S.C.E.N.A

Anaerobic supernatant

N₂O

Raw data

N₂O

Knowledge discovery and data-mining

Cycle 1

An Aer Anox

N₂O end of anoxic > 0.6 mg/L

Prediction of initial aerobic N₂O accumulation

Cycle 2

An Aer Anox

N₂O prediction

Knowledge discovery and data-mining

Raw data

N₂O
A knowledge discovery framework to predict the N\textsubscript{2}O emissions in the wastewater sector

V. Vasilaki\textsuperscript{1}, V. Conca\textsuperscript{2}, N. Frison\textsuperscript{2}, A.L. Eusebi\textsuperscript{3}, F. Fatone\textsuperscript{3}, E. Katsou\textsuperscript{1*}

\textsuperscript{1}Department of Civil & Environmental Engineering, Brunel University London, Uxbridge UB8 3PH, UK,  
\textsuperscript{2}Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy,  
\textsuperscript{3}Department SIMAU, Faculty of Engineering, Polytechnic University of Marche, Via Brecce Bianche 12, Ancona, Italy  

*Corresponding author. Email: evina.katsou@brunel.ac.uk

Keywords: short-cut enhanced nutrients abatement - SCENA; long-term dissolved N\textsubscript{2}O and energy consumption monitoring, knowledge discovery and data mining
Abstract

Data Analytics is being deployed to predict the dissolved nitrous oxide (N$_2$O) concentration in a full-scale sidestream sequence batch reactor (SBR) treating the anaerobic supernatant. On average, the N$_2$O emissions are equal to 7.6% of the NH$_4$-N load and can contribute up to 97% to the operational carbon footprint of the studied nitritation-denitritation and via-nitrite enhanced biological phosphorus removal process (S.C.E.N.A). The analysis showed that average aerobic dissolved N$_2$O concentration could significantly vary under similar influent loads, dissolved oxygen (DO), pH and removal efficiencies. A combination of density-based clustering, support vector machine (SVM), and support vector regression (SVR) models were deployed to estimate the dissolved N$_2$O concentration and behaviour in the different phases of the SBR system.

The results of the study reveal that the aerobic dissolved N$_2$O concentration is correlated with the drop of average aerobic conductivity rate (spearman correlation coefficient equal to 0.7), the DO (spearman correlation coefficient equal to -0.7) and the changes of conductivity between sequential cycles. Additionally, operational conditions resulting in low aerobic N$_2$O accumulation (<0.6 mg/L) were identified; step-feeding, control of initial NH$_4^+$ concentrations and aeration duration can mitigate the N$_2$O peaks observed in the system. The N$_2$O emissions during aeration shows correlation with the stripping of accumulated N$_2$O from the previous anoxic cycle. The analysis shows that N$_2$O is always consumed after the depletion of NO$_2^-$ during denitritation (after the “nitrite knee”). Based on these findings SVM classifiers were constructed to predict whether dissolved N$_2$O will be consumed during the anoxic and anaerobic phases and SVR models were trained to predict the N$_2$O concentration at the end of the anaerobic phase and the average dissolved N$_2$O concentration during aeration. The proposed approach accurately predicts the N$_2$O emissions as a latent parameter from other low-cost sensors that are traditionally deployed in biological batch processes.
1. Introduction

In recent years the sustainability and operational efficiency of wastewater treatment plants (WWTPs) have come to the fore (Liu et al., 2018). Several biological technologies such as partial-nitritation – anammox (anaerobic ammonium oxidation) have emerged, towards the efficient, low-cost treatment of high-strength municipal wastewater streams (Lackner et al., 2014; Zhou et al., 2018). The anaerobic supernatant is a by-product of dewatering of the anaerobic digestion effluent and represents less than 1-2% of the total influent flow in the WWTP. It contains 10–30% of the N load and 20–30% of the P load (Janus and van der Roest, 1997; van Loosdrecht and Salem, 2006). Sidestream treatment of the anaerobic supernatant can contribute to the reduction of energy consumption for N-removal, decrease of nitrogen loads in the secondary treatment, and the minimisation of risks related to exceeding effluent regulatory requirements of nitrogen concentrations in the water line of WWTPs (Eskicioglu et al., 2018). However, the performance and environmental evaluation of different sidestream technologies is still under investigation (Eskicioglu et al., 2018; Rodriguez-Garcia et al., 2014).

SCENA (Short-Cut Enhanced Nutrient Abatement) is a new sidestream process, that combines the conversion of \( \text{NH}_4^+ \) to \( \text{NO}_2^- \) under aerobic conditions (nitritation) with the subsequent reduction of \( \text{NO}_2^- \) to nitrogen gas and enhanced biological phosphorus uptake by polyphosphate-accumulating organisms (DPAOs) in a sequencing batch reactor (SBR) (Frison et al., 2015). External volatile fatty acids (VFAs), are produced via acidogenic fermentation of the primary and secondary sludge on-site and dosed into the SBR. In a recent study, Longo et al. (2016), quantified the environmental and cost benefits and impacts of the integration of the SCENA process in a full-scale WWTP. They reported major energy savings for aeration after the integration of sidestream SCENA process. The direct \( \text{N}_2\text{O} \) emissions were equal to 1.42% of the
influent N-load. Short-term monitoring campaigns were implemented, while the effect of operational conditions on N\textsubscript{2}O generation was not investigated.

N\textsubscript{2}O is a potent cause of global warming, its global warming potential is 265 - 298 times more than that of CO\textsubscript{2} (IPCC, 2013). The emission of N\textsubscript{2}O in full-scale sidestream partial-nitrification/partial-nitritation–anammox or nitrification-denitrification systems range from 0.17% to 5.1% of the influent N-load (average equal to ~2.1% of the N-load is emitted (Vasilaki et al., 2019). Schaubroeck et al. (2015) showed that N\textsubscript{2}O emissions from a full-scale sidestream DEMON process in Austria were significantly higher than the direct N\textsubscript{2}O emissions from the mainstream treatment in a full-scale WWTP. On average, 0.256 g N\textsubscript{2}O were emitted compared to 0.005 g emitted in the secondary treatment per m\textsuperscript{3} treated wastewater. The increased direct N\textsubscript{2}O emissions can be mainly attributed to low DO concentrations, higher ammonia oxidation rates (AOR) and NO\textsubscript{2}\textsuperscript{-} build-up (Desloover et al., 2011; Kampschreur et al., 2008); conditions that also prevail in the SCENA process. The variability of EF reported in sidestream technologies can be partially attributed to both complex relationships between emitted N\textsubscript{2}O and operational conditions and different configurations (i.e. SBR, continuous systems), loads (i.e. NH\textsubscript{4}\textsuperscript{+} concentrations), feeding strategies and operational control (i.e. DO set-points). Additionally, different interactions between operational variables trigger a different response of N\textsubscript{2}O generation. For instance, in a recent modelling study of a granular one-stage partial-nitritation-anammox reactor, Wan et al. (2019) showed that higher temperatures resulted in increased N\textsubscript{2}O emissions in the presence of COD (chemical oxygen demand) and in decreased N\textsubscript{2}O emissions in the absence of COD (due to increased anammox activity and reduction of NO\textsubscript{2}\textsuperscript{-} accumulation in higher temperature). Additionally, the long-term temporal variations of direct N\textsubscript{2}O emissions were not adequately assessed in sidestream technologies; the majority of the monitoring campaigns in sidestream reactors lasted less than 5 days (Vasilaki et al., 2019).
The digitalisation of water services and the data-driven knowledge discovery from wastewater treatment plant may increase the resilience of water utilities under climate change and other water-related challenges (Sarni et al., 2019). Recent studies have provided extensive overviews of the use of data-driven techniques in the wastewater sector for different applications including the development of soft-sensors, fault prediction and multi-objective optimisation of control of water utilities (Corominas et al., 2018a; Haimi et al., 2013; Newhart et al., 2019). Data-mining and extraction of the information hidden in the raw sensor signals can facilitate the identification of patterns and hidden structures and reveal significant information on the behaviour of N$_2$O emissions in continuous wastewater treatment processes (Vasilaki et al., 2018). The SBR in the SCENA process is multiphase (i.e. anaerobic, aerobic, anoxic conditions) applying different operational variables (unsynchronised data), non-linear and subject to different disturbances, such as influent compositions and fermentation liquid characteristics. Moreover, SBR process data are based on a 3d-structure that consists of the number of i) cycles, ii) variables and iii) sampling points within each cycle. Therefore, the identification of process abnormalities and patterns can be complicated. N$_2$O emissions could be affected by both within-cycle and between-cycle batch dynamics.

In this study, sensor and laboratory analyses data from a full-scale SCENA SBR were analysed to provide insights on the N$_2$O emissions behaviour and generation. A structured approach was followed for knowledge discovery from the available dataset using a combination of abnormal events detection, classification and regression techniques. The objectives of the study were to i) investigate whether the sensors integrated in the system (i.e. conductivity, pH) can provide actionable information on the dynamics of N$_2$O emissions, ii) detect hotspots for the accumulation and emission of N$_2$O and iii) develop data-driven regression and classification
models to predict the dissolved N$_2$O behaviour and concentration for the different phases (anaerobic, aerobic, anoxic) of the SBR.

2. Materials and Methods

2.1 Process description and data origin

The Carbonera plant is designed to treat domestic wastewater of a population equivalent of 40,000 (dry weather flow equal to 10,000 m$^3$/d). After screening and degritting and primary sedimentation, the effluent from the primary clarifier is sent to a Schreiber reactor (single basin – working volume 4671 m$^3$). Schreiber reactor effluent is pumped to two secondary clarifiers (2260 m$^3$ each) and subsequently to the tertiary treatment unit for disinfection and filtration before final discharge in the Melma River.

Waste activated sludge (WAS) generated by the biological treatment is recycled to the primary sedimentation unit and mixed with primary sludge. The final concentration of the thickened mixed sludge is around 5% total solids (TS). About 90% of the mixed thickened sludge is fed to an anaerobic digestion unit (1800 m$^3$ working volume). Digestate is dewatered by a centrifuge with the addition of polyelectrolyte; the solid fraction is mechanically composted and used as agricultural fertilizer. The anaerobic supernatant is sent to the equalization tank (of 90 m$^3$) in the SCENA system for the biological N and P removal.

The remaining portion of mixed sludge (10%) is fed to a sequencing batch fermentation reactor (SBFR) with hydraulic retention time (HRT) equal to 5 days. The SBFR is operated under mesophilic condition (37°C) for the alkaline fermentation of thickened sewage sludge and the on-site production of carbon source enriched of VFAs (mainly acetic and propionic acids). Daily, 10 m$^3$ of fermentation sludge are extracted and replaced with fresh thickened sludge. The solid/liquid separation of the fermented sewage sludge is carried out by a screw-press (SCAE), generating ~2-
4 m³/h of fermentation liquid rich of VFAs (in total, ~ 10.5 m³/d). The latter is collected in a storage tank of 20 m³ and automatically dosed during the anaerobic and anoxic phases of a short-cut sequencing batch reactor (SBR) based on pH and conductivity sensors. The solid fermented fraction (13-15% TS based) is mixed with the thickened mixed sludge and fed to the anaerobic digestor.

The anaerobic supernatant is treated in an SBR with a maximal working volume of 70 m³ (3-4 cycles daily). The SBR is fed with ~10-15 m³ of anaerobic supernatant in each cycle that is treated via nitrite enhanced phosphorus removal associated with nitritation-denitritation (S.C.E.N.A process). The typical SBR (Figure 1) cycle consists of feeding (6-8 min), anaerobic conditions (30 min), aerobic conditions (200-240 min), anoxic (~60-140 min), settling (30 min) and discharge (8 min). The sensors integrated in the SBR include: pH, Dissolved Oxygen (DO), conductivity, Oxidation Reduction Potential (ORP), mixed liquor suspended Solids (MLSS) and temperature. Conductivity and pH are used to control the length of the aerobic and anoxic phases and the carbon source dosage. Additionally, variable frequency driver is used to control the air flow-rate of the blowers, maintaining the dissolved oxygen during aerobic phase in the range of 1.0 to 1.5 mg/L. The aeration system consists of volumetric blowers (nominal power 11 kW) and n80 diffusers (INVENT), providing ~500 m³/h of compressed air at 400 mbar of pressure. The treated supernatant is recirculated back to the WWTP headworks.

A monitoring campaign was conducted in the sidestream line at Carbonera WWTP treatment plant for approximately 4 months (January 2019 – April 2019). Dissolved N₂O concentrations were measured using a polarographic Clark-type electrode (Unisense, Aarhus, Denmark). To supplement the long-term monitoring campaign with Unisense probes, N₂O emissions in the headspace of the SBR reactor, were also continuously monitored with MIR9000CLD analyser (Environment Italia S.p.A.) during March – April 2019. Details of the monitoring campaign, N₂O
emissions’ calculation and laboratory analyses are provided in the supplementary material (S1-S3).

2.2 Data analysis

2.2.1 Methodological Framework

Figure 2 summarises the methodological framework of the study. Phase one includes preliminary analysis of the collected data. Features extraction and density-based clustering was applied (Ester et al., 1996), to isolate abnormal cycles. The methodology and results of abnormal cycles’ isolation are given in the supplementary material (section S4). In phase two, the behaviour of N₂O emissions and dissolved N₂O concentration during normal operation was investigated; efforts were focused to identify dependencies with the operational dataset and laboratory analyses. Finally, in phase three, classification and regression models were trained to predict the behaviour of aerobic dissolved N₂O concentration in the different cycles. A support vector machine classification (SVM) and regression (SVR) models were constructed (Cortes and Vapnik, 1995).

The first step for the prediction of the average aerobic dissolved N₂O concentration included the training of an SVM classifier (ANOXSVM) to predict whether dissolved N₂O will be consumed during the anoxic phase. This was significant, given that accumulated dissolved N₂O in the beginning of the aerobic phase, will be stripped during aeration. All cycles were divided in two classes: class anoxA (dissolved N₂O < 0.6 mg/L) and class anoxB (dissolved N₂O > 0.6 mg/L). The dissolved N₂O concentration threshold was set equal to 0.6 mg/L, since in ~88% of these cases, N₂O was consumed by the end of subsequent anaerobic phase. In cycles belonging to class anoxA, no N₂O carryover was assumed. It is important to note that the term anaerobic phase, is
used to describe the first operational phase of the SBR (Figure 1) within each cycle and is not necessarily representative of the actual conditions in the reactor.

Subsequently, an SVM classifier (ANSVM) was trained to predict if dissolved N\textsubscript{2}O will be consumed in the subsequent anaerobic phase. The threshold of N\textsubscript{2}O at the end of the anaerobic phase was set equal to 2.6 mg/L (sensor calibration limit). Therefore, anaerobic phases with accumulated N\textsubscript{2}O were classified in two groups: class anaerA (N\textsubscript{2}O concentration < 2.6 mg/L) and anaerB (N\textsubscript{2}O concentration > 2.6 mg/L). Cycles belonging to anaerA class, were used to train an SVR model (ANSVR) to predict the dissolved N\textsubscript{2}O concentration at the end of the anaerobic phase.

Finally, an SVR model was trained to predict the average N\textsubscript{2}O concentration during the aerobic phase (AERSVR), utilizing the ANSVR model predictions for cycles with initial aerobic N\textsubscript{2}O less than 2.6 mg/L. Finally, the aerobic SVR model was also tested to cycles belonging in class anaerB (N\textsubscript{2}O concentration > 2.6 mg/L). In anaerB cycles, initial aerobic N\textsubscript{2}O accumulation exceeds the calibration limit of the sensor. Additionally, aerobic N\textsubscript{2}O accumulation starts before completion of the stripping of pre-existing dissolved N\textsubscript{2}O. In these cases, the average dissolved N\textsubscript{2}O concentration of the cycle, was calculated considering the period from the first minimum of dissolved N\textsubscript{2}O concentration until the end of aeration (or after 30 min if a local minimum did not exist). Additionally, initial N\textsubscript{2}O accumulation was assumed to be equal to 0.6 mg/L (average minimum after initial N\textsubscript{2}O stripping observed in these cycles).

In practice, the methodology followed was not linear as it is illustrated in Figure 2; it involves several backward and forward loops between the different steps. The feedback loops were necessary to leverage the knowledge discovered and adjust the data-preparation (i.e. new features extraction, different pre-processing) and mining phases.
Support vector machines (SVMs) are a range of supervised non-parametric classification and regression algorithms that have various applications in several fields including hydrology (Raghavendra and Deka, 2014), bioinformatics (Byvatov and Schneider, 2003) and wastewater (Corominas et al., 2018). For instance, in wastewater, support vector regression (SVR) has been successfully applied to data generated from mechanistic modelling of biological processes (Fang et al., 2011) or to experimental data (Seshan et al., 2014) to predict reactors’ performance.

SVM classification and SVR models were constructed to predict the behaviour of dissolved N\textsubscript{2}O production/consumption in different phases of the SBR operation (Figure 2). SVM aims to define an optimum separating hyperplane in the feature space that maximizes the margin between two different classes. Classes with large margins are clearly separable and provide a ‘safety’ for the generalisation of the algorithm when applied to new points. In practical applications, the overlapping of a number of data belonging to the two classes, is common. Therefore, soft margins are introduced to allow a number of misclassifications to identify feasible solutions when the training dataset is not strictly linearly separable. Similarly, in the SVR case, the aim of the method is to identify the hyperplane that has the minimum distance to all data points. A complete description of the SVM and SVR algorithms is provided in the supplementary material. Radial basis function (RBF) was selected to construct the models in this study. The ‘kernel trick’ and enables SVMs to operate even in infinite feature space (where data are mapped), without in practice executing calculations there (Luts et al., 2010).

The algorithms were implemented with the kernlab package (Karatzoglou et al., 2004) in R software. Repeated 10-fold cross validation (3 repetitions) was applied to select the cost and
gamma (γ) regularization parameters over a grid-search with the caret package (Kuhn, 2008). The cost determines the penalty of misclassified instances or instances violating the maximal margin whereas γ determines the amplitude of the kernel. The dataset was randomly divided into test and train, with 70% of the available data used for training the SVM model and 30% used for testing.

In the classification case, over-sampling was applied for the minority classes within the 10-fold cross validation loop (before training). Local models were developed based on observations from each phase of the SBR reactor instead of the dataset from the duration of the whole cycle. The underlying characteristics and dependencies of the operational variables vary between anoxic, aerobic and anaerobic conditions. Additionally, the performance of the system under different phases within the cycle can also vary. There are significant benefits in the development of local phase-based models. The behaviour of dissolved N$_2$O and triggering operational conditions vary between the different phases; local models enable to investigate the phase-based dependency structures that would not be possible using the whole cycle dataset. The performance of the classification SVM models were evaluated based on accuracy and kappa and from the sensitivity and specificity as described in the supplementary material (S3.1). Similarly, the regression models were evaluated considering the root mean squared error (RMSE) and R-squared ($R^2$) (S3.1).

3. Results and discussion

3.1 SCENA performance

The SBR treats up to 43 kg of N/day of anaerobic supernatant, which results in a volumetric nitrogen loading rate up to 0.78 kgN/m$^3$ day. The performance of the SBR reactor in terms of NH$_4$-N removal, was stable during the monitoring campaign. During system’s normal operation (January 2019 - April 2019), the average removal efficiency of NH$_4$-N, TN and PO$_4$-P was 78%.
~77% and 84% respectively. Influent and effluent concentrations of the SCENA system for the duration of the monitoring campaign are provided in Table 1. A detailed description of the abnormal cycles isolated is provided in the supplementary material.

[Table 1]

### 3.2 N₂O Emission factor

N₂O emissions were measured using a gas analyser (March – April 2019); on average ~0.8 kg of N₂O-N was emitted in each cycle, equivalent to 7.6% of the NH₄-N load in the SBRR. In terms of the NH₄-N removed the N₂O EF was equal to 11% (±4). The emissions during the aerobic phase were considered. N₂O emissions exhibited significant variability ranging from 0.14 kg N₂O-N/cycle (1.3% of NH₄-N load) to ~2 kg N₂O-N/cycle (19% of NH₄-N load) as shown in Figure 3 (a). Emission peaks higher than 1.5 kg N₂O-N/cycle and the increasing trend observed close to the end of the monitoring campaign coincide with peaks in the conductivity change in the aerobic phase of the cycles (Figure 3 (b)). Laboratory analyses performed approximately four times per week, did not demonstrate any significant changes in the influent COD, NH₄-N loads and removal efficiencies linked with the increasing trend of the emissions observed in Figure 3 (a). Given the wide range of the N₂O emissions observed in the system, in the following sections, efforts were focused to identify triggering operational conditions.

[Figure 3]

### 3.3 Energy consumption vs N₂O emissions

The operational carbon footprint of the sidestream line was estimated using the direct GHG emissions (from N₂O) and electricity consumption. The electricity consumption was relatively steady over the monitoring period; on average ~5.4 kWh was consumed in the SBR for the removal of 1 kg of NH₄-N from the anaerobic supernatant. The average energy consumption of
the SBR represented ~77% of the total electricity consumption of the SCENA system. On average ~48.7 kg of CO$_{2eq}$ are generated for the removal of 1 kg of NH$_4$-N due to the direct N$_2$O emissions and electricity consumption in the system. The contribution of the total N$_2$O emissions to the operational carbon footprint of the S.C.E.N.A process ranged from 66.7% to 96.8% when all the equipment electricity consumption (i.e. fermenter, dynamic thickener) were considered. Given the variability of the N$_2$O emissions observed in the system (Figure 3) the kg of CO$_{2eq}$ emitted per kg of NH$_4$-N removed ranged between 9.5 kg CO$_{2eq}$ to 117.7 kg CO$_{2eq}$. Figure 4 (a), shows the average operational carbon footprint (considering direct N$_2$O emissions and electricity consumption) of the SCENA system for two cases with different ranges of N$_2$O emissions. In the first case (26/03), a considerable amount of N$_2$O was emitted, equal to ~10.5% of the influent NH$_4$-N load. In the second case, the emissions were significantly lower, equal to ~4% of the influent NH$_4$-N load. Both cases are characterised by similar influent NH$_4$-N concentrations, phase duration, temperature and ammonia removal efficiencies (~79%). The DO concentration is equal to ~1 mg/L. In case 1, the operational carbon footprint of the process is ~136% higher compared to case 2. This example shows that under similar conditions (considering laboratory analyses, average pH and DO), dissolved N$_2$O concentrations can vary significantly in the studied system. Investigation of the behaviour of conductivity during the two aerobic phases, showed higher conductivity and pH decrease in case one (~ 510 μS/cm and ~1 respectively) compared to case two (~350 μS/cm and 0.7 respectively) (Figure 4 (b) and (c)). Additionally, the initial aerobic ORP in case 2, was higher (~43 mV) compared to case 1 (~274 mV) (Figure 4 (b)). Therefore, efforts to understand the N$_2$O triggering operational conditions and mitigate GHG emissions, should consider the dynamic in-cycle behaviour of the variables monitored in the system. The relationship between the operational variables (i.e. DO, NH$_4$-N concentration, ORP, conductivity) will be discussed in the following sections.
3.4 Variability of N$_2$O emissions during normal operation

N$_2$O was emitted during aeration phase in all cycles and correlated significantly with the dissolved N$_2$O accumulation. One representative cycle profile for the dissolved N$_2$O concentration and N$_2$O emissions in cycles starting without dissolved N$_2$O accumulation from the previous cycle is shown in Figure 5, together with the DO, NH$_4$-N, conductivity, ORP and pH.

ORP at the beginning of the aerobic phase shows a correlation with the DO, whereas N$_2$O accumulation is minimum. Dissolved N$_2$O increases in the first 60-70 min of aeration (a small change in the pH slope can be seen coinciding with the peak of accumulated N$_2$O) indicating that the generated N$_2$O generation is higher than the stripped N$_2$O. N$_2$O accumulation shows a decreasing trend after ~90 minutes of aeration. Subsequently dissolved N$_2$O concentration increases when aeration stops, and the anoxic phase starts. This shows that production of N$_2$O continues under decreasing DO and until DO depletion. The calibration range of the dissolved N$_2$O probe is between 0 - 2.6 mg/L. Therefore, the accumulation of dissolved N$_2$O can be higher than the peak shown in Figure 5. During the anoxic phase, pH increases rapidly during the dosage of fermentation liquid, followed by a slow decrease upon the end of carbon dosage phase. A sudden change in the ORP signal slope (‘nitrite knee’) indicates the depletion of nitrite whereas TN still exists in the form of N$_2$O. Accumulated N$_2$O is subsequently depleted rapidly after NO$_2^-$ depletion.
3.5 The pattern of N$_2$O emissions

Offline data from laboratory studies and the ranges of the operational variables were analysed in order to investigate significant changes that contribute to high accumulation of dissolved N$_2$O concentration and high N$_2$O emissions. Figure 6 (a) shows the daily average dissolved N$_2$O concentration (coloured points) during aerobic phase versus conductivity at the end of aerobic phase and the effluent NH$_4$-N concentration. Conductivity is significantly related and can be linked with the NH$_4$-N concentration in the reactor (spearman correlation coefficient equal to 0.97). High average aerobic dissolved N$_2$O concentration (>1.5 mg/L) was mainly observed with NH$_4$-N concentrations lower than 150 mg/L and higher than 300 mg/L in the effluent of the SBR. Additionally, the spearman correlation coefficient between dissolved N$_2$O and average aerobic conductivity decrease rate (μS/cm/min) was equal to -0.7 and N$_2$O concentration peaks were observed for conductivity decrease rate > 1.8 μS/cm/min. The latter indicates that higher emissions occur under high ammonia removal efficiency that can be linked with higher ammonia oxidation rates (AOR) (i.e. due to pH values observed ~ 8) triggering the NH$_2$OH oxidation pathway or higher than average NO$_2$-N accumulation (triggering nitrifier denitrification pathway). Domingo-Félez et al., (2014) found that N$_2$O production rates were positively correlated with the extant nitrification rate in a single-stage nitritation/Anammox reactor. Similarly, Law et al. (2011) identified a linear relationship between AOR and N$_2$O emissions in a partial nitritation SBR reactor treating the reject water from anaerobic digestion. Law et al. (2011) suggested that is attributed to higher accumulation of the ammonium oxidation intermediates (hydroxylamine (NH$_2$OH) and nitrosyl radical (NOH)) leading to faster N$_2$O formation or to the increased use of electrons reducing nitrite to nitric oxide (nitrifier denitrification pathway) under low DO concentrations. High nitrite accumulation has been also linked with elevated N$_2$O
emissions and the nitrifier denitrification pathway, especially under low DO concentrations (Tallec et al., 2006; Kampschreur et al., 2008; Desloover et al., 2011; Peng et al., 2015; Massara et al., 2017; Law et al., 2012). For instance, Peng et al. (2017) and Kampschreur et al. (2009), in a nitritation-denitrification SBR and a full-scale single stage nitritation-Anammox reactor respectively, identified linear relationship between nitrite accumulation and \( \text{N}_2 \text{O} \) emissions at DO levels below 1.5 mg/L. Similarly, Tallec et al (2006) in a nitrifying activated sludge observed eightfold increase of \( \text{N}_2 \text{O} \) emissions with the addition of nitrite pulses (10 mg/L) at DO equal to 1 mg/L. Therefore, both hydroxylamine oxidation and the nitrifier denitrification are possible during aeration in the investigated SBR.

The average dissolved \( \text{N}_2 \text{O} \) concentration during the aerobic phase of different cycles varied significantly in relation to the average DO concentration. Figure 6 (b), shows that the dissolved \( \text{N}_2 \text{O} \) concentration peaks coincided with average DO concentrations less than 0.9 to 1 mg/L. The spearman correlation coefficient between dissolved \( \text{N}_2 \text{O} \) and DO concentrations was equal to -0.7. The coloured points in the Figure, represent the ORP at the end of the aerobic phase; ORP is higher than 40 mV in the majority of the cycles with average aerobic dissolved \( \text{N}_2 \text{O} \) concentration less than 1 mg/L. Only cycles without dissolved \( \text{N}_2 \text{O} \) accumulation from the previous anoxic phase are shown in the graph. Stenström et al. (2014) showed decreasing DO concentrations lower than 1–1.5 mg/L are linked with higher nitrite accumulation and are positively correlated with \( \text{N}_2 \text{O} \) emissions during nitrification in a full-scale predenitrification-nitrification SBR treating anaerobic supernatant. Similarly, Pijuan et al., (2014) reported an increase of \( \text{N}_2 \text{O} \) emissions in a nitritation reactor with the reduction of DO from 4 to <1 mg/L. During the monitoring period, blowers operated at maximum flow-rate. Therefore, the presence of residual biodegradable COD concentration in the aerobic, is expected to decrease DO concentration. Similarly, higher influent \( \text{NH}_4^+ \) loads or higher ammonia oxidation rates (that can
also result in increased NO\textsubscript{2} accumulation) can impact the DO concentration in the system. The dissolved N\textsubscript{2}O concentration can be affected by a combination of variables; therefore, it cannot be deduced that the decreased DO is the sole contributing factor triggering the increased N\textsubscript{2}O generation observed.

[Figure 6]

3.6 Impact of accumulated N\textsubscript{2}O in the end of anoxic and anaerobic phase

Several parameters have been reported to affect the N\textsubscript{2}O accumulation under anoxic conditions, such as the inhibition of the nitrous oxide reductase (Nos) by free nitrous acid (FNA) or high accumulation of NO\textsubscript{2}, the electron competition between electron acceptors and the type of carbon source (Itokawa et al., 2001; Pan et al., 2013; Zhou et al., 2008; Zhu and Chen, 2011). Additionally, low values of COD/N can result in incomplete denitritation and therefore, N\textsubscript{2}O accumulation via the heterotrophic denitrification pathway during the anoxic phase of the SBR. Accumulated N\textsubscript{2}O in the end of the anoxic phase is stripped in the subsequent cycle, increasing the N\textsubscript{2}O emissions. Caranto et al. (2016) have recently showed that N\textsubscript{2}O can be the main product of anaerobic NH\textsubscript{2}OH oxidation catalysed by the cytochrome P460 in N. europaea. The latter can be an evidence of the biological N\textsubscript{2}O generation under limited DO and high NH\textsubscript{3} concentrations, both conditions occurring in the target system in the during the transition from aerobic to anoxic phases when N\textsubscript{2}O accumulation rapidly increases.

In this study, the average soluble COD concentration in the fermentation liquid was equal to 13082 mg COD/L over the monitoring period (Table 1). Overall, in >27% of the examined cycles the N\textsubscript{2}O was completely consumed by the end of the anoxic phase. Zhu and Chen, (2011), showed that the use of sludge alkaline fermentation as carbon source in an anaerobic-aerobic system treating high-strength stream, can reduce the N\textsubscript{2}O production by up to 68.7% compared to
alternative carbon sources (i.e. acetic acid). On the other hand, Li et al., (2013a) in a process utilizing PHA as internal carbon source, observed higher N$_2$O production and reduction rates at higher influent COD concentrations linked with higher anaerobic PHA synthesis (ranging from 100 to 500 mg/L). The higher N$_2$O production rates were attributed to the accumulated NO$_2^-$ inhibiting the N$_2$O reduction.

The dissolved N$_2$O concentration in the anoxic phase exceeded the calibration limit of the sensors; only cycles in which “nitrite knee” was observed and N$_2$O reduced to values lower than 2.6 mg/L could be investigated. Therefore, the effect on NO$_2^-$ in anoxic N$_2$O generation could not be studied. However, studies have shown that elevated NO$_2^-$ concentrations during denitrification can reduce the denitrification rate and increase the N$_2$O accumulation (Schulthess et al., 1995). The electron competition between nitrite reductase NIR, nitric oxide reductase (NOR) and nitrous oxide reductase (NOS) is intensified under high NO$_2^-$ concentrations; NOS is less competitive under limitation of electron donor and this will result in N$_2$O accumulation (Pan et al., 2013; Ren et al., 2019).

Based on the profiles shown in Figure 5, N$_2$O was always consumed after the depletion of NO$_2^-$ during denitrification; specifically, dissolved N$_2$O concentration decreased after the “nitrite knee”. Gabarró et al. (2014), studied a partial-nitritation reactor treating landfill leachate, and operated under alternating aerobic/anoxic conditions to allow heterotrophic denitrification. The authors demonstrated that significant N$_2$O accumulation was observed during anoxic periods. NO$_2^-$ denitrification rate was higher under both biodegradable COD limiting conditions and after acetate addition compared to N$_2$O reduction; N$_2$O reduction rate was maximum after NO$_2^-$ removal (similar to what was observed in this study). In denitrifying phosphorus removal processes, Li et al. (2013) showed that the N$_2$O accumulation can be higher compared to conventional denitrification; the authors suggested that in the electron competition between
denitrifying enzymes and PHA, \( \text{N}_2\text{O} \) reductase is less competitive. On the other hand, Ribera-Guardia et al. (2016) investigated the electron competition during denitrification (PHA as the sole carbon source) of enriched dPAO and dGAO biomass and found that higher \( \text{N}_2\text{O} \) accumulation in the latter culture. Additionally, the last step of denitrification was inhibited in dGAO cultures (\( \text{N}_2\text{O} \) accumulation up to \( \sim84\% \) of the N-reduced), under high levels of \( \text{NO}_2^- \) (\( \sim15\text{mgN/gVSS} \)).

Wang et al., (2015) demonstrated that during denitrifying phosphorus removal, mitigation of \( \text{NO}_2^- \) accumulation is possible via continuous dosage of phosphate and nitrate. Wang et al., (2011), showed that optimisation of the synthesis of PHA during the anaerobic phase can mitigate the \( \text{N}_2\text{O} \) production during the anoxic phase leading to complete denitrification.

In the system, \( \text{N}_2\text{O} \) emissions and dissolved \( \text{N}_2\text{O} \) concentration at the aerobic phase is strongly related with incomplete denitrification in the previous cycle. In \( \sim26\% \) of the cycles with incomplete denitrification, the \( \text{N}_2\text{O} \) concentration did not decrease below \( \sim2\text{mg/L} \) in the anaerobic phase and therefore the stripping of accumulated \( \text{N}_2\text{O} \) in the subsequent aerobic phase was substantial.

Figure 7 (a) shows representative profiles of the dissolved \( \text{N}_2\text{O} \) concentration and the \( \text{N}_2\text{O} \) emissions based on different initial concentrations of \( \text{N}_2\text{O} \) in the beginning of the aerobic phase. The profiles of the ORP, DO and pH are comparable in the preset cycles (Figure 7 (b)). In cycle B \( \sim0.56\text{kgN} \) of \( \text{N}_2\text{O} \) were emitted during the aerobic phase, whereas in cycle A \( \text{N}_2\text{O} \) emissions are equal to \( 0.33\text{kgN} \) (given the duration of these cycles is not equal only 220 min were considered). The initial dissolved \( \text{N}_2\text{O} \) concentration in cycles A and B is equal to 0.27 and \( >2.6\text{mg/L} \) respectively. The \( \text{N}_2\text{O} \) emissions increased significantly due to the accumulated \( \text{N}_2\text{O} \) at the beginning of the previous anoxic phase that was stripped at the beginning of aeration.

Overall, in \( \sim72\% \) of the cycles, the dissolved \( \text{N}_2\text{O} \) concentration at the beginning of the anaerobic phase was higher than 0.3 mg/L. In cycles with dissolved \( \text{N}_2\text{O} \) concentration higher than 0.3
mg/L at the beginning of the anaerobic phase, the change in dissolved N$_2$O concentration during the anaerobic phase was highly correlated with the ORP at the beginning of the anaerobic phase. Additionally, the spearman correlation coefficient between the magnitude of the ORP reduction and magnitude of the dissolved N$_2$O reduction was equal to 0.7. Figure 8 shows the boxplots of dissolved N$_2$O reduction in relation to initial anaerobic ORP and ORP change for two cases: i) negligible dissolved N$_2$O change mainly due to influent dilution or anaerobic dissolved N$_2$O concentration > 2.6 mg/L, and ii) occasions with N$_2$O reduction during the anaerobic phase. In Figure 8 (a) only occasions with ORP decrease higher than -50 mV are shown. The presence of nitrites in the bulk liquid during the (anaerobic) phase affected the ORP. NO$_2$-N depletion in the bulk liquid resulted in a sharp “nitrite knee” in the ORP profile (similar to the one observed during the anoxic phase. Therefore, higher ORP change was expected in cycles with NO$_2$-N depletion and N$_2$O consumption during the anaerobic phase.

Anaerobic phase term, is used to describe the first operational phase of the SBR (Figure 1) within each cycle and might not represent the actual conditions in the reactor. For instance, ORP ~ -80 mV in the anaerobic phase of the SBR indicates anoxic conditions, due to residual NO$_2$N concentration from the previous anoxic phase of the reactor.

3.7 Prediction and control of N$_2$O accumulation in the anoxic and anaerobic phases

As discussed in section 3.6, the behaviour of ORP was significantly related with the behaviour of NO$_2^-$ and consequentially of the dissolved N$_2$O concentration during the anaerobic phase. Therefore, in the ANSVM model, features related with the ORP profile were mainly used (Table
2). Similarly, there was a strong link with the ORP behaviour and the nitrite “knee” with the N\textsubscript{2}O accumulation during the anoxic phase. The features considered in ANOX SVM model are shown in Table 2.

The classification matrices for train and test datasets of the ANSVM and ANOX SVM models are presented in Table 3. The average classification accuracy for the ANOX SVM model, was equal to 99% and 97% for the test and validation datasets. Similar results were obtained for the anaerobic phase with 95% and 98% accuracy in the train and test datasets respectively.

Jaramillo et al. (2018) developed an SVM classifier to estimate online the end of partial nitrification in a laboratory aerobic-anoxic SBR based on features extracted from pH and DO sensors over time-windows, resulting in 7.52% reduction in the operational time. In this study, the main focus was to estimate offline the behaviour of N\textsubscript{2}O emissions based on historical batch data. The results from this study indicate that ORP and pH sensor data can be used to detect the consumption of N\textsubscript{2}O during the nitritation/nitrification in SBR reactors. The results show that knowledge-based feature-extraction and SVM classification could help in explaining the behaviour of the system and potentially optimise the control to consider the consumption of accumulated N\textsubscript{2}O (i.e. in this system the denitrification can be stopped after the local maximum of the ORP rate after the nitrite “knee” in all the cycles investigated.)

Figures 9 (a) and (b) illustrate the predicted and measured N\textsubscript{2}O concentration at the end of the anaerobic phase (ANSVR model). The SVR parameters were optimised based on the root mean square error using the train dataset. RMSE of the SVR model was equal to 0.11 and 0.1 mg N\textsubscript{2}O-N/L for the train and test datasets respectively (R-squared equal to 0.85 and 0.75 respectively).
As shown in Figure 9 (b) the simulation results follow the behaviour of the actual dissolved N\textsubscript{2}O concentrations observed. One of the major factors affecting the performance is the limited number of data points, but the prediction is still accurate.

[Figure 9]

3.8 Prediction of the N\textsubscript{2}O concentration in aerobic phase

The input features are shown in Table 4 and were selected based on the identified influential variables. The N\textsubscript{2}O predicted values of the ANSVR model were used (anaerP). The procedure followed for the selection of model parameters was similar to the respective one followed for the anaerobic phase. Additionally, ANSVR test dataset cycles, were identified and used in AERSVR test dataset A. The model was also applied in anaerB cycles (test dataset B).

[Table 4]

Figure 10 (a), shows the predicted and measured average aerobic N\textsubscript{2}O concentration for the trained and test datasets. RMSE of the SVR model was equal to 0.06 and 0.11 mg N\textsubscript{2}O-N/L for the train dataset and test dataset A respectively, whereas the R-squared was equal to 0.94 and 0.82 (Figure 10 (a) and (b)).

[Figure 10]

The RMSE of the predicted values for the test dataset B, was equal to 0.29 mg N\textsubscript{2}O-N/L and the R-squared was equal to 0.72 (Figure 10 (a)). The AERSVR model underpredicted the average dissolved N\textsubscript{2}O concentration of test B dataset. This is expected given that in test B dataset cycles, the initial aerobic N\textsubscript{2}O accumulation exceeds the sensor calibration limit. Therefore, on many occasions the initial aerobic N\textsubscript{2}O accumulation was also underestimated (section 2.2.1 - anaerB cycles). An example is shown if Figure 11. In cycle A, the average dissolved N\textsubscript{2}O concentration
(calculated as discussed in section 2.2.1 for anaerB cycles) is equal to 1.33 mg/L. The AERSVR model predicted 0.87 mg/L underestimating the actual concentration (considering initial accumulation equal to 0.6 mg/L). In cycle B, the AERSVR model predicted N$_2$O concentration equal to 0.61 mg/L (considering initial accumulation equal to 0.6); the observed average dissolved N$_2$O concentration (after the local minimum), was equal to 0.6 mg/L.

The results show that under the investigated operational conditions, the framework shown in Figure 2 can provide a good estimation of the real dissolved N$_2$O behaviour and concentration observed during the different phases of SBR operation. Instabilities in the performance of machine learning models due to changes in the operational conditions in wastewater bioreactors have been reported in the literature (Shi and Xu, 2018). Therefore, long-term datasets and investigation of different patterns and dependencies should be investigated before model construction.

### 3.9 Mitigation strategy

During aerobic phases, elevated average dissolved N$_2$O concentration was linked with DO less than 1 mg/L and increased conductivity decrease rates (conductivity values represent NH$_4$-N concentration values in the reactor). Therefore, cycles with increased conductivity decrease rate indicate higher NH$_4$-N removal efficiency and NO$_2$N accumulation. Dissolved N$_2$O concentrations lower than 0.6 mg/L were identified in cycles with average DO concentration equal to ~1.36 mg/L, and conductivity decrease rate > 1.8 μS/cm/min. Increasing the reactor DO concentration to values higher than 1.3 mg/L can result in decreased aerobic N$_2$O generation (Law et al., 2012). However, with the current anaerobic supernatant feeding strategy, blowers operate at maximum flowrate, so it is not possible to increase the aeration in the system.
On the other hand, the implementation of a step-feeding strategy could foster the reduction of N$_2$O emissions thanks to the lower NH$_4$-N and free ammonia (FA) concentration at the beginning of the cycle, which has been recognized as a triggering factor for N$_2$O production (Desloover et al., 2012). Conductivity at the end of the cycle can act as surrogate to estimate the effluent NH$_4$-N concentration of the reactor and optimize the anaerobic supernatant feeding load. Consequently, the aerobic initial NH$_4$-N concentration could be controlled to avoid either FA accumulation or high AOR with subsequent N$_2$O generation.

Additionally, frequent alternation of aerobic/anoxic phases can be introduced in order to avoid high nitrite accumulation. The impact of nitrite concentration on N$_2$O production can be also minimized by ensuring adequate DO levels within the reactor to inhibit the nitrifiers denitrification pathway (Blum et al., 2018; Law et al., 2013). Rodriguez-Caballero et al. (2015) reported that in a full-scale SBR treating municipal wastewater, intermittent aeration (alternation between 20–30 min oxic and anoxic) led to a minimization of N$_2$O compared to long oxic periods that enhanced N$_2$O emission. The authors related this behaviour to the presence of shorter aeration times with subsequently lower nitrite accumulation and N$_2$O production.

In addition, Su et al. (2019) reported that slightly acidic or neutral pH in nitritation reactors (at values that do not inhibit microbial activity) can decrease N$_2$O generation by up to seven times. Based on the pH profiles observed in this study, regulation of aerobic (alkalinity consumption) phase duration can be also considered to control the pH at lower levels.

The developed models can be used to estimate rapidly and precisely the hard-to-measure N$_2$O concentrations during aeration and detect N$_2$O accumulation in non-aerated phases. Additionally, it can alert operators about cycles with anoxic and anaerobic N$_2$O accumulation and elevated aerobic N$_2$O concentrations, that require modifications to the system operation. The ANOXSVM
model can predict if N$_2$O is consumed in anoxic phases or if anoxic duration should be extended. Thus, additional provision of fermentation liquid can be performed to promote N$_2$O consumption through denitrification, when after 70-90 minutes the anoxic SVM model still indicates incomplete denitrification.

This study provides evidence on the relationship of DO, ORP and conductivity and pH with the dissolved N$_2$O concentration (in terms of correlation coefficients, behaviour and thresholds that indicate specific ranges of N$_2$O accumulation). These findings together with the models developed in this study, can be the basis for the development of intelligent control algorithms to integrate emissions control in sidestream SBR reactors performing nitritation/partial nitritation or other systems similar to S.C.E.N.A. Moreover, features based on ORP, pH, DO and conductivity measurements in wastewater SBR processes, that can be used to predict dissolved N$_2$O concentrations have been identified. The developed framework can be also tested in continuous processes for the data-driven prediction of N$_2$O emissions.

**Conclusions**

Knowledge discovery and data-mining techniques were employed to extract useful information about the dynamic behaviour of N$_2$O, and to predict the behaviour of dissolved N$_2$O concentration in a full-scale SBR reactor treating the anaerobic supernatant. The main conclusions are summarized as follows:

- The N$_2$O emissions in SCENA process varies from 1.3% to 19% of NH$_4$-N load, therefore they can contribute considerably to the operational carbon footprint of the process (~90% on average).
- Average aerobic dissolved N$_2$O concentration could significantly under similar influent loads, DO, pH and removal efficiencies. Extracting information from the dynamic in-
cycle behaviour of the variables monitored in the system is a significant step towards understanding N$_2$O behaviour.

- Aerobic dissolved N$_2$O concentration peaks (>1 mg/L), were observed in cycles with average DO concentrations less than 0.9-1 mg/L and ORP concentration at the end of the aerobic phase less than 40 mV. Conductivity was correlated with the reactor NH$_4$-N concentration (0.97). N$_2$O peaks were also observed in cycles with elevated decrease of conductivity during aeration. Step-feeding, control of initial NH$_4$-N concentrations and control of pH via the regulation of aerobic phase duration can mitigate the N$_2$O peaks observed in this study.

- The accumulation of N$_2$O at the end of the SBR anoxic phase was stripped in the subsequent aerobic phase and had a significant impact on the amount of N$_2$O emitted. The accumulated N$_2$O was consumed rapidly after nitrite ‘knee’ that was linked with the nitrite depletion. The ANOXSVM model can be used to detect if anoxic duration should be extended or additional fermentation liquid provided to enhance N$_2$O consumption in anoxic phases.

- This study shows that low-cost sensors, conventionally used to monitor SBR systems (i.e. pH, DO, ORP), have good capabilities to predict the dissolved N$_2$O behaviour and concentrations when couple with knowledge discovery techniques. The AERSVR model, showed reliable estimations of the aerobic N$_2$O concentration and can provide guidance to WWTPs operators, on whether N$_2$O levels are acceptable or mitigation actions are required.

Acknowledgements
This research was supported by the Horizon 2020 research and innovation program SMART-Plant (grant agreement No 690323).
References


IPCC, 2013. The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. USA: Cambridge University Press.


Table 1: Influent and effluent concentrations of the SCENA system

<table>
<thead>
<tr>
<th>Parameter</th>
<th>unit</th>
<th>mean</th>
<th>Sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBR Influent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH$_4$-N</td>
<td>mg/L</td>
<td>992.5</td>
<td>90</td>
</tr>
<tr>
<td>PO$_4$-P</td>
<td>mg/L</td>
<td>30.8</td>
<td>6.9</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>8.2</td>
<td>0.2</td>
</tr>
<tr>
<td>sCOD</td>
<td>mg/L</td>
<td>1111.7</td>
<td>562</td>
</tr>
<tr>
<td>Flow-rate</td>
<td>m$^3$/d</td>
<td>30 (8.4 per cycle)</td>
<td>2.2</td>
</tr>
<tr>
<td>Gas flow-rate</td>
<td>m$^3$/h</td>
<td>450 (170 - 520)</td>
<td>78</td>
</tr>
<tr>
<td>Dimensions</td>
<td>mxmxm</td>
<td>8 x 3.5 x 2.5</td>
<td></td>
</tr>
<tr>
<td>SBR Effluent</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH$_4$-N</td>
<td>mg/L</td>
<td>214.7</td>
<td>80.93</td>
</tr>
<tr>
<td>NO$_2$-N</td>
<td>mg/L</td>
<td>3.23</td>
<td>9.7</td>
</tr>
<tr>
<td>NO$_3$-N</td>
<td>mg/L</td>
<td>0.28</td>
<td>0.34</td>
</tr>
<tr>
<td>PO$_4$-P</td>
<td>mg/L</td>
<td>6.78</td>
<td>2.22</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>8.04</td>
<td>0.3</td>
</tr>
<tr>
<td>SBR Reactor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLSS</td>
<td>g/L</td>
<td>5.05</td>
<td>0.87</td>
</tr>
<tr>
<td>HRT</td>
<td>d$^{-1}$</td>
<td>1.05</td>
<td></td>
</tr>
<tr>
<td>SRT</td>
<td>d$^{-1}$</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>7.7</td>
<td>0.5</td>
</tr>
<tr>
<td>T</td>
<td>°C</td>
<td>30.02</td>
<td>1.56</td>
</tr>
<tr>
<td>Fermentation Unit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH$_4$-N</td>
<td>mg/L</td>
<td>715</td>
<td>72.6</td>
</tr>
<tr>
<td>PO$_4$-P</td>
<td>mg/L</td>
<td>86</td>
<td>12</td>
</tr>
<tr>
<td>pH</td>
<td></td>
<td>5.6</td>
<td>0.6</td>
</tr>
<tr>
<td>T</td>
<td>°C</td>
<td>36</td>
<td>5.1</td>
</tr>
<tr>
<td>sCOD</td>
<td>mg/L</td>
<td>13082</td>
<td>2228</td>
</tr>
<tr>
<td>Ferm_Hac</td>
<td>mg/L</td>
<td>3250</td>
<td>546</td>
</tr>
<tr>
<td>Ferm_HPr</td>
<td>mg/L</td>
<td>2281</td>
<td>588</td>
</tr>
<tr>
<td>Ferm_Hbut</td>
<td>mg/L</td>
<td>1347</td>
<td>196</td>
</tr>
<tr>
<td>Flow-rate to SBR</td>
<td>m$^3$/cycle</td>
<td>7.45 (~2.41 per cycle)</td>
<td>3.0</td>
</tr>
</tbody>
</table>
Table 2: Features used in the classification algorithm to predict the accumulation of dissolved $\text{N}_2\text{O}$ at the end of the anoxic and anaerobic phases

<table>
<thead>
<tr>
<th>Anaerobic</th>
<th>Anoxic</th>
<th>Anaerobic regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORP phase initial</td>
<td>Last ORP value</td>
<td>ORP phase initial</td>
</tr>
<tr>
<td>ORP change</td>
<td>ORP change</td>
<td>ORP change</td>
</tr>
<tr>
<td>First local maximum ORP first derivative</td>
<td>Mean pH</td>
<td></td>
</tr>
<tr>
<td>Local minimum of ORP first derivative after first local maximum ORP first derivative</td>
<td>Difference between first local maximum (after carbon dosage) and subsequent local minimum of the ORP first derivative</td>
<td>pH phase initial</td>
</tr>
<tr>
<td>Duration between first local maximum and subsequent local minimum of the ORP first derivative</td>
<td>Duration of carbon dosage</td>
<td>Time of ORP first derivative minimum/duration of phase</td>
</tr>
<tr>
<td>pH phase initial</td>
<td>Duration between first local maximum (after carbon dosage) and subsequent local minimum of the ORP first derivative</td>
<td></td>
</tr>
<tr>
<td>Time local minimum ORP first derivative/Phase duration</td>
<td>Last ORP first derivative</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Features selected in the SVR model for the aerobic phase

<table>
<thead>
<tr>
<th>Aerobic Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average conductivity rate</td>
</tr>
<tr>
<td>ORP end of aeration</td>
</tr>
<tr>
<td>ORP increase during aeration</td>
</tr>
<tr>
<td>Conductivity at the beginning of aeration</td>
</tr>
<tr>
<td>Average DO</td>
</tr>
<tr>
<td>pH at the beginning of aeration</td>
</tr>
<tr>
<td>Conductivity increase (based on the conductivity at the end of the aerobic phase of the previous cycle)</td>
</tr>
<tr>
<td>pH change during aeration</td>
</tr>
<tr>
<td>Initial aerobicN$_2$O concentration (based on ANSVR predictions)</td>
</tr>
</tbody>
</table>
Table 3: SVM classification results anoxic and anaerobic phases

<table>
<thead>
<tr>
<th>Phase</th>
<th>Dataset</th>
<th>Misclassified</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Accuracy (%)</th>
<th>Kappa</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anoxic phase</td>
<td>Train</td>
<td>anoxA: 1</td>
<td>1</td>
<td>0.99</td>
<td>99</td>
<td>0.97</td>
<td>anoxA: Final dissolved N₂O concentration end of anoxic &lt; 0.6 mg/L</td>
</tr>
<tr>
<td>cycle N</td>
<td>Test</td>
<td>anoxA: 1</td>
<td>1</td>
<td>0.98</td>
<td>98</td>
<td>0.92</td>
<td>anoxB: Final dissolved N₂O concentration end of anoxic &gt; 0.6 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>anoxB: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaerobic phase</td>
<td>Train</td>
<td>anaerA: 2</td>
<td>0.98</td>
<td>0.97</td>
<td>97</td>
<td>0.94</td>
<td>anaerA: N₂O end of anaerobic &gt; 2.6 mg/L</td>
</tr>
<tr>
<td>cycle N+1</td>
<td>Test</td>
<td>anaerA: 1</td>
<td>1</td>
<td>0.97</td>
<td>98</td>
<td>0.95</td>
<td>anaerB: N₂O end of anaerobic &lt; 2.6 mg/L</td>
</tr>
<tr>
<td></td>
<td></td>
<td>anaerB: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1: Schematic representation of a complete cycle in the S.C.E.N.A process and datasets used in the analysis.
Figure 10: (a) Predicted vs measured dissolved N2O concentration (AERSVR) in the aerobic phase for the train dataset, the test dataset A and the test dataset B and (b) comparison of predicted and measured dissolved N2O concentration for the test dataset B.
Figure 11: An example of dissolved N2O profiles for cycles belonging to anaerB cycles (test dataset B). The red points represent the first point considered for the calculation of the average aerobic N$_2$O accumulation (as described in section 2.2.1). Data points in the beginning of aeration exceeding sensor calibration limits are not shown.
Figure 2: Methodological Framework followed in the study
Figure 3: (a) N$_2$O emissions and (b) aerobic phase conductivity decrease, during monitoring campaign (gas analyser, March-April)
Operating Conditions

<table>
<thead>
<tr>
<th></th>
<th>Unit</th>
<th>26/03</th>
<th>11/04</th>
</tr>
</thead>
<tbody>
<tr>
<td>Influent NH$_4$-N</td>
<td>mg/L</td>
<td>1186.3</td>
<td>1154.13</td>
</tr>
<tr>
<td>DO</td>
<td>mg/L</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Anaerobic/Aerobic/Anoxic duration</td>
<td>h</td>
<td>0.5/3.6/1.5</td>
<td>0.5/3.7/1.5</td>
</tr>
<tr>
<td>Effluent NH$_4$-N</td>
<td>mg/L</td>
<td>211.1</td>
<td>185.9</td>
</tr>
<tr>
<td>Effluent NO$_2$-N</td>
<td>mg/L</td>
<td>~0</td>
<td>1.6</td>
</tr>
<tr>
<td>Temperature</td>
<td>°C</td>
<td>31.7</td>
<td>32.1</td>
</tr>
</tbody>
</table>

Figure 4: (a) Example of the effect of N$_2$O emissions in the operational carbon footprint for two cases, (b) aerobic profiles of conductivity, ORP and (c) DO for the two cases shown in (a)
Figure 5: Representative cycle profile for the (a) dissolved $\text{N}_2\text{O}$ concentration, $\text{N}_2\text{O}$ emissions, conductivity, DO, (b) ORP and pH, and (c) $\text{NH}_4\text{-N}$, $\text{NO}_2\text{-N}$ and $\text{PO}_4\text{-P}$ concentrations
Figure 6: (a) Daily average conductivity at the end of the aerobic phase versus effluent NH$_4$-N concentration (coloured points: average dissolved N$_2$O accumulated in the aerobic phase), (b) Aerobic average accumulated dissolved N$_2$O in respect to DO concentration; only cycles without initial N$_2$O accumulation from the previous anoxic cycle are shown (coloured points: ORP at the end of the aerobic phase)
Figure 7: (a) Representative profiles of dissolved N$_2$O concentration based on different initial concentrations of N$_2$O in the beginning of the aerobic phase and (b) ORP and DO profiles.
Figure 8: Box-plots of the (a) initial anaerobic ORP and (b) the ORP change during the anaerobic phase for cycles with and without N$_2$O consumption (Class 0: no significant N$_2$O consumption or anaerobic N$_2$O concentration > 2.6 mg/L; Class 2: significant N$_2$O consumption).
Figure 9: (a) Predicted vs measured dissolved N$_2$O concentration in the end of the anaerobic phase (ANSVR) for the test and train datasets and (b) comparison of predicted and measured dissolved N$_2$O concentration for the test dataset.
Highlights

- S.C.E.N.A NO emissions contribute up to 97.3% to the operational carbon footprint
- Stripping of accumulated NO from the previous anoxic cycle increases emissions
- Aerobic dissolved NO concentration is correlated with DO and conductivity
- Conductivity can be used to control SBR NH4+ concentration and NO
- NO can be estimated as a latent parameter from other low-cost sensors
Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: