the case for the composition of wastewater treatment streams. Hence only in volumetric and mass flow rates the measurement variability was accounted for, but not in measured concentrations. Additionally, the high variability of flow measurements (around 50% relative standard deviation) includes process dynamics which is disputable given the fact the steady state is a prerequisite for the applied method of data reconciliation.

Schraa et al. (2006) does mention data reconciliation citing Crowe (1996) but focuses on sensor fault detection. He did investigate data reconciliation in an earlier publication (Schraa and Crowe, 1998) when he was not yet involved with wastewater treatment.

Very recently two papers on redundancy classification and fault detection based on mass balances where published by

[1] The application of CUSUM charts had originally been labeled “dynamic mass balancing” to differentiate from the established approaches. But because it does not actually target kinetic rates this naming will be avoided in the future.