Application of data reconciliation to a dynamically operated wastewater treatment process with off-gas measurements

Le, Quan H.; Verheijen, Peter J.T.; van Loosdrecht, Mark C.M.; Volcke, Eveline I.P.

DOI
10.1039/d2ew00006g

Publication date
2022

Document Version
Final published version

Published in
Environmental Science: Water Research and Technology

Environmental Science
Water Research & Technology

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Environ. Sci.: Water Res. Technol.*, 2022, **8**, 2114

Application of data reconciliation to a dynamically operated wastewater treatment process with off-gas measurements†

Quan H. Le,[a] Peter J. T. Verheijen,[b] Mark C. M. van Loosdrecht[b] and Eveline I. P. Volcke*[a]

This study deals with the application of data reconciliation to wastewater treatment processes which are subject to dynamic conditions and therefore do not reach a steady-state behaviour *sensu stricto*. The SHARON partial nitritation process, which is operated cyclically with alternating aerated and anoxic periods, is studied as an example. The collected data long-term dynamic data set was split up into data subsets corresponding with different pseudo-steady-state operations, which allowed a better gross error detection. Mass balances were set up taking into account off-gas measurements besides liquid phase measurements and including kinetic relations between measurements based on the biological conversions in the reactor. As a result, a higher number of variables could be reconciled, more key variables could be identified, and gross error detection was facilitated. In order to draw conclusions on the process performance in a shorter period of operation, *e.g.*, on the N_2O emission factor, the average value of the whole data set should be used with caution. The strong dependence of infiltrated air on the aeration regime and gross error in grab sampling (magnitude of 20%) had a substantial impact on calculating N_2O emission. It is recommended that the process performance indicators are derived and checked separately for steady state data subsets to guarantee reliable outcomes.

Received 2nd January 2022,
Accepted 19th May 2022

DOI: 10.1039/d2ew00006g

rsc.li/es-water

Water impact

Data reconciliation is a proven tool to improve data quality and detect gross errors, but its application for wastewater treatment processes remains limited. This study applies data reconciliation to a full-scale SHARON partial nitritation process. A clear added value of including off-gas analysis and kinetic relations was demonstrated, as well as of splitting up the available data in steady state subsets. Moreover, data reconciliation including off-gas analyses allowed more reliable information on N_2O emissions to be drawn.

Introduction

Wastewater treatment processes (WWTPs) treat wastewater collected from households and industries so it can be safely discharged into a receiving water body. Even though the primary goal remains to protect human health and avoid environmental hazards, a paradigm shift is taking place, increasingly regarding WWTPs as resource recovery facilities.[1] Nowadays, WWTPs must fulfil multiple objectives: to comply with (increasingly tighter) discharge standards, to maintain and improve process stability, to guarantee safe operations, to minimize the use of energy and resources, to limit greenhouse gas emissions and at the same time aim at recovering energy and materials. Thanks to instrumentation revolution in the water sector, a large amount of data is collected from various sources, from either routine off-line measurements or online sensors and actuators. Modern small wastewater treatment plants (WWTPs) generate up to 500 signals, whereas larger ones typically register over 30 000.[2] These data provide essential information for WWTP operation, to provide early warning of disturbances and process changes, to track relevant variables and process performance indicators and as the basis of control actions. Still, the success of (advanced) process monitoring, control and optimisation strategies stands or falls with the availability of reliable data.

Depending on the objectives, available historical data are complemented with additional data obtained through one or

[a] *BioCo Research Group, Department of Green Chemistry and Technology, Ghent University, 9000 Ghent, Belgium. E-mail: eveline.volcke@ugent.be*
[b] *Department of Biotechnology, Delft University of Technology, 2600 AA Delft, The Netherlands*

† Electronic supplementary information (ESI) available: (A) Materials and methods: (A1) data reconciliation and gross error detection procedure and (A2) mass balances – stoichiometric and conversion factors & (B) raw and reconciled data set with off-gas data. See DOI: https://doi.org/10.1039/d2ew00006g

more intensive monitoring campaigns. Nevertheless, collecting more data does not necessarily lead to more information. A stepwise experimental design procedure to determine sets of additional measurements that guarantee the identifiability of key process variables was proposed by Le *et al.*[3] Their approach relies on the application of mass balances to relate various measured and unmeasured variables.

Once data are collected, their consistency can be evaluated by data reconciliation, a proven technique to evaluate the consistency of collected data,[4,5] which has been widely applied in (bio)chemical engineering for decades.[6,7] It involves a procedure of optimally adjusting estimates for variables such that these estimates satisfy the conservation laws, *e.g.* in the form of mass balances, and other constraints[4] and are therefore more accurate than the original values. Data reconciliation is often accompanied by statistical tests for gross error detection (measurement validation), which verifies whether the deviation between each estimate and its measurement is acceptable compared to the measurement error.

Redundancy and steady-state are the two preconditions that must be satisfied for the successful application of data reconciliation.[8] The steady-state precondition is typically the most important hurdle for WWTPs. For many other industries, the plant is typically operated for hours or days in a region around a nominal steady-state operating point, so data reconciliation can be easily applied over a period of pseudo-steady-state operation.[8] In contrast, wastewater treatment processes typically face a dynamic input, which contains both diurnal and seasonal patterns. Meijer *et al.*[9] proposed a filtering procedure to select a suitable data set that reflects the pseudo-steady operation of a WWTP. The procedure filtered out any data related to rain events and this filter was considered applicable only for WWTPs in The Netherlands. The filtered data set was then considered suitable for applying data reconciliation. Rieger *et al.*[10] suggested that an intensive measuring campaign should be started after several weeks (*i.e.*, the duration of 2 to 3 times the solids retention time) of stable operation, *i.e.* operation without significant changes in flows, recycles, precipitant dosage, and so on, to obtain reasonable and more reliable data that reflect the typical WWTP operation condition. The data should be collected over a period within about one solids retention time (SRT) to maintain consistency.[9]

With regard to the measurements, one common aspect of the application of data reconciliation to a typical WWTP is that the collected information focuses on the liquid phase. In contrast, the use of off-gas data in data reconciliation of WWTPs is not commonly applied due to the practical limitation of off-gas data collection, requiring closed reactors or the use of an off-gas collection system, preferably connected to an online gas phase analyser. Initially, off-gas data from WWTPs were only used to assess the performance of aeration in activated sludge tanks.[11] Over the last decade, off-gas data has become important to estimate emissions of methane and nitrous oxide, both potent greenhouse gases [*e.g.*, ref. 12 and 13]. Since the metabolism of the microorganisms responsible for the biological conversions is directly related to the gaseous phase through the components O_2, CO_2, N_2 or N_2O, measurements of these gases at WWTPs can be directly connected to these key biological processes. As a result, they have considerable potential for monitoring, controlling and optimising the wastewater treatment process.[14,15] In brief, the off-gas data can give valuable insights into the biological conversions during wastewater treatment. Still, the added value of using WWTP off-gas data for data reconciliation has not yet been demonstrated.

This contribution focuses on the application of data reconciliation to evaluate the consistency of measurements and to gain more process insights into a dynamically operated wastewater treatment process, which is monitored not only with liquid phase measurements but also off-gas measurements. The approach was demonstrated through a case study: a SHARON process for the treatment of ammonium-rich wastewater. Steady-state detection was performed to evaluate the collected data set and split it into smaller data subsets characterized by pseudo-steady-state operation. The applied mass balances involved both liquid phase and off-gas measurements. Some assumptions related to kinetic relations between variables were made and evaluated. Mass balance-based data reconciliation was applied to the whole dataset as well as to each data subset. Based on the reconciled data, a process-specific analysis of the reactor operation was performed, especially concerning greenhouse gas emissions.

Reactor description and measured data

The full-scale SHARON partial nitritation reactor under study is part of the sludge handling facility of the WWTP Dokhaven, Rotterdam and treats the reject water from anaerobic digestion. In the SHARON process, oxidation of ammonium to nitrite is established, while further oxidation to nitrate is prevented by keeping the temperature sufficiently high (35 °C) and the aerobic retention time sufficiently low.[16] The SHARON process is operated without sludge retention, so the hydraulic retention time (HRT) equals the sludge retention time (SRT). This reactor is a completely covered reactor with a constant liquid volume of 1500 m^3 and a headspace volume of approximately 300 m^3, which was permanently kept under pressure (−200 Pa) to prevent odour emissions.[17] However, due to the prevailing under-pressure, ambient air infiltrates into the headspace of the reactor with an unknown flow rate.[18]

Available data of a three-week monitoring campaign[18] consisted of standard operation cycles as well as dedicated experiments concerning prolonged aeration, prolonged anoxic conditions, lowered dissolved oxygen (DO) conditions and shortened cycles. The standard operation cycles were 120-minute-cycles consisting of an aerated period (ON, DO

setpoint = 2 g O_2 m^{-3}) and an anoxic period (OFF), to keep a constant aerobic retention time of 1.35 days despite the varying influent flow rates. The SHARON reactor inflow rate varied between 0 (no flow) and 41 m^3 h^{-1} over the three-week monitoring period but was relatively constant during individual cycles. The liquid feed flow rate (Q_{in}), aeration flow rate (Q_{aer}), off-gas flow rate (Q_{off}), reactor temperature, pH and DO were logged every minute by a SCADA (supervisory control and data acquisition) system. Grab samples of the influent and effluent of the SHARON reactor were taken daily and analysed by standard photometric cuvette tests (Hach-Lange): the influent was analysed for total ammonium (NH_{in}), and the effluent was analysed for total ammonium (NH_{eff}), total nitrite (NO_{2eff}), and nitrate concentrations (NO_{3eff}). The influent was further assumed to contain a constant biodegradable COD (bCOD) value of 100 g m^{-3}, as confirmed by weekly BOD measurements.[13] The aeration gas was analysed for oxygen (O_2) and carbon dioxide (CO_2) while the off-gas was analysed for O_2, CO_2, nitrous oxide (N_2O) and nitric oxide (NO), through an Emerson MLT4 Rosemount FTIR analyser, preceded by a condenser. Fig. 1 presents the layout of measurement and involved measured variables. Given that the reactor volume is constant, the influent flow rate equals the effluent flow rate (Q_{in} = Q_{eff}). Furthermore, the aeration air and infiltration air have the same (atmospheric) composition, so O_{2aer} = O_{2inf} and CO_{2aer} = CO_{2inf}.

Data reconciliation

Main goals and key variables

The data reconciliation procedure proposed by Le,[19] based on the work of Verheijen,[20] was applied to the collected data set (see ESI† A1).

The main goal in this study was to evaluate the consistency of the collected data set through data reconciliation and, more specifically to evaluate the following output variables: (a) conversion efficiency of total ammonium nitrogen (NH = NH_4^+ + NH_3) to nitrite, (b) molar ratio of the total inorganic carbon (TIC = CO_3^{2-} + HCO^{3-} + H_2CO_3) to the total ammonium in the influent, (c) air infiltration flow rate and (d) N_2O emission factors (EF = ratio of mass flow of N_2O in off-gas to the mass flow of incoming total ammonium).

This main goal was then translated into the following list of key variables that needed to be reconciled by data reconciliation:

- The influent flow rate (Q_{in}) and total influent ammonium concentration (NH_{in}); related to main goals (a), (b) and (d).
- The total ammonium (NH_{eff}), total nitrite (NO_{2eff}), and nitrate concentrations (NO_{3eff}) in the effluent; related to main goal (a).
- The influent molar ratio of total inorganic carbon to total ammonium ratio $(TIC:NH)_{in}$; related to main goal (b).
- The infiltrated air flow rate (Q_{inf}); related to main goal (c).
- The concentration of N_2O in the off-gas (N_2O_{off}) and the flow rate of off-gas (Q_{off}) to evaluate N_2O emission factors; related to main goal (d).

Overall, there were 9 key variables, 6 of which were measured (Q_{in}, NH_{in}, NH_{eff}, NO_{2eff}, NO_{3eff}, N_2O_{off}) and 3 of which were unmeasured (Q_{inf}, Q_{off}, $(TIC:NH)_{in}$.) Besides, there were 7 more measurements available which did not concern key variables ($bCOD_{in}$, Q_{aer}, O_{aer}, CO_{aer}, O_{off}, CO_{off}, NO_{off}). The goal of data reconciliation was to find better estimates for the measured key variables and to calculate the unmeasured key variables, from the values of other measured variables and taking into account constraints in the form of mass balances. The objective function of the data reconciliation problem is defined as the weighted least squares of the distance between the measurements vector and the vector of reconciled values weighted by the measurement error. More details are provided in the ESI† A1. The procedure was implemented in Matlab 2014b (MathWorks®).

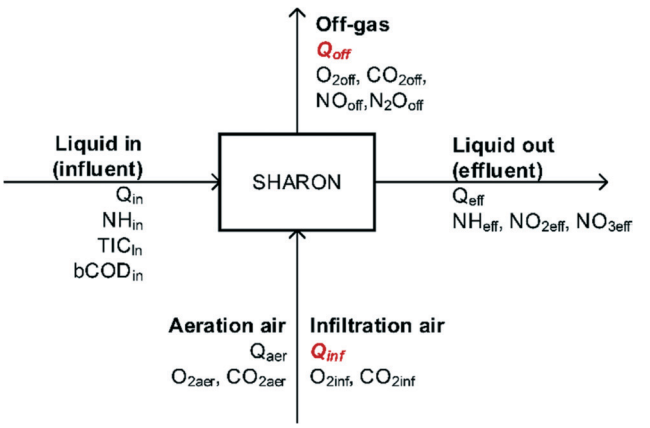

Fig. 1 Reactor mass flows and related variables. The variables in bold italic are unmeasured ones.

Data pre-processing, steady-state detection and subset selection

Steady-state operation is one of the prerequisites for data reconciliation. The five available online measurements directly impacting or reflecting reactor operation in steady-state detection are: (1) influent flow rate, (2) pH, (3) DO, (4) aeration regime and (5) operation temperature. The aeration flow rate was directly related to the DO setpoint, it was not used for steady state detection; the off-gas flow rate was not considered for this purpose either since it resulted from the airflow rate.

Time windows were selected based on (1) a consistent influent flow rate (variation within 10%) and extended as long as the following criteria were met: (2) pH in the range of 6–7, (3) fixed DO setpoint, (4) constant cycle length (120 ± 2 min for standard operation cycles; 60 min for shortened aeration cycles or continuous aeration, see further) and (5) operation temperature of 35 ± 2 °C. Selected time windows were required to contain at least ten cycles.

Each data subset that fulfilled the criteria was split from the raw data set to form a smaller subset of data. From the selected data subsets, average measurement values were calculated to be used in data reconciliation. To this end, 11 consecutive subsets (namely S1–S11) were identified corresponding to 11 pseudo-steady-state operation periods (Fig. 2). These subsets represented 67% of the whole data set. The remaining data did not fulfil the abovementioned criteria (1)–(5) and were therefore omitted from the analysis. The limiting criteria were the ones concerning the influent flow rate (1) and the cycle length (4).

- S1, S3, S4, S5, S7, S10 and S11 were typical operation cycles with a length of 120 min. DO was about 2–2.1 g m^{-3} and pH varied in a very narrow range of 6.6–6.8.

- S2 is a period with shortened cycles with only 60 min per cycle.

- S6 is a period, in which the reactor had a long non-feeding in combination with an anoxic period that lasted 19 hours.

- S8 is a period in which the reactor had non-interrupted aeration.

- S9 was chosen for comparison even though the feeding flow rate in this period was not constant.

The feed flow rate and aeration regime were the main factors determining the splitting of the data set (Table 1) since pH and DO were rather constant in these periods. The average measurement values of S2 are presented in Table 3 as an example. Data of other subsets are presented in ESI† B. These data were used as input for data reconciliation.

Mass balances including off-gas measurements

Considering off-gas measurements, four mass balances were set up over the reactor, related to total nitrogen, COD, carbon, O_2 and nitrogen gas (N_2) (Table 2 and ESI† A). The mass balances consider steady-state conditions and express that for each of the four conserved quantities, the sum of their inflow through the liquid input (*i.e.*, the influent), the aeration air and the infiltration air, and their production through (biological) conversion of conserved quantities, must equal the sum of their outflow in the effluent and in the off-gas and their consumption during biological conversions (see Fig. 1). The conversion factors involved in the mass balances are the stoichiometric coefficients for the biological conversions of nitrifiers and heterotrophs (detailed in ESI† A – Table A1). The mass balances thus include kinetic relations between measurements based on the biological conversions in the reactor.

The total nitrogen balance (m1). The total nitrogen balance (m1, unit: mol h^{-1}) over the reactor expresses that the mass flow of incoming total nitrogen in the form of Kjeldahl nitrogen (Q_{in}·NH_{in}·1.053) in the liquid input equals the sum of the mass flows of outgoing nitrogen in the forms of Kjeldahl nitrogen (Q_{eff}·NH_{eff}·1.053), nitrite (Q_{eff}·NO_{2eff}) and nitrate (Q_{eff}·NO_{3ef}) in the effluent and nitric oxide (Q_{off}·K_{off}·NO_{off}) and nitrous oxide (Q_{off}·K_{off}·N_2O_{off}) in the off-gas. The total Kjeldahl nitrogen concentration was not measured but was estimated as 1.053 of the influent total ammonium concentration, as in Mampaey *et al.*[13] The nitrogen incorporation into the biomass was neglected. Since the influent originated from an anaerobic digester (HRT = 40 days), it was assumed that the influent did not contain nitrite or nitrate. Denitrification by heterotrophs was considered negligible given the short SRT (about 2 days), the low biodegradable organic carbon concentration in the influent and aerobic conditions with DO ≈ 2–3 g m^{-3}.[13] K_{aer} and K_{off} are unit conversion factors for off-gas measurements (from ppm to mol m^{-3}). These conversion factors are temperature dependent and are detailed in ESI† A (Table A2).

The oxygen balance (m2). The oxygen balance (m2, unit: mol h^{-1}) was set up by taking into account the mass flow of supplied oxygen from the aeration air and oxygen in the infiltrated air ((Q_{aer} + Q_{inf})·O_{2aer}·K_{aer}), considering the same-atmospheric-oxygen concentration in both flows, the mass flow of oxygen that left the reactor in the off-gas (Q_{off}·O_{2off}·K_{off}), oxygen that was consumed for ammonium oxidation to nitrite, nitrate, nitrous oxide and nitrous oxide (Q_{in}·(1.434·NO_{2eff} + 1.934·NO_{3eff}) − Q_{off}·(N_2O_{off} + 1.25·NO_{off})·K_{off}) and the mass flow of oxygen for biodegradable COD conversion (0.428·$bCOD_{in}$·Q_{in}). The consumption of oxygen for ammonium oxidation was based on the oxygen required for the formation of nitrite (1.434),

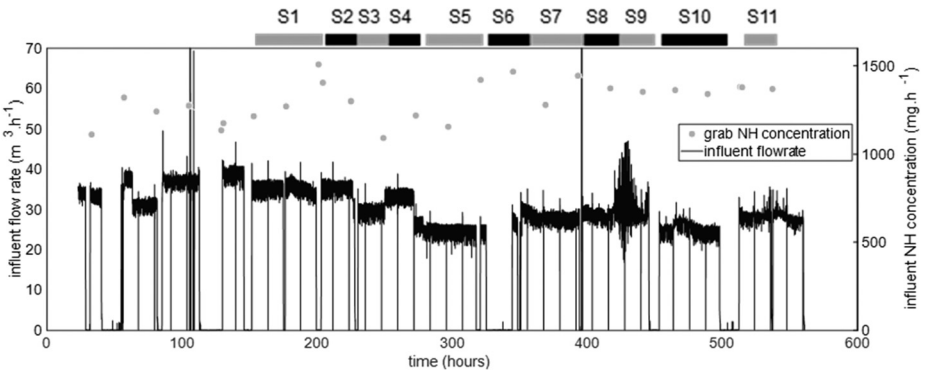

Fig. 2 Detected steady-state operation of SHARON. Filled dots denote liquid sampling.

Table 1 Influent flow rate and operating conditions (average values) for selected data subsets

		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	All cycles
Influent flow rate	[$m^3\ h^{-1}$]	34.6	36.5	30.8	33.5	25.6	7.3	29	28.7	28.1	25.7	27.3	25.0
pH		6.6	6.6	6.6	6.6	6.6	6.7	6.7	6.6	6.9	6.7	6.7	6.9
DO	[$g\ L^{-1}$]	2.1	2.1	2.1	1.9	2.1	2.3	2.1	2.2	2.1	2.1	2.1	2.2
Cycle length	Aeration ON [min]	96	49	81	91	68	Low aeration	76	Continuously aerated	75	68	75	57.2%
	Aeration OFF [min]	24	11	39	29	52		44		45	52	45	42.8%
Length	[Hours]	42.5	21.8	19.4	21.8	41.1	32.8	41.2	23.9	25.3	43.1	23.1	504.2

nitrate (1.934), nitrous oxide (1) and nitric oxide (1.25), according to the stoichiometry of the corresponding (biological) reactions (ESI† – Table A1). Oxygen requirement to consume the influent organic carbon COD was calculated from the biological reaction of heterotrophs (Table A1 in ESI†).

Carbon (C) balance (m3). In the carbon (C) balance (m3, unit: mol h^{-1}), the incoming C-species were the TIC (related to the molar ratio of the total inorganic carbon as follows: $(TIC:NH)_{in}\cdot Q_{in}\cdot 1.053\cdot NH_{in}$) and the CO_2 from the aeration and infiltrated air ($(Q_{aer} + Q_{inf})\cdot CO_{2aer}\cdot K_{aer}$) considering the same-atmospheric-oxygen concentration in both flows. CO_2 left the reactor in the off-gas ($Q_{off}\cdot CO_{2off}\cdot K_{off}$). It was reasonably assumed that all inorganic carbon was stripped as CO_2.[16] Therefore, the liquid outflow of inorganic carbon was neglected (effluent pH was 6.2–6.9), and so was the accumulation of inorganic carbon in the reactor (steady state). CO_2 was produced during the conversion of the organic carbon present in the liquid influent by heterotrophs ($0.303\cdot COD_{in}\cdot Q_{in}$) and consumed by nitrifiers ($-0.067\cdot Q_{in}\cdot(NO_{2eff} + NO_{3eff})$). The origin of the stoichiometric coefficients is detailed in ESI† A (Table A1).

The nitrogen gas (N_2) balance (m4). The nitrogen gas (N_2) balance (m4, unit: mol h^{-1}) in the gas phase was set up given that heterotrophic denitrification was considered negligible. Therefore, N_2 could be seen as a conserved quantity and its mass balance was separately presented, expressing that the total mass flow of incoming N_2 in aeration and infiltrated air, $(Q_{aer} + Q_{inf})\cdot(10^6 - O_{2aer} - CO_{2aer})$ equals the N_2 in the off-gas, $Q_{off}\cdot(10^6 - O_{2off} - CO_{2off})\cdot K_{off}$. In the latter expressions, the fraction (in ppm) of the nitrogen in the gas phase was calculated as the complement of the oxygen and CO_2 fractions.

Mass balances without off-gas measurement

Only two mass balances could be set up without using the information of the off-gas measurements.

Table 2 Overview of the mass balances considered in this study. Each mass balance m_i (i = 1,6) is set up by expressing that the sum of all components in the corresponding row must equal zero. Positive signs indicate inflow or production of a component, negative signs indicate outflow or consumption. The resulting mass balances can be further simplified considering $Q_{in} = Q_{eff}$, $O_{2aer} = O_{2inf}$ and $CO_{2aer} = CO_{2inf}$ (see Fig. 1) and are listed in ESI† A

→ Reactor (in/out)flows & conversion terms	Liquid in (influent)	Liquid out (effluent)	Conversion	Aeration air	Infiltration air	Off-gas
↓ Conserved quantity						
With off-gas measurements						
m1: total nitrogen [mol h^{-1}]	$Q_{in}\cdot NH_{in}\cdot 1.053$	$-\ Q_{eff}\cdot NH_{eff}\cdot 1.053$ $-\ Q_{eff}\cdot NO_{2eff}$ $-\ Q_{eff}\cdot NO_{3eff}$				$-\ Q_{off}\cdot K_{off}\cdot NO_{off}$ $-\ Q_{off}\cdot K_{off}\cdot N_2O_{off}$
m2: oxygen [mol h^{-1}]			$-\ 1.434\cdot NO_{2eff}\cdot Q_{eff}$ $-\ 1.934\cdot NO_{3eff}\cdot Q_{eff}$ $-\ 0.428\cdot bCOD_{in}\cdot Q_{in}$ $-\ Q_{off}\cdot 1.25\cdot NO_{off}\cdot K_{off}$ $-\ Q_{off}\cdot N_2O_{off}\cdot K_{off}$	$Q_{aer}\cdot O_{2aer}\cdot K_{aer}$	$Q_{inf}\cdot O_{2inf}\cdot K_{aer}$	$-\ Q_{off}\cdot O_{2off}\cdot K_{off}$
m3: carbon [mol h^{-1}]	$(TIC:NH)_{in}$ $\cdot Q_{in}\cdot 1.053\cdot NH_{in}$		$0.303\cdot bCOD_{in}\cdot Q_{in}$ $-\ 0.067\cdot Q_{in}\cdot NO_{2eff}$ $-\ 0.067\cdot Q_{in}\cdot NO_{3eff}$	$Q_{aer}\cdot CO_{2aer}\cdot K_{aer}$	$Q_{inf}\cdot CO_{2inf}\cdot K_{aer}$	$-\ Q_{off}\cdot CO_{2off}\cdot K_{off}$
m4: N_2 [mol h^{-1}]				$Q_{aer}\cdot K_{aer}\cdot(10^6 - O_{2aer} - CO_{2aer})$	$Q_{inf}\cdot K_{aer}\cdot(10^6 - O_{2inf} - CO_{2inf})$	$-\ Q_{off}\cdot K_{off}\cdot(10^6 - O_{2off} - CO_{2off})$
Without off-gas measurements						
m5: total nitrogen [mol h^{-1}]	$Q_{in}\cdot NH_{in}\cdot 1.053$	$-\ Q_{eff}\cdot NH_{eff}\cdot 1.053$ $-\ Q_{eff}\cdot NO_{2eff}$ $-\ Q_{eff}\cdot NO_{3eff}$				
m6: oxygen [mol h^{-1}]			OC $-\ 1.434\cdot NO_{2eff}\cdot Q_{eff}$ $-\ 1.934\cdot NO_{3eff}\cdot Q_{eff}$ $-\ 0.428\cdot bCOD_{in}\cdot Q_{in}$			

Table 3 Identified variables (*i.e.*, reconciled and checked for gross errors) of subset S2, both for including and for not considering (grey-shaded rows) off-gas data. Key variables are indicated in bold

#	Measured variable [unit]	Measured value	s.e. $\sqrt{\text{var}(y)}$	Reconciled value	s.e. $\sqrt{\text{var}(\bar{x})}$	Δ (%)	i (%)
1	CO_{2aer} [ppm]	550	12	550	12	100.00	0.00
2	CO_{2off} [ppm]	304 96	159	30 496	159	100.00	0.00
3	$bCOD_{in}$ [mole per m^3]	2.78[a]	0.01	2.78	0.01	100.00	0.00
				2.78	0.01	100.00	0.00
4	**N_2O_{off}** [ppm]	**296**	**9**	**296**	**9**	**99.96**	**0.00**
5	**NH_{eff}** [mole per m^3]	**44.8**	**2.2**	**43.1**	**2.1**	**96.16**	**6.72**
				43.2	**2.08**	**96.6**	**6.81**
6	**NH_{in}** [mole per m^3]	**96.6**	**4.8**	**104.6**	**3.0**	**108.28**	**37.02**
				103.7	**3.0**	**107**	**37.6**
7	**NO_{2eff}** [mole per m^3]	**66.0**	**3.3**	**62.4**	**2.8**	**94.50**	**14.37**
				62.8	**2.8**	**95.2**	**13.8**
8	**NO_{3eff}** [mole per m^3]	**0.8**	**0.0**	**0.8**	**0.04**	**99.93**	**0.00**
				0.8	**0.04**	**95.2**	**13.8**
9	NO_{off} [ppm]	44	0.	44	0	100.00	0.00
10	O_{2aer} [ppm]	210 280	13	210 280	13	100.00	0.00
11	O_{2off} [ppm]	187 767	93	187 766	93	100.00	0.00
12	Q_{aer} [m^3 h^{-1}]	2292	35	2292	35	100.00	0.00
13	**Q_{in}** [m^3 h^{-1}]	**36.48**	**0.05**	**36.5**	**0.05**	**100.00**	**0.00**
				36.5	**0.05**	**100.00**	**0.00**

#	Unmeasured variable [unit]	Calculated from raw data	s.e.	Reconciled value	s.e.	Δ (%)	i (%)
14	**Q_{inf}** [m^3 h^{-1}]	**1924**	**209**	**1671**	**179**	**86.86**	**14.38**
15	**Q_{off}** [m^3 h^{-1}]	**4438**	**218**	**4201**	**187**	**94.65**	**14.22**
16	**$(TIC:NH)_{in}$ [mol mol^{-1}]**	**1.46**	**0.10**	**1.27**	**0.04**	**87.38**	**61.35**

std = standard error of the mean; Δ (%) = ratio of reconciled value to measured value; i (%) = precision improvement (reconciled & checked by gross error detection).[a] Assuming a composition $CH_{1.5}O_{0.5}$, corresponding with a COD content of 36 gCOD mol^{-1}.

The total nitrogen balance (m5). The total nitrogen balance (m5) over the reactor expresses that the mass flow of total incoming nitrogen in the form of total Kjeldahl nitrogen (Q_{in}·NH_{in}·1.053) in the liquid phase equals the mass flows of the outgoing nitrogen in the forms of ammonia, nitrite and nitrate (Q_{in}·(1.053·NH_{eff} + NO_{2eff} + NO_{3eff})) in the effluent (liquid phase). In contrast to m1, this mass balance neglects the formation of NO and N_2O, out of necessity because no off-gas data are available.

The oxygen balance (m6). The oxygen balance (m6) was formed by taking into account that the mass flow of oxygen required for nitrification and COD removal (OC) equals the required oxygen for ammonium oxidation to nitrite and nitrate (Q_{in}·(1.434·NO_{2eff} + 1.934·NO_{3eff})) and for biodegradable COD conversion (0.428·$bCOD_{in}$·Q_{in}). It was assumed that the mass flow of oxygen for conversion of ammonium to NO and N_2O was negligible. In case no off-gas data are available, the oxygen consumption (OC) is an additional variable which needs to be identified (calculated), based on this mass balance.

Performance indicators

Data reconciliation and gross error detection were applied to the 11 subsets for two scenarios: considering mass balances including off-gas measurements as well as considering mass balances without off-gas measurements. The results were subsequently compared. The number of reconciled key variables and the precision improvement of key variables were used as performance indicators in this comparison.

The precision improvement of the key variables (i_x) is defined as the ratio of the difference between the variance of the measurement (var(y)) and the variance of the reconciled value (var($\bar{x}$)) to the variance of the measurement, expressed in percentage. The precision improvement i_x is also referred to as the effect of balancing.[7]

$$i_x = \frac{\text{var}(y) - \text{var}(\bar{x})}{\text{var}(y)} \times 100$$

The value of i_x is always between 0 and 100 and is typically positive, which means that a better estimate (*i.e.*, characterized by a smaller variance) is found for the key variable. The higher the i_x, the more improved the key variables got through data reconciliation.

Results and discussion

Consistency of the data set and reconciled data with and without off-gas data

The results of data reconciliation of subset S2 with and without off-gas data are presented as an example (Table 3). The results for the whole dataset and for the other subsets, when including off-gas data, are presented in ESI† B.

Considering off-gas data in data reconciliation resulted in a higher number of variables which could be reconciled (Table 3 shows 16 reconciled values with off-gas data and 6 reconciled values without off-gas data). All key variables (Q_{in}, NH_{in}, NH_{eff}, NO_{2eff}, NO_{3eff}, N_2O_{off}, Q_{inf}, Q_{off}, $(TIC:NH)_{in}$.) were reconciled when off-gas data were available, while the key variables related to off-gas measurements (N_2O_{off}, Q_{inf}, Q_{off}) and the key variable $(TIC:NH)_{in}$ could not be calculated when off-gas data were not available. Also the calculation of the N_2O emission factor ($EF = Q_{off}{\cdot}N_2O_{off}/Q_{in}{\cdot}NH_{in}{\cdot}1.053$ expressed as %) could only be calculated using off-gas measurements. The reconciled value of all off-gas measurements was similar to the corresponding raw data.

With regard to the results of gross error detection of S2, the global test did identify gross error in the data set. The measurement test indicated a large discrepancy between the reconciled values and the measured values and the nodal test indicated a large residue in the nitrogen balance in these subsets. Since all tests signalled a gross error, data set S2 was concluded to contain gross error(s) in the concentration measurements. Overall, the result of gross error detection was quite similar with or without off-gas data concerning detecting the gross errors in the liquid phase measurements (see Table B2b, ESI† B for more details of gross error detection results).

By examining the reconciled results in Table 3, the gross error could be allocated to the grab sampling measurements of the influent ammonium (NH_{in}), effluent ammonium (NH_{eff}) or effluent nitrite (NO_{2eff}) as they exhibited the largest differences between reconciled and measured values (by checking Δ (%) value, Table 3). More importantly, by examining the raw data of S2 (data not shown), it was clear that the total mass flow rate of nitrogen in the influent in the form of ammonium at some measurement points was about 20% lower than the total mass flow rate of nitrogen in the effluent in forms of ammonium, nitrite and nitrate.

The gross error detection (with and without off-gas) also signalled gross errors in subsets S6, S7, S8, S9 and S10. These subsets exhibited a significant deviation between the reconciled values and measured values of grab samples (Fig. 3). For example, the total mass flow of nitrogen in the influent of subset S10 was 12% higher than that of the effluent, and the total mass flow of nitrogen in the effluent of S6 was 12% lower than the one in the effluent (see Tables B6b, B7b, B8b, B9b and B10b, ESI† B for more details). The impact of the assumptions in setting up mass balances on the results of data reconciliation was investigated through a sensitivity analysis. More specifically, the value of the output variables was calculated when varying the values of the influent Kjeldahl-nitrogen to ammonium ratio, the influent bCOD concentration, the percentage of CO_2 stripped and some stoichiometric conversion factors (depending on assumed biomass yields) within realistic ranges (Table 4). The results are elaborated in what follows.

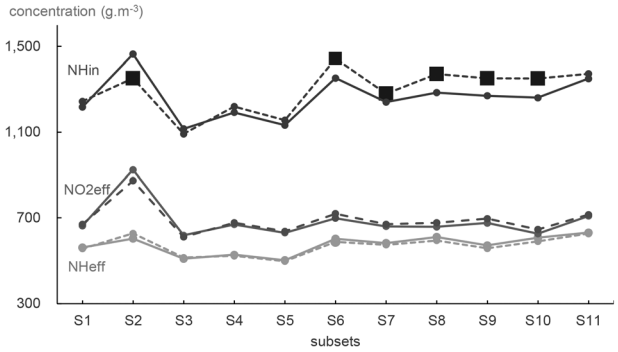

Fig. 3 Measured (dashed line) and reconciled (continuous line) values of influent ammonium (NH_{in}, square), effluent ammonium (NH_{eff}, circle) and effluent nitrite (NO_{2eff}, triangle) using off-gas data. Big squares indicate a potential gross error in the measurements.

Ammonium conversion efficiency

Considering the raw data subsets, the conversion efficiency of ammonium to nitrite varied from about 46% to 68% (Fig. 4A). The ammonium conversion efficiencies in the periods of S6 and S10 were low (46 & 48%, respectively), while their value was rather high in S2 (68%). These unexpected conversion efficiencies calculated from the raw data in periods S2, S6 and S10 were potentially the consequence of the gross errors in the ammonium measurements during these periods.

In contrast to raw data, the reconciled data presented a consistent ammonium conversion efficiency, which varied in the range of 51–60% between the subsets (Fig. 4A). For the periods S2, S6 and S10 with potential gross errors, the conversion efficiencies of 68%, 46% and 48% calculated from the raw data were reconciled to be 60%, 53% and 51%, respectively. The nitrite : ammonium ratio in the effluent ($NH_{eff}:NO_{2eff}$) was found to be about 1.06–1.38 (Fig. 4B), which is suitable for the subsequent anammox reactor.

The average conversion efficiency of ammonium to nitrite for the whole dataset was 52%. This value was only slightly affected (range 51–56%) by the Kjeldahl to ammonium ratio assumed in setting up the mass balances (Table 4). It was not affected by the assumed influent bCOD concentration, the amount of CO_2 stripped and the given stoichiometric conversion factors – at least not within the realistic parameter ranges considered.

Molar ratio of total inorganic carbon to ammonium in the influent

The raw data set showed a relatively wide range of the molar ratio of total inorganic carbon to (total) ammonium in the influent, *i.e.*, $(TIC:NH)_{in}$, with average values between subsets ranging from 0.94 to 1.46. Using the reconciled data, the average $(TIC:NH)_{in}$ ratio was in the range of 1.04–1.27 for all subsets (Fig. 5). The average value of $(TIC:NH)_{in}$ for the whole monitoring campaign was 1.17 ± 0.04, which agreed quite well with a typical range of 1–1.4 for this type of influent reported in the literature.[21]

Table 4 Influence of assumptions on the estimated values of output variable, namely the conversion efficiency of ammonium to nitrite, ratio of total inorganic carbon to (total) ammonium in the influent, infiltrated air, and N_2O emission factor (EF). Average values for the whole dataset (grouping all data subsets)

Assumptions	Default value	Range	Ammonium conversion efficiency [–]	$(TIC:NH)_{in}$ [mol mol^{-1}]	Infiltration air [m^3 h^{-1}]	N_2O EF [%]
Reference value of output variables (for default assumed values)			0.52 ± 0.02	1.17 ± 0.04	1454 ± 140	3.67 ± 0.13
Influent Kjeldahl-nitrogen to ammonium ratio	1.053	1–1.20	0.51–0.56	1.19–1.09	1425–1511	3.75–3.43
Influent bCOD (mol m^{-3})	100	0–300	0.52	1.16–1.18	1403–1556	3.61–3.78
Stripped CO_2	100%	75–100%	0.52	1.55–1.17	1453–1454	3.67
Stoichiometric conversion factor of ammonium to nitrite (mol mol^{-1})	1.43	1.4–1.5[a]	0.52	1.14–1.21	1382–1591	3.58–3.82
Stoichiometric conversion factor from bCOD to CO_2 (mol mol^{-1})	0.428	0.15–0.5[b]	0.52	1.15–1.17	1420–1462	3.63–3.68
Stoichiometric conversion factor from bCOD to O_2 (mol mol^{-1})	0.303	0.15–0.5[b]	0.52	1.17–1.16	1454	3.67

[a] Based on the yield of ammonium oxidizing bacteria of 0.1–0.2 gCOD gN^{-1}. [b] Based on the yield of heterotrophs of 0.5–0.85 gCOD gCOD^{-1}.

The $(TIC:NH)_{in}$ ratios were based on the assumption that all CO_2 was stripped from the liquid phase during aeration, thus neutralizing protons produced during nitrification. The reconciled value of the $(TIC:NH)_{in}$ ratio was strongly affected by the assumption related to stripped CO_2: the reconciled value increased from 1.17 to 1.55 as the stripped CO_2 decreased from 100% to 75% (Table 4). Indeed, a lower fraction of CO_2 stripped requires a higher influent TIC concentration to correspond with a given CO_2 off-gas concentration. The $(TIC:NH)_{in}$ ratio was positively correlated to the stoichiometric conversion factor of ammonium to nitrite (Table 4). The latter can be explained as follows: a higher stoichiometric nitrite production is accompanied by a higher proton production, resulting in more CO_2 stripped, which originates from a higher influent TIC concentration.

Based on the raw data, no correlation was found between the $(TIC:NH)_{in}$ ratio and the conversion efficiency of ammonium to nitrite. From the reconciled data, a high correlation of 0.84 in between the $(TIC:NH)_{in}$ ratio and the conversion efficiency of ammonium to nitrite was observed. The latter result is consistent with previous observations reported in the literature that the $(TIC:NH)_{in}$ ratio determined the ratio of ammonium to nitrite in the effluent.[16,22–24] The reconciled data indicated that the reactor was in stable operation while raw data indicate some periods with high variability.

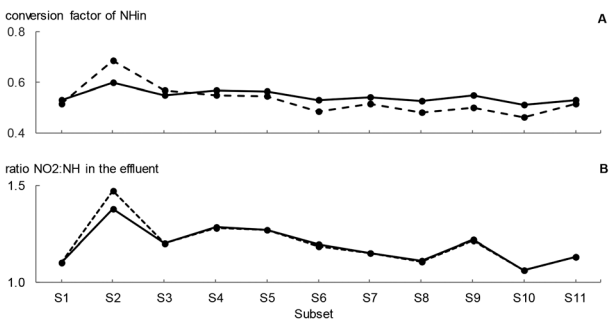

Fig. 4 (A) The conversion efficiency of ammonium in the effluent to nitrite and (B) the ratio of ammonium to nitrite in the effluent. The dashed line represents values from raw data, the continuous line represents values from balanced data. Average values for each data subset.

Air infiltration

The average infiltrated air flow rate for the whole monitoring period was reconciled to be Q_{inf} = 1454 ± 140 m^3 h^{-1} (Fig. 6), which equalled to about 90% of aeration air supplied into the reactor (Q_{aer} = 1653 ± 9 m^3 h^{-1}). The off-gas flow rate (Q_{off}) was also reconciled to be 3287 ± 147 m^3 h^{-1} on average. Considering the individual data subsets, the infiltrated air flow rate was in the range of 600–2500 m^3 h^{-1}, which represented 21–76% of the off-gas flow rate (Fig. 6). The infiltrated air was inversely proportional to the average aeration airflow rate with a correlation factor of 0.92. The infiltration flow rates were much higher in the periods with lower aeration (Fig. 6). Especially in subset S6, the reactor was unaerated for an extended period of almost 24 h, and the infiltrated air accounted for 76% of the off-gas flow rate. On the other hand, in subset S8, the reactor was continuously aerated for the whole period, and the infiltrated air was the lowest compared to that of the other subsets and represented only 21% of the off-gas.

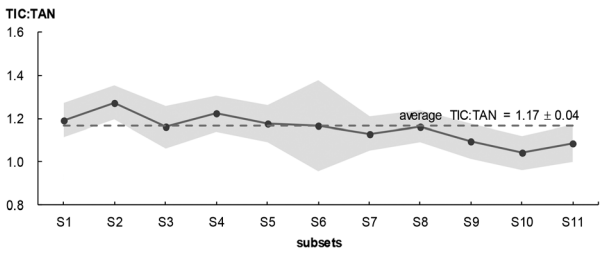

Fig. 5 Influent molar ratio of total inorganic carbon to total ammonium, $(TIC:NH)_{in}$ – average values for each data subset. The grey shade indicates the standard error of the mean.

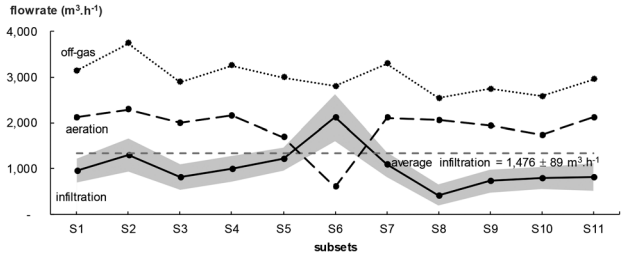

Fig. 6 Aeration air (dashed line), infiltrated air (continuous line) and off-gas (dotted line) flow rates – average values for each data subset. The shaded grey band represents the uncertainty of the reconciled infiltration air.

The reconciled value of the infiltrated air flow rate (Q_{inf}) varied in the range of ±7% (Table 4), depending on different assumptions in setting up the mass balances such as the influent Kjeldahl nitrogen to ammonium ratio, the bCOD in the influent, the stripped CO_2 and the conversion factor of ammonium to nitrite.

N_2O emission factor

The N_2O emission factor (EF) was calculated from raw data to be 3.78% and reconciled by data reconciliation to be 3.67%. This factor means that 3.67% of the incoming nitrogen in the form of ammonium was converted to N_2O, which is equivalent to the emission rate of 1196 gN_2O–N h^{-1} or 28.72 kg N_2O–N d^{-1} throughout the whole monitoring campaign. However, examining the reconciled data of each subset revealed the dynamics of N_2O emission, which was closely related to the aeration regime (Fig. 7). There was a high correlation (of about 0.9) between the length of the anoxic phase of the aeration cycle and the emission factor with either raw or reconciled data. The longer the anoxic time per cycle lasted, the higher the emission factor (see Fig. 7).

When the reactor was interrupted by a non-feeding period in combination with a long unaerated period such as in periods S6 (19 h), S9* (7 h between periods S9 and S10) and S10* (15 h between periods S10 and S11), high emission factors of 5–9% were observed (Fig. 7). Notably, since the reactor was operated with a lowered DO set point of 0.6 g m^{-3} in period S6, the emission factor could surge up to 8.65 ± 0.74%. In contrast, when the reactor was continuously aerated in period S8, a very low emission factor of 1.46 ± 0.08% was found. Table 4 shows that the reconciled value of the emission factor was slightly influenced by the assumptions during setting up the mass balances and varied in a narrow range of ±5%.

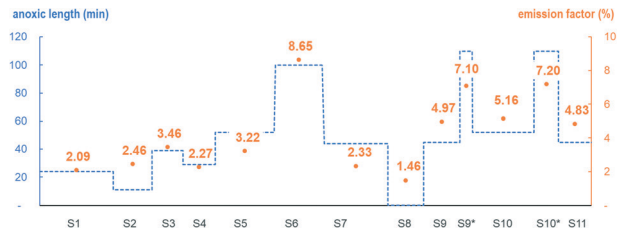

Fig. 7 The N_2O emission factor (value with orange dot) and length of the anoxic phase (dashed line) – average values for each data subset. Data subsets marked with (*) in the *x*-axis label denote periods with a long anoxic phase combined with non-feeding.

Prolonging the aeration time relative to the anoxic time and minimising the non-feeding and anoxic periods were considered as mitigation measures to reduce N_2O emission, be it at the higher expense for aeration. The observation of a high N_2O emission factor of 7% or higher, as the result of interrupted feeding combined with non-aeration, was also reported previously. Mampaey *et al.*[13] reported that there was an increased formation of N_2O under anoxic conditions, which contributes to 80% of the emitted N_2O in regular operation. The formed N_2O was then stripped during the aerated condition. Desloover *et al.*[25] reported an emission factor of 5.1–6.6% under the standard operation of a partial nitritation sequencing batch reactor (SBR), operated at a low DO setpoint (0.75 g O_2 m^3) and a short SRT of 1.7 d. This emission factor increased to up to 9% after a period of non-feeding.

In the study of Mampaey *et al.*[13] on the same reactor, the reported N_2O emission factors for different operation conditions were doubtful since they were calculated based on a fixed infiltrated air flow rate and potential erroneous measurements in the liquid phase. For example, the reported EF of 1.8% of S2 was calculated from a fixed infiltrated air flow rate of 910 m^3 h^{-1}, which was an average flow rate for the whole monitoring campaign. With reconciled data, the infiltrated air flow rate in S2 was 1674 m^3 h^{-1}, which was 80% higher than the previously reported values. Besides, the incoming nitrogen measurements used in calculating the emission factor of S2 suffered from gross error. The gross error detection and the examining of the raw data confirmed that the mass flow rate of the total nitrogen in the influent of S2, represented by ammonium measurements, was 10–20% lower than the mass flow rate of total nitrogen in the effluent, represented by ammonium, nitrite and nitrate measurements.

In sum, the time-variation of the infiltrated air flow rate strongly depended on the aeration regime and several measurements of nitrogen in the liquid phase were subject to a gross error of magnitude of 20%. As a result, the advantage of shortening the aeration cycle to reduce N_2O emissions reported by Mampaey *et al.*[13] could not be confirmed.

Dealing with dynamics

Separately checking the data quality for each steady state sub-period allowed drawing additional conclusions on the short-term process performance. Indeed, gross errors occurring during a shorter period of a dynamic process are not always visible from long-term average process data.

This is illustrated by two critical remarks when drawing a conclusion about the N_2O emission factor for shorter periods

related to the dynamics of the infiltrated air and the quality of grab sampling.

First, for a typical operation condition, the infiltrated air flow rate was inversely proportional to the aeration air flow rate (correlation factor of 0.92). It was the result of actively drawing of air out of the headspace of the reactor and maintaining a constant negative headspace pressure of 200 Pa to control odour. Since the dynamics of air infiltration influenced the calculation of the N_2O mass flow rate in the off-gas, using an average infiltrated air flow rate of the whole measuring campaign to calculate the emission factor for a shorter period is not recommended.

Second, while the average data of the whole data set only showed a small gap (1–1.3%) in the nitrogen mass balances, confronting the reconciled results with the raw data for each subset revealed that some gaps may be as large as 20%. Since the incoming nitrogen in the form of ammonium directly affects the calculation of the N_2O emission, the nitrogen mass balances of the shorter periods should be checked in order to calculate and draw a conclusion regarding the N_2O emission. Despite the fact that the imbalance of nitrogen in the liquid phase could be noticed by examining the raw data, the gross errors related to these measurements may not always be visible even to the expert eye, especially in dealing with large and complicated data sets. Data reconciliation offers a more structured approach to pinpoint the errors and to eliminate the interferences of using raw data for process analysis.

Added value of including off-gas data and kinetic relations

The potential for data reconciliation and gross error detection to gain process insights was significantly increased by adding off-gas measurements on top of measurements in the liquid phase. Without off-gas measurements, only mass balances for total nitrogen and oxygen could be set up and the latter could only be used to calculate the oxygen consumption (OC). Considering off-gas measurements allowed addition of mass balances for carbon and nitrogen gas (m3 and m4 in Table 2). Besides, the mass balance for total nitrogen was more precise when the gas phase concentrations of NO and N_2O could be included (m2 *versus* m6 in Table 2). As a result, more variables could be reconciled when considering off-gas data than in case this was not done (16 *versus* 6), including all key variables. Besides, the results of the gross error detection had a higher probability considering off-gas data.

The added value of gas phase O_2, CO_2 and N_2 measurements to gain information on the biological activity in the system was previously pointed out by Hellinga *et al.*[14] and Leu *et al.*[15] Two well-known applications of off-gas measurements in data reconciliation were from Strous *et al.*[26] and Lotti *et al.*[27] In these studies, long-term operating data from a lab-scale reactor, including off-gas measurement of CO_2, was used in data reconciliation to calculate the stoichiometry of conversion of carbon and nitrogen compounds by anammox biomass.

The mass balances with off-gas measurements include kinetic relations (stoichiometric conversions factors) between process variables. These types of constraints are widespread in the application of data reconciliation in the power industry such as the use of equations related to the ellipse law or isentropic efficiency calculation.[28,29] The latter can potentially produce bias in the result of data reconciliation.[20] However, the impacts of these kinetic coefficients and their corresponding assumptions were verified to be small, and the results were justified (Table 4).

Still, the contribution of off-gas measurements in this study, however, was relatively limited because there was a lack of redundant measurements in the collected data set. The air infiltrated into the headspace of the reactor was unavoidable and unmeasurable. Nonetheless, the unmeasured off-gas flow rate was the main weakness in the monitoring campaign. These two unknown measurements in the off-gas data reduced the redundancy of the system of mass balances (results not shown), which potentially prevented more interesting process insights to be found.

Concerning the monitoring planning, the collected data would be more useful if some additional measurements such as influent COD, alkalinity, total Kjeldahl nitrogen, *etc.* could have been done to maximise the potential of off-gas measurements to verify the quality of the collected data using the mass balances. Even though the impact of these additional measurements would be small compared to the unknown off-gas flow rate, they would definitely contribute to the redundancy of the collected data set. These additional measurements could be efficiently planned using the procedure proposed by Le *et al.*[3]

Conclusions

The possibility of applying data reconciliation to dynamically operated processes and the added value of including off-gas data besides liquid phase measurements were demonstrated by applying data reconciliation to a full-scale SHARON partial nitritation reactor.

Methodological advances in data reconciliation were realized dealing with process dynamics and including off-gas data. Splitting up the long-term dynamic data set in steady state data subsets resulted in a better detection of gross errors, which may not always be visible even to the expert eye. In order to draw conclusions on the process performance in a shorter period of operation, the average value of the whole data set should be used with caution and steady state data subsets should be analyzed instead. Including off-gas data and considering kinetic relations between variables allowed the set-up of additional mass balances. As a result, a higher number of variables could be reconciled, more key variables could be identified, and gross error detection was facilitated.

Process-specific information was gained as well, especially concerning N_2O emissions. The N_2O emission dynamics were closely related to the aeration regime, showing a high

correlation between the length of the unaerated phase and the emission factor. The advantage of shortening aeration cycles to reduce N_2O emissions reported in a previous study could not be confirmed due to the gross error in the grab samples and the dynamics of infiltrated air.

Conflicts of interest

There are no conflicts to declare.

References

1 W. Mo and Q. Zhang, Energy-nutrients-water nexus: Integrated resource recovery in municipal wastewater treatment plants, *J. Environ. Manage.*, 2013, **127**, 255–267.
2 G. Olsson, B. Carlsson, J. Comas, J. Copp, K. V. Gernaey, P. Ingildsen, U. Jeppsson, C. Kim, L. Rieger, I. Rodriguez-Roda, J. P. Steyer, I. Takacs, P. A. Vanrolleghem, A. Vargas, Z. Yuan and L. Amand, Instrumentation, control and automation in wastewater - from London 1973 to Narbonne 2013, *Water Sci. Technol.*, 2014, **69**, 1373–1385.
3 Q. H. Le, P. J. T. Verheijen, M. C. M. van Loosdrecht and E. I. P. Volcke, Experimental design for evaluating WWTP data by linear mass balances, *Water Res.*, 2018, **142**, 415–425.
4 C. M. Crowe, Data reconciliation - Progress and challenges, *J. Process Control*, 1996, **6**, 89–98.
5 D. B. Ozyurt and R. W. Pike, Theory and practice of simultaneous data reconciliation and gross error detection for chemical processes, *Comput. Chem. Eng.*, 2004, **28**, 381–402.
6 F. Madron, V. Veverka and V. Vanecek, Statistical-Analysis of Material Balance of a Chemical Reactor, *AIChE J.*, 1977, **23**, 482–486.
7 R. T. J. M. van der Heijden, B. Romein, J. J. Heijnen, C. Hellinga and K. C. A. M. Luyben, Linear Constraint Relations in Biochemical Reaction Systems 3. Sequential Application of Data Reconciliation for Sensitive Detection of Systematic-Errors, *Biotechnol. Bioeng.*, 1994, **44**, 781–791.
8 S. Narasimhan and C. Jordache, *Data reconciliation and gross error detection: an intelligent use of process data*, Annals of the New York Academy of Sciences, Gulf Publishing Company, Houston, Texas, US, 2000.
9 S. C. F. Meijer, R. N. A. van Kempen and K. J. Appeldoorn, Plant upgrade using big-data and reconciliation techniques, *Applications of Activated Sludge Models*, IWA publishing, 2015, pp. 357–410.
10 L. Rieger, I. Takacs, K. Villez, H. Siegrist, P. Lessard, P. A. Vanrolleghem and Y. Comeau, Data reconciliation for wastewater treatment plant simulation studies-planning for high-quality data and typical sources of errors, *Water Environ. Res.*, 2010, **82**, 426–433.
11 D. Rosso, L. E. Larson and M. K. Stenstrom, Aeration of large-scale municipal wastewater treatment plants: State of the art, *Water Sci. Technol.*, 2008, **57**, 973–978.
12 M. R. J. Daelman, E. M. van Voorthuizen, L. G. J. M. van Dongen, E. I. P. Volcke and M. C. M. van Loosdrecht, Methane and nitrous oxide emissions from municipal wastewater treatment – results from a long-term study, *Water Sci. Technol.*, 2013, **67**, 2350–2355.
13 K. E. Mampaey, M. K. De Kreuk, U. G. J. M. van Dongen, M. C. M. van Loosdrecht and E. I. P. Volcke, Identifying N_2O formation and emissions from a full-scale partial nitritation reactor, *Water Res.*, 2016, **88**, 575–585.
14 C. Hellinga, P. Vanrolleghem, M. C. M. van Loosdrecht and J. J. Heijnen, The potential of off-gas analyses for monitoring wastewater treatment plants, *Water Sci. Technol.*, 1996, **33**, 13–23.
15 S. Y. Leu, J. A. Libra and M. K. Stenstrom, Monitoring off-gas O_2/CO_2 to predict nitrification performance in activated sludge processes, *Water Res.*, 2010, **44**, 3434–3444.
16 C. Hellinga, A. A. J. C. Schellen, J. W. Mulder, M. C. M. van Loosdrecht and J. J. Heijnen, The SHARON process: An innovative method for nitrogen removal from ammonium-rich waste water, *Water Sci. Technol.*, 1998, **37**, 135–142.
17 J. W. Mulder, M. C. M. van Loosdrecht, C. Hellinga and R. van Kempen, Full-scale application of the SHARON process for treatment of rejection water of digested sludge dewatering, *Water Sci. Technol.*, 2001, **43**, 127–134.
18 K. E. Mampaey, U. G. J. M. van Dongen, M. C. M. van Loosdrecht and E. I. P. Volcke, Novel method for online monitoring of dissolved N_2O concentrations through a gas stripping device, *Environ. Technol.*, 2015, **36**, 1680–1690.
19 Q. H. Le, Mass-Balance-based Experimental Design and Data Reconciliation for Wastewater Treatment Processes, *PhD thesis*, Department of Green Chemistry and Technology, Ghent University, Belgium, 2019.
20 P. J. T. Verheijen, Data reconciliation and error detection, in *The Metabolic Pathway Engineering Handbook: Fundamentals*, CRC Press/Taylor & Francis, Boca Raton, 2010.
21 S. W. H. Van Hulle, H. J. P. Vandeweyer, B. D. Meesschaert, P. A. Vanrolleghem, P. Dejans and A. Dumoulin, Engineering aspects and practical application of autotrophic nitrogen removal from nitrogen rich streams, *Chem. Eng. J.*, 2010, **162**, 1–20.
22 U. van Dongen, M. S. M. Jetten and M. C. M. van Loosdrecht, The SHARON-Anammox process for treatment of ammonium rich wastewater, *Water Sci. Technol.*, 2001, **44**, 153–160.
23 R. van Kempen, J. W. Mulder, C. A. Uijterlinde and M. C. M. Loosdrecht, Overview: full scale experience of the SHARON process for treatment of rejection water of digested sludge dewatering, *Water Sci. Technol.*, 2001, **44**, 145–152.
24 E. I. P. Volcke, M. C. M. van Loosdrecht and P. A. Vanrolleghem, Controlling the nitrite: Ammonium ratio in a SHARON reactor in view of its coupling with an Anammox process, *Water Sci. Technol.*, 2006, **53**, 45–54.
25 J. Desloover, H. De Clippeleir, P. Boeckx, G. Du Laing, J. Colsen, W. Verstraete and S. E. Vlaeminck, Floc-based sequential partial nitritation and anammox at full scale with contrasting N_2O emissions, *Water Res.*, 2011, **45**, 2811–2821.

26 M. Strous, J. J. Heijnen, J. G. Kuenen and M. S. M. Jetten, The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms, *Appl. Microbiol. Biotechnol.*, 1998, **50**, 589–596.
27 T. Lotti, R. Kleerebezem, C. Lubello and M. C. M. van Loosdrecht, Physiological and kinetic characterization of a suspended cell anammox culture, *Water Res.*, 2014, **60**, 1–14.
28 S. Guo, P. Liu and Z. Li, Estimation of exhaust steam enthalpy and steam wetness fraction for steam turbines based on data reconciliation with characteristic constraints, *Comput. Chem. Eng.*, 2016, **93**, 25–35.
29 X. Jiang, P. Liu and Z. Li, A data reconciliation based framework for integrated sensor and equipment performance monitoring in power plants, *Appl. Energy*, 2014, **134**, 270–282.