Relating \( \text{N}_2\text{O} \) emissions during biological nitrogen removal with operating conditions using multivariate statistical techniques

V. Vasilaki \(^a\), E.I.P. Volcke \(^b\), A.K. Nandi \(^c\), M.C.M. van Loosdrecht \(^d\), E. Katsou \(^a\), *

\(^a\) Department of Civil & Environmental Engineering, Brunel University London, Uxbridge UB8 3PH, UK
\(^b\) Department of Green Chemistry and Technology, Ghent University, Coupure Links 653, 9000 Gent, Belgium
\(^c\) Department of Electronic and Computer Engineering, Brunel University London, Uxbridge UB8 3PH, UK
\(^d\) Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands

\section*{Abstract}
Multivariate statistical analysis was applied to investigate the dependencies and underlying patterns between \( \text{N}_2\text{O} \) emissions and online operational variables (dissolved oxygen and nitrogen component concentrations, temperature and influent flow-rate) during biological nitrogen removal from wastewater. The system under study was a full-scale reactor, for which hourly sensor data were available. The 15-month long monitoring campaign was divided into 10 sub-periods based on the profile of \( \text{N}_2\text{O} \) emissions, using Binary Segmentation. The dependencies between operating variables and \( \text{N}_2\text{O} \) emissions fluctuated according to Spearman’s rank correlation. The correlation between \( \text{N}_2\text{O} \) emissions and nitrite concentrations ranged between 0.51 and 0.78. Correlation >0.7 between \( \text{N}_2\text{O} \) emissions and nitrate concentrations was observed at sub-periods with average temperature lower than 12 \( ^\circ \text{C} \). Hierarchical k-means clustering and principal component analysis linked \( \text{N}_2\text{O} \) emission peaks with precipitation events and ammonium concentrations higher than 2 mg/L, especially in sub-periods characterized by low \( \text{N}_2\text{O} \) fluxes. Additionally, the highest ranges of measured \( \text{N}_2\text{O} \) fluxes belonged to clusters corresponding with NO\(_3\)-N concentration less than 1 mg/L in the upstream plug-flow reactor (middle of oxic zone), indicating slow nitriﬁcation rates. The results showed that the range of \( \text{N}_2\text{O} \) emissions partially depends on the prior behavior of the system. The principal component analysis validated the findings from the clustering analysis and showed that ammonium, nitrate, nitrite and temperature explained a considerable percentage of the variance in the system for the majority of the sub-periods. The applied statistical methods, linked the different ranges of emissions with the system variables, provided insights on the effect of operating conditions on \( \text{N}_2\text{O} \) emissions in each sub-period and can be integrated into \( \text{N}_2\text{O} \) emissions data processing at wastewater treatment plants.

© 2018 Published by Elsevier Ltd.

1. Introduction

The increasing demand to reduce the carbon footprint of municipal wastewater treatment plants (WWTPs) by reducing greenhouse gas (GHG) emissions and energy consumption, is posing new challenges for the water industry (Flores-Alsina et al., 2014). The climate change pressures prompt the quantification and minimization of GHG emissions generated in WWTPs (Haas et al., 2014). Three main sources of GHG emissions prevail in WWTPs (Monteith et al., 2005; Mannina et al., 2016): (i) the direct emissions mainly linked to biological processes, (ii) the indirect internal emissions generated by the use of imported energy to the plants, and (iii) the indirect external emissions associated with the sources that are controlled outside the WWTPs (e.g. chemicals production, disposal of sewage sludge, transportation). The GHGs emitted into the atmosphere from biological wastewater treatment processes are carbon dioxide (CO\(_2\)), methane (CH\(_4\)) and nitrous oxide (N\(_2\)O) (Kampf et al., 2009b).

With the potential contribution of 265 times more than CO\(_2\) for a 100-year time horizon to global warming (IPCC, 2013), N\(_2\)O is a potent GHG and the most significant contributor to ozone depletion (Ravishankara et al., 2009). WWTPs are significant generators of N\(_2\)O and are responsible for 3.1% of the N\(_2\)O emissions in Europe (EEA Report, 2017). N\(_2\)O is generated mainly during the autotrophic nitrification and heterotrophic denitrification (Kampf et al.,...
Abbreviations

AOR Ammonia oxidation rate
CH₄ Methane
CO₂ Carbon dioxide
DO Dissolved oxygen
GHG Greenhouse gas
N₂O Nitrous oxide
NH₄-N Ammonium nitrogen
NO₂-N Nitrite nitrogen
NO₃-N Nitrate nitrogen
PC Principal component
PCA Principal component analysis
PLS Partial least squares
TN Total nitrogen
WWTP Wastewater treatment plant

and can contribute up to 78% (Daelman et al., 2013) of the footprint of a WWTP’s operation. Recent studies have focused on the understanding, quantification, control and minimization of N₂O emissions (Aboobakar et al., 2013; Mampaey et al., 2016; Pan et al., 2016). However, several studies have resulted in contradicting findings on the influence of operating and environmental variables on N₂O generation (Li et al., 2016; Massara et al., 2017). For instance, several studies have reported increasing N₂O emissions with decreasing DO concentrations during nitrification (Kampschreur et al., 2009b). However, Rodriguez-Caballero et al. (2014) found that N₂O emission profiles in a full-scale biological reactor did not change even for DO variations higher than 1.5 mg/L. The latter, was attributed to the high nitrification efficiency and the potential biomass adaptation to continuously varying DO concentrations. Results from real-field N₂O monitoring campaigns cannot fully explain long-term causes of N₂O emissions and the combined effect of operating, environmental and external factors that influence the biological systems (Jonsson et al., 2015). Long-term full-scale monitoring campaigns have shown that N₂O fluxes are highly dynamic with significant diurnal fluctuations and seasonal variations; however, the dynamics cannot be fully explained (Daelman et al., 2015; Kosonen et al., 2016).

Several mechanistic process models describing N₂O emissions from wastewater treatment plants have been developed over the last few years (Massara et al., 2017). While they have been successfully applied to identify N₂O formation mechanisms and pathways from experimental data (Ni et al., 2015; Poquet et al., 2016), their validation and calibration to long-term process data remains a challenge. Domingo-Feliz and Smets (2016) reported that substrate affinity constants for NO₂ and NO reduction in existing N₂O models differ by a factor of about 100. Additionally, calibration of models under specific operational conditions (i.e. dry weather) can affect their performance and accuracy when the system varies (Gernaey et al., 2004; Guo and Vanrolleghem, 2014). Moreover, full-scale N₂O emission data show long-term trends that cannot be explained by commonly available operational data (Daelman et al., 2015) but are possibly caused by microbial population changes, which are hard to catch with the current models, typically describing single functional groups with fixed parameter sets. Multivariate statistical techniques are capable of identifying relationships between N₂O emissions and a multitude of influencing factors, at the same time identifying various operating sub-periods for which this behaviour may differ. This will lead to increased understanding of experimental data, on its turn facilitating the application, calibration and validation of mechanistic models. As such, multivariate statistical techniques maximize the information acquired from N₂O monitoring campaign data.

Statistical techniques have been used for the analysis of data from full-scale monitoring campaigns, to identify interconnections between operating and environmental variables on the one hand and N₂O formation on the other hand. Through multiple linear regression analyses, Aboobakar et al. (2013) showed dependencies between N₂O emissions and nitrogen load, temperature and dissolved oxygen (DO) in various compartments of a plug-flow reactor for biological nitrogen removal. Multi-regression analysis of one year of data with bi-monthly sampling frequency, coming from a full-scale SBR (Sun et al., 2013) indicated negative correlation between N₂O emissions and temperature, while COD/N ratio lower than 6 resulted in higher emissions. Brotto et al. (2015) used Spearman’s rank correlation to explain the behavior of N₂O emissions in an activated sludge process. The analysis showed negative correlation between N₂O emissions and pH but positive correlation between N₂O fluxes and temperature. However, most of the studies did not consider continuous long-term operational data, while further analysis is required to gain a better understanding on the dynamics and trade-offs between N₂O generation and the online monitored and controlled process variables.

Multivariate analysis and data mining have been proposed to be a suitable method for the identification of patterns and hidden relationships within WWTP data (Rosén and Lexmond, 2001) and can be applied to provide insights on the combined effect of operational variables on N₂O emissions in full-scale systems. Chemometric techniques have been applied to the wastewater treatment sector for 40 years (Rosén and Olsson, 1998), enabling the visualization and interpretation of the multi-dimensional interrelations of the operational variables monitored in biological processes (Platikanov et al., 2014). Their application can (i) improve the efficiency of process monitoring (Mirin and Wahab, 2014) and provide further insights of the biological processes (Moon et al., 2009), (ii) identify and isolate process faults (Haimi et al., 2016; Liu et al., 2014; Maere et al., 2012; Rosen and Yuan, 2001), and sensor faults (Lee et al., 2004), and iii) predict significant operating variables in the biological systems that affect performance (Rustum et al., 2008). Furthermore, the gradual implementation of online sensors to monitor important parameters in the biological treatment train of WWTPs results in the production of time series, which require the application of specific statistical tools for their interpretation. The most widely applied approaches include methods aiming to reduce the dimensionality of large data-sets (i.e., principal component analysis (PCA), partial least squares (PLS)) and data clustering techniques (i.e., hierarchical clustering, k-means clustering) (Haimi et al., 2013). However, there are limited studies investigating the behavior of N₂O emissions with the application of multivariate statistical techniques, especially utilizing online operational data in long-term monitoring.

The aim of this work is to investigate whether widely applied multivariate statistical techniques can be applied to the online data collected from real-field N₂O monitoring campaigns in order to gain a better understanding on the dynamic behavior of N₂O emissions and explain the combined effect of the operating variables monitored in wastewater treatment processes on N₂O emissions. Hourly data from the operating variables monitored online and N₂O emissions data in a full-scale carousel reactor from the long-term monitoring campaign published by Daelman et al. (2015) were used for the analysis. A statistical methodological approach was developed, applying changepoint detection techniques to identify changes in the N₂O fluxes behavior combined with hierarchical k-means clustering and PCA, to provide insights on N₂O emissions patterns and generation pathways.
2. Materials and methods

2.1. Process description and data origin

This work was based on the data obtained by Daelman et al. (2015) for the Kralingseveer WWTP, consisting of a plug-flow reactor followed by two carrousel reactors in parallel (Fig. 1). The plant treated 80,000 m³ d⁻¹ of domestic wastewater from a combined sewer system. The carrousel reactors were characterized by alternating anoxic/oxic zones; aeration was performed through surface aerators, which were manipulated to control the ammonium concentration in the effluent. Aerator 1 operates under on/off pattern, being on when the ammonium concentration was higher than 1.2 mg N/L, while surface aerators 2 and 3 were always operational to keep the solids from settling but operated at maximum capacity when the ammonium concentration became higher than 0.6 and 0.9 mg/L, respectively. Over the monitoring period the average total nitrogen (TN) removal efficiency was 81 ± 10%; the average COD removal efficiency was equal to 87 ± 5%.

Ammonium nitrogen (NH₄-N), nitrate nitrogen (NO₃-N) and DO were monitored in the middle of the second oxic zone in the plug flow reactor (location 1, Fig. 1). The carrousel reactors were equipped with NH₄-N, temperature probes, and 3 DO probes (DO1, DO2, DO3) (locations 2, 3, 4, Fig. 1). The Northern carrousel reactor was also equipped with a nitrite probe. All the reactors were covered, and the off-gas was collected in ducts and pumped to a Servomex gas analyzer, where N₂O was measured. Table S1 lists all the variables monitored online (Supplementary material). The data matrix developed consists of the variables monitored in the carrousel reactor (DO, NH₄-N, NO₃-N, NO₂-N, N₂O), the influent flow-rate, as well as the NH₄-N and NO₃-N concentrations from the plug-flow reactor. 24 h composite samples of influent and effluent, available about every 6 days, were used to support the analysis. Fig. 2, summarizes the methodological framework applied to the online database.

2.2. Methodological framework for data analysis

The monitoring period was divided into distinct sub-periods based on the profile of N₂O fluxes in the carrousel reactor. Spearman’s correlation analysis, k-means clustering, hierarchical clustering, and Principal component analysis were applied to the database. The application of clustering algorithms facilitated the identification of operational modes that have historically resulted in specific ranges of N₂O emissions. The PCA reduced the dimensionality of the data-set transforming the sensor signals into useful knowledge that can be easily interpreted. The methodological framework is extensively described in the following sub-sections.

The data-driven approach enabled the utilization of the information and patterns embedded in the real-time monitored variables (from the system sensors) in the biological processes and GHG measurements. Multivariate statistical analysis is an alternative to univariate analysis that is commonly applied for the analysis of WWTP data. It enables the identification of patterns and interrelations in data-sets by examining multiple variables simultaneously (Olsson et al., 2014). R software was used for the statistical analysis (R Core Team, 2017). The complete list of packages used is provided in the supplementary material (Table S2).

2.2.1. Preliminary data processing

The preliminary data analysis included: (i) data synchronization under the same time-stamp, and ii) removal of duplicate and unreliable measurements (multiple readings at the same time stamp for the same sensor). The data were aggregated into hourly averages in order to compensate for the missing data due to variation in sampling frequency between the different variables monitored. Exponential moving average imputation was applied when less than 24 consequential data were missing for each variable. Longer periods of missing data were excluded from the analysis.

2.2.2. Binary segmentation changepoint detection

Given a series of data, change point analysis investigates abrupt changes in a data-series when specific properties change (i.e., mean

![Fig. 1. Layout of Kralingseveer WWTP with Plug-flow and Carrousel reactors, adapted from Daelman et al. (2015).](image-url)
and variance) (Kawahara and Sugiyama, 2012). The Binary Segmentation (Scott and Knott, 1974) is a widely applied and computationally efficient changepoint detection algorithm (Killick et al., 2012). The algorithm employs initially single changepoint detection method to the complete data-set as described in (Killick and Eckley, 2014). If a changepoint is identified the procedure is repeated to the two new segments formed; before and after the changepoint. The process continues splitting the data until there are no more changepoints identified. The computational cost of the algorithm is of the order of $O(n \log n)$ with $n$ being the number of data in the data-set and therefore it is applicable in large data-sets. A distribution-free test statistic was applied based on the work of Chen and Gupta (1997). The penalty for the changepoints identification was equal to $\log(n)$. The algorithm requires independent data points. Therefore, first difference transformation of the $N_2O$ timeseries was performed and changes in variance were identified by the Binary segmentation algorithm. The profile of the $N_2O$ emissions was highly variable during the monitoring campaign. Binary segmentation enabled the identification of the sub-periods characterized by different $N_2O$ emissions’ profile.

2.2.3. Spearman’s rank correlation

Spearman’s rank correlation coefficient (Spearman, 1904) was used to detect bivariate temporal monotonic trends among the system variables for the different sub-periods; it served as a measure of the association strength. This method is based on the rank of the values and therefore, is less sensitive to outliers than Pearson’s correlation. P values lower than 0.01 were considered to be significant.

2.2.4. Hierarchical k-means clustering

Clustering techniques are widely applied in data mining in order to identify and group the underling patterns that exist in high dimensional data sets (Jain, 2010). K-means clustering (Hartigan and Wong, 1979) is a recognized clustering algorithm (Haimi at al., 2013). K-means clustering was applied to categorize the data in groups of similar observations and to investigate the patterns of $N_2O$ emission fluxes, based on Euclidean distance. K-means algorithm begins with the selection of k random centroids of the same dimension within the original data. Data-points are compared and assigned to the nearest centroid. During each iteration, the nearest data to each centroid are re-defined and centroids are recalculated in a way that squared distances of all points within a cluster to the cluster’s centroid are minimized. However, the randomly selected initial centroids can result into locally optimized clustering results (Abu-Jamous et al., 2015). Therefore, hierarchical k-means clustering that was proposed by Arai and Barakbah (2007), was applied to the dataset. In this method agglomerative hierarchical clustering (Kaufman and Rousseuw, 1990) is applied for the selection of the centroids; Ward’s method is used in order to divide the dataset in clusters (Ward, 1963). The data were normalized before the analysis. NBclust package in R (Charrad et al., 2014) was used to select the number of clusters in each sub-period. The package applies a number cluster validity

![Fig. 2. Methodology followed in the current study for data processing and visualization.](image-url)
Hierarchical k-means clustering was applied to the carrousel reactor data matrix from the different sub-periods identified through binary segmentation, to investigate whether different temporal patterns of the operating variables were responsible for the different behavior of $N_2O$ emissions. Hierarchical k-means clustering enabled i) the detection of frequency and persistence of extreme ranges of operating variables, and ii) the comparison of the operational modes between the plug-low and carrousel reactor. Ammonium and nitrate probes in the plug-flow reactor were included in the analysis, since they can provide indirect feedback in terms of the carrousel reactor influent and additional information for the operational behavior of the system. However, the analysis was repeated excluding plug-flow variables (NH$_4$-N and NO$_3$-N). Graphical comparisons of the clustered data-points versus time and boxplots of the variables in each identified cluster are displayed in the results' section.

### 2.2.5. Principal component analysis

Principal component analysis (PCA) (Jolliffe, 2002) was applied to the dataset in an effort to reduce the dimensionality of the data by eliminating a small proportion of variance in the data. PCA transforms the original correlated measured variables to uncorrelated variables, i.e., Principal components (PCs), explaining the maximum observed variability. The principal components are linear combinations of the original data variables. The loadings of the variables in each principal component can map their relationship with the respective principal component. PC scores are a linear combination of the data, weighted by the PC loadings for each variable. The scores of the principal components map the different samples in the new dimensional space of the principal components facilitating the investigation of the different relationships between the variables. The data matrices (X) consisting of J columns (variables) and I data rows (number of observations) were normalized with mean equal to 0 and standard deviation equal to 1. Each column of $X$, $x_j = (x_{ij}, ..., x_{Ij})^T$, $j = 1, ..., J$, represents a vector in the 1-dimensional space. In PCA, eigenvalue decomposition is used to factorize the data matrix $X$ ($I \times J$) and to map the data matrix to a reduced dimensional space:

$$X = TP^T + E$$

where, $T$: matrix ($I \times S$) representing the score of the principal components, $S$: the number of principal components selected, $P$: matrix ($J \times S$) representing the loadings and $E$: matrix of residuals.

The biplot of the first 2 PCs was used in order to visualize the combined behavior of significant variables that affect the system. The biplots enabled the simultaneous visualization of i) the variables' loadings in the first two principal components, ii) the scores of the first two principal components, and iii) the different clusters. The temporal variations of the PC scores enabled the identification of occasions in which the behavior of the system changes. PCA was applied to the data matrix of the carrousel reactor excluding $N_2O$ emissions time series, i) to identify the most significant variables that affect the system, ii) to analyze the structure of the sensor data, iii) to investigate if changes in the relationship of the system coincide with changes in the $N_2O$ emissions profile, and iv) to validate the results from hierarchical clustering. $N_2O$ emissions time series were excluded from the PCA in order to investigate the relationship between the PC scores and $N_2O$ emissions and to examine which PCs are most significantly linked to the behavior of $N_2O$ emissions.

### 3. Results and discussion

#### 3.1. $N_2O$ emissions profile and main dependencies

The profile of all the variables monitored was fluctuating during the monitoring period, which can justify the different profiles of $N_2O$ emissions that resulted from the Binary Segmentation algorithm. Overall, high ranges of emissions were reported when nitrate concentration in the plug-flow reactor was low, whereas periods with lower ammonium concentrations in the plug-flow reactor were linked with lower $N_2O$ emissions.

Table 1 shows the average values and standard deviations of the variables monitored online and offline in the Northern carrousel and plug-flow reactors. $N_2O$ fluxes peaked in March 2011 followed by a period characterized by very low $N_2O$ emissions. Gradual decrease was observed until November 2011 and negligible emissions again until January 2011 (Fig. 3). The application of Binary Segmentation algorithm to the $N_2O$ emissions of the Northern carrousel reactor identified 9 change-points that correspond to 10 sub-periods with distinct variance of the $N_2O$ timeseries first difference (Fig. 3). The analysis identified abrupt temporal changes in the emission dynamics that indicate changes in the underlying mechanisms or environmental conditions responsible for the $N_2O$ formation.

Offline data were analyzed in the different sub-periods in order to investigate significant changes that can contribute to the high $N_2O$ emissions in sub-periods 4 and 5. The average COD concentration in the influent of the plug-flow reactor (effluent of primary sedimentation) was $239 \pm 80$ mg COD/L over the 15-month monitoring period. The average plug-flow reactor influent and carrousel reactor effluent concentrations of COD, TKN, BOD, TP and the effluent pH for all sub-periods are given in the supplementary material (Table S3). In sub-period 5, 27% increase in the influent COD concentration to the plug flow reactor (compared to average value) was observed, which could be attributed to less precipitation events and to the consequently lower average influent flow-rate during this sub-period. Laboratory analyses did not show significant seasonal changes in the plug-flow COD loading (19,934 ± 13310 kg COD/day). The COD loading in sub-period 4 (16,160 ± 2546 kg COD/day) was 17% less than in sub-period 1. TKN and TP loadings were reduced in sub-period 4 compared to sub-period, by 11% and 12% respectively. The COD:TKN:TP ratio remained quite stable, ranging between 1:0.17:0.02 (sub-period 2) and 1:0.20:0.03 (sub-period 4).

Fig. 4 shows the different COD to TKN ratios measured for all the sub-periods. There were cases with lower than average COD/TKN in the influent of the plug-flow reactor that coincided with increased $N_2O$ emissions, particularly in sub-periods 4 and 5. However, low ranges of COD/TKN (<5) in sub-periods 1, 2, 7 and 6 corresponded with low $N_2O$ emissions. These observations indicate that limitation of COD cannot be considered the sole contributor of $N_2O$ emissions via heterotrophic denitrification in sub-periods 4 and 5.

The COD removal efficiency remained relatively steady during the monitoring campaign ranging from 79% (sub-period 8) to 91% (sub-period 5). The range of TN and TP removal efficiencies ranged from 73% (sub-periods 1 and 9) to 92% (sub-period 5) and from 67% (sub-period 7) to 87% (sub-period 4) respectively. The effluent pH was steady (~8) and did not show seasonal variability that could influence the generation of $N_2O$ emissions. On the other hand, a significant variation is observed for all variables monitored online by analyzing at the complete database. Table 2 summarizes the average values and standard deviations of the online monitored variables considered in the analysis for the target periods. In the carrousel reactor, the nitrite concentration is relatively high in sub-period 4 (average = 2.6 mg/L) and in the first
part of sub-period 10 (average = 2.1 mg/L). The average temperature in both cases is ~13°C. In biological reactors operating in continuous mode, appreciable (> 2 mg N/L) nitrite concentrations are usually not observed, since nitrite is directly oxidized by nitrite oxidizing bacteria into nitrate. However, in certain cases, high nitrite concentrations in biological processes have been observed, which have been linked with low temperatures that affect N2O reductase during denitrification enhancing N2O production.

Table 1
Average value and standard deviation (std) of variables monitored in the Northern carrousel reactor (C: carrousel reactor, N: Northern, PF: plug-flow reactor).

<table>
<thead>
<tr>
<th>Online variables</th>
<th>Average</th>
<th>Std</th>
<th>Offline variables</th>
<th>Average</th>
<th>Std</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2O (kg/h)</td>
<td>1.4</td>
<td>2.1</td>
<td>COD influent (mg COD/L)</td>
<td>238.8</td>
<td>79.5</td>
</tr>
<tr>
<td>NH4-N C (mg/L)</td>
<td>1.63</td>
<td>2.2</td>
<td>TKN influent (mg/L)</td>
<td>42.1</td>
<td>10.0</td>
</tr>
<tr>
<td>NO2-N C (mg/L)</td>
<td>5.8</td>
<td>4</td>
<td>TP influent (mg/L)</td>
<td>7.0</td>
<td>2.1</td>
</tr>
<tr>
<td>NO3-N C (mg/L)</td>
<td>1.2</td>
<td>1.1</td>
<td>Flow-rate (m³/d)</td>
<td>85,898</td>
<td>41,786</td>
</tr>
<tr>
<td>DO1 (mg/L)</td>
<td>0.6</td>
<td>0.9</td>
<td>COD effluent (mg/L)</td>
<td>36.9</td>
<td>6.9</td>
</tr>
<tr>
<td>DO2 (mg/L)</td>
<td>0.8</td>
<td>0.9</td>
<td>TKN effluent (mg/L)</td>
<td>2.8</td>
<td>1.2</td>
</tr>
<tr>
<td>DO3 (mg/L)</td>
<td>1.9</td>
<td>0.6</td>
<td>TP effluent (mg/L)</td>
<td>1.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>16</td>
<td>3.5</td>
<td>pH effluent</td>
<td>8.0</td>
<td>0.2</td>
</tr>
<tr>
<td>N2O PF (kg/h)</td>
<td>0.71</td>
<td>1.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH4-N PF (mg/L)</td>
<td>12.41</td>
<td>5.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO3-N PF (mg/L)</td>
<td>2.38</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Influent Flow-rate (m³/h)</td>
<td>3973</td>
<td>2375</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DO PF (mg/L)</td>
<td>2.61</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3. (top): N2O emissions profile in the Northern Carrousel reactor (grey area: periods with missing N2O data); (bottom): First difference of the N2O emissions timeseries (blue line) showing the sub-periods identified by the application of binary segmentation (grey area: periods with missing N2O data, blue dotted lines: changepoints identified by the algorithm, red horizontal lines: standard deviation in each sub-period). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
Analyzing the whole profile, the emissions tended to be low at higher temperatures (sub-periods 6, 7, and 8). Higher emissions were also observed, though, at temperature higher than 18°C and low nitrite concentrations (i.e., sub-period 5). Ahn et al. (2010) demonstrated that N2O emissions can be significant at higher temperatures due to the higher enzymatic activities of the bio-processes producing N2O. In the carrousel reactor during sub-periods 4 and 5, the temperature increases from 11.8 to 20°C.

Low N2O emissions were also observed when ammonium concentration was lower than 13 mg/L and nitrate was higher than 2.5 mg/L in the plug-flow reactor. The probe was located in the middle of the second oxic zone; thus, lower ammonium concentrations in the plug-flow reactor can indicate less ammonium loads in the carrousel reactor.

The analysis of the variables’ ranges for the N2O emission profiles provides limited insight on the dependencies between the system variables monitored online, which is further analyzed in the following sections.

3.2. Spearman’s rank correlation analysis for carrousel reactor

The application of Spearman’s rank correlation coefficient to the data of the carrousel reactor did not identify significant correlations between the N2O emissions and the operating variables. The lack of monotonic univariate dependencies could be attributed to i) the temporal fluctuations of the influent characteristics, ii) the continuous variability in the operating conditions of the reactors, and iii) the seasonal variations of the environmental conditions in wastewater treatment processes. Fluctuating correlation coefficients between N2O emissions and carrousel reactor variables were identified (Supplementary, Figs. S1:S2). The findings are in line with the study of Kosonen et al. (2016). The authors compared the results from two monitoring periods at the same biological system and identified different relationships between N2O emissions and BOD7/ATU loads.

The correlation coefficient between nitrite and N2O emissions ranged from 0.78 (sub-period 7) to 0.51 (sub-period 9). As a general remark, nitrite was correlated with N2O emissions in sub-periods 4, 6 and 7, while lower correlation was observed during sub-periods 5, 8 and 9. N2O emissions and NO3-N concentration in the carrousel reactor exhibited a positive correlation with coefficient higher than 0.7 for sub-periods 2 (Figs. 5), 4 and 10 (the temperature was lower than 13°C in all cases). N2O emissions and NO3-N concentrations followed similar diurnal patterns, wherein peaks in nitrate concentration coincided with peaks in N2O emissions (Daelman et al., 2015). The accumulation of nitrate is potentially linked with higher nitrification than denitrification rates. This is in line with Daelman et al. (2015), considering that the nitrate utilization rate in these sub-periods is affected by the low temperatures

(Holtan-Hartwig et al., 2002; Adouani et al., 2015).

Table 2

<table>
<thead>
<tr>
<th>Variables</th>
<th>N2O (kg/h)</th>
<th>NO3-C</th>
<th>N (mg/l)</th>
<th>NO3-N</th>
<th>PF (mg/l)</th>
<th>NH4-N</th>
<th>C (mg/l)</th>
<th>NH4-N</th>
<th>PF (mg/l)</th>
<th>NO2-N</th>
<th>Ca (mg/l)</th>
<th>Temperature (°C)</th>
<th>DO1 (mg/l)</th>
<th>DO2 (mg/l)</th>
<th>DO3 (mg/l)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
<td>Std</td>
<td>Std</td>
<td>Std</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.1</td>
<td>6.1</td>
<td>3.1</td>
<td>1.8</td>
<td>1.6</td>
<td>1.8</td>
<td>2.67</td>
<td>11.4</td>
<td>4.1</td>
<td>15.7</td>
<td>1.4</td>
<td>0.62</td>
<td>0.7</td>
<td>0.62</td>
</tr>
<tr>
<td>2</td>
<td>0.6</td>
<td>0.6</td>
<td>7.2</td>
<td>3.1</td>
<td>2.5</td>
<td>2</td>
<td>1.5</td>
<td>1.7</td>
<td>13</td>
<td>4</td>
<td>11.2</td>
<td>1.0</td>
<td>0.77</td>
<td>1</td>
<td>1.31</td>
</tr>
<tr>
<td>3</td>
<td>2.7</td>
<td>1.4</td>
<td>6.1</td>
<td>3.2</td>
<td>1.6</td>
<td>2.1</td>
<td>1.6</td>
<td>2.1</td>
<td>15.2</td>
<td>4.5</td>
<td>11.5</td>
<td>0.7</td>
<td>0.67</td>
<td>0.8</td>
<td>1.49</td>
</tr>
<tr>
<td>4</td>
<td>5.6</td>
<td>2.6</td>
<td>3</td>
<td>0.1</td>
<td>0.5</td>
<td>0.7</td>
<td>1.3</td>
<td>1.6</td>
<td>15</td>
<td>4.8</td>
<td>2.6</td>
<td>1.9</td>
<td>12.9</td>
<td>1.1</td>
<td>0.64</td>
</tr>
<tr>
<td>5</td>
<td>2.6</td>
<td>2.2</td>
<td>4.3</td>
<td>4.2</td>
<td>3.1</td>
<td>1.9</td>
<td>1.3</td>
<td>2</td>
<td>11.5</td>
<td>5.2</td>
<td>0.8</td>
<td>1</td>
<td>18.2</td>
<td>1.7</td>
<td>0.34</td>
</tr>
<tr>
<td>6</td>
<td>0.8</td>
<td>1.4</td>
<td>3.3</td>
<td>3.2</td>
<td>2.3</td>
<td>1.9</td>
<td>2</td>
<td>3.1</td>
<td>14.7</td>
<td>6.1</td>
<td>0.5</td>
<td>0.5</td>
<td>20</td>
<td>1.0</td>
<td>0.42</td>
</tr>
<tr>
<td>7</td>
<td>0.2</td>
<td>0.3</td>
<td>7.2</td>
<td>5</td>
<td>2.8</td>
<td>2.4</td>
<td>2</td>
<td>3.1</td>
<td>9.8</td>
<td>5.2</td>
<td>0.6</td>
<td>0.4</td>
<td>20</td>
<td>0.7</td>
<td>0.42</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>0.2</td>
<td>10.1</td>
<td>5.7</td>
<td>5.2</td>
<td>2.6</td>
<td>1.4</td>
<td>1</td>
<td>9.6</td>
<td>5.5</td>
<td>0.8</td>
<td>0.5</td>
<td>19.6</td>
<td>0.5</td>
<td>0.27</td>
</tr>
<tr>
<td>9</td>
<td>0.1</td>
<td>0.2</td>
<td>7.9</td>
<td>3.6</td>
<td>2.8</td>
<td>2.8</td>
<td>2</td>
<td>2</td>
<td>13.2</td>
<td>5.4</td>
<td>1.9</td>
<td>0.8</td>
<td>12.9</td>
<td>2.1</td>
<td>1.12</td>
</tr>
<tr>
<td>10</td>
<td>1.3</td>
<td>1.1</td>
<td>6.3</td>
<td>3.5</td>
<td>1.4</td>
<td>0.9</td>
<td>1.6</td>
<td>3.7</td>
<td>16.4</td>
<td>4.3</td>
<td>2.1</td>
<td>0.9</td>
<td>13</td>
<td>0.7</td>
<td>0.58</td>
</tr>
</tbody>
</table>

* NO2-N concentration was monitored between 11/03/2011 and 19/01/2012.

Fig. 4. COD/TKN (offline data) for each sub-period.
average NO₃-N concentration in the plug-flow reactor was positively correlated with DO₁ (Fig. 5). Given that the surface aerator in the location of DO₁ probe is manipulated to control the ammonium concentration in the effluent, ammonium peaks trigger the surface aerators to start.

The correlation coefficient between any two of the system variables did not remain stable between the different sub-periods. Fig. 5 shows the correlograms for sub-periods 2 and 5. These sub-periods were characterized by low and high ranges of N₂O emissions and temperature respectively (Table 2). In sub-period 2, the average NO₃-N concentration in the plug-flow reactor was equal to 2.5 mg/L (Table 2) and correlated negatively with the influent flow-rate (~ - 0.63) (Fig. 5). In sub-period 5 the behavior of nitrate concentration (average equal to 2.1 mg/L) was mainly correlated with DO₁ only in sub-periods 4 and 5. However, the profiles of these two variables showed that in the majority of the sub-periods, abrupt and rapid increase of influent flow-rate (i.e., precipitation events) coincided with increase of the NH₄-N. However, the NH₄-N concentration reduced more rapidly in the system than the influent flow-rate. For example, in sub-period 3 the correlation coefficient between NH₄-N in the plug-flow reactor and influent flow-rate was 0.26. However, when days with significant precipitation events (and thus high influent flow-rate) were omitted, the correlation coefficient was equal to 0.58. The latter shows that, in this example, the lack of correlation between these two variables is most likely to be an indication that the interrelationships are not monotonic and that the method is not appropriate to identify complex relationships within the data.

3.3. Hierarchical k-means clustering

The application of hierarchical k-means clustering enabled the categorization of the different ranges of the operating variables and N₂O emissions within each sub-period. Hierarchical k-means clustering analysis was repeated excluding NH₄-N and NO₃-N concentrations in the plug-flow reactor. The results showed that the majority of the data points were allocated to the same clusters for each sub-period even when the NH₄-N and NO₃-N concentrations in the plug-flow reactor were excluded. In the majority of the sub-periods (i.e. sub-periods 1–6) more than 85% of the data points were assigned to the same cluster. It can be concluded that specific patterns and ranges of NH₄-N and NO₃-N monitored in plug-flow reactor, systematically resulted in specific responses to the carrousel reactor. The latter is supported by the Spearman’s rank correlation analysis, where high correlations were observed between the variables in the two reactors for several sub-periods. For example, the correlation coefficient between NH₄-N in the plug-flow and carrousel reactors is higher than 0.7 for sub-periods 1 to 7. The similarity of the clusters for all the sub-periods is shown in Table S4 in the Supporting Material.

The range of N₂O emissions was differentiated in the majority of the sub-periods. In all the sub-periods, two major clusters were identified characterized by significant differences in the NH₄-N and NO₃-N concentrations in the plug-flow reactor. In the majority of
the sub-periods they represented the diurnal variability of the system nutrient concentrations and inflow-flux rate. Additionally, clustering distinguished occasions with high influent flow-rate and ammonium concentration in the carousel reactor, which can be an indication of precipitation events. In sub-periods characterized by low average N₂O emissions (i.e., 1, 2, 7, 8, and 9), clusters with increased N₂O emissions (yet relatively low) were mainly linked to higher loading rates due to the expected diurnal variability or to precipitation events. However, N₂O emissions higher than 3.8 kg/h were observed when the average NO₃-N concentration was constantly lower than 1 mg/L in the plug-flow reactor and the NO₃-N concentration was lower than 4 mg/L in the carousel reactor. Table 3 compares the clustered average values for all the variables in sub-period 2 (average N₂O emissions equal to 0.6 kg/h — Tables 2 and 4) (average N₂O emissions equal to 5.6 kg/h — Table 2). The average value of N₂O emissions for a set of clusters in a specific sub-period (from Table 3) can be found taking into account the number of data-points in the individual clusters. Sub-period 4 was characterized by very low NO₃-N concentration in the middle of the oxic zone in the plug-flow reactor. The latter indicates lower oxidation of ammonia to nitrate or insufficient DO in the plug-flow nitrification lane. This can lead to higher NH₄-N loading in the plug-flow reactor. On the other hand, higher nitrification rates in the plug-flow reactor (i.e. sub-period 2) resulted in lower N₂O emissions in the carousel reactor. The average values of all the variables in each cluster during all the sub-periods are given as supplementary material (Table S5).

In clusters 2 and 16 the averages of operating variables had similar ranges (Table 3). However, in these two occasions the N₂O emissions were different (0.09 and 0.01). Similarly, in clusters 5 and 6 the averages of operating variables were similar yet the emissions were different (0.09 and 0.01). Similarly, in clusters 1 and 2 the averages of operating variables were different but the N₂O emissions were similar (0.21 and 0.24). Such observations indicate the underlying complexities of the interdependencies. Additionally, it can be concluded that the range of N₂O emissions can partially depend on the preceding operational mode of the system. The PCA analysis showed that N₂O emissions’ peaks are related with NH₄-N and N₂O fluxes at 00:00 am. The influent flow-rates, NH₄-N and NO₃-N concentrations in the plug-flow reactor also were similar in these two occasions. The average N₂O fluxes were equal to 0.44 and 2.01 kg/h for occasion 1 and 2 respectively. More extensive data are required for quantitative investigation.

### Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄-N PF</td>
<td>-0.28</td>
<td>0.47</td>
<td>-0.24</td>
<td>0.29</td>
</tr>
<tr>
<td>NO₃-N PF</td>
<td>0.36</td>
<td>0.21</td>
<td>0.14</td>
<td>-0.67</td>
</tr>
<tr>
<td>Influent</td>
<td>-0.38</td>
<td>-0.31</td>
<td>-0.09</td>
<td>-0.37</td>
</tr>
<tr>
<td>NH₄-N C</td>
<td>-0.34</td>
<td>0.03</td>
<td>-0.59</td>
<td>-0.29</td>
</tr>
<tr>
<td>NO₃-N C</td>
<td>-0.04</td>
<td>0.58</td>
<td>0.21</td>
<td>-0.31</td>
</tr>
<tr>
<td>DO1</td>
<td>-0.43</td>
<td>0.06</td>
<td>-0.15</td>
<td>-0.18</td>
</tr>
<tr>
<td>DO2</td>
<td>-0.40</td>
<td>0.08</td>
<td>0.48</td>
<td>-0.17</td>
</tr>
<tr>
<td>DO3</td>
<td>-0.37</td>
<td>0.21</td>
<td>0.40</td>
<td>0.28</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.22</td>
<td>0.49</td>
<td>-0.33</td>
<td>0.11</td>
</tr>
</tbody>
</table>

*NO₂-N concentration was monitored between 11/03/2011 and 19/01/2012.*

## 3.4. Principal component analysis in the carousel reactor

PCA was applied to transform the original correlated measured variables to uncorrelated variables (Principal components) and explain the maximum observed variability. In sub-periods with low emissions (1, 2, 7, 8, and 9) the PCA analysis showed that N₂O emissions’ peaks are related with NH₄-N and influent flow-rate peaks in the carousel reactor and with the effect of the diurnal variability of these variables’ loading rates.

The current section discusses the PCA results for sub-period 2, as an example. The results for all the sub-periods are given in the supplementary material (Tables S6–S13, Figs. S4–S29). The application of PCA reduced the dimensionality of the data with 4 principal components (PCs) explaining ~86% of the total variance (PC1 = 39%, PC2 = 26%, PC3 = 12%, and PC4 = 9%). Loadings for the system variables in the 4 PCs are given in Table 4. The loadings of each component are an indication of the variation in the variables explained by a specific component. Influent flow-rate, ammonium concentration in the carousel reactor (NH₄-N C) and the three DO (DO1, DO2 and DO3) concentrations had the highest negative loadings in PC1. This means that the first principal component increased with the increase of these variables. Nitrate concentration (NO₃-N PF) in the plug-flow reactor has a relatively high positive loading in PC1 (0.36). Therefore, PC1 describes how the
carrousel reactor responds to the behavior of the upstream plug-flow reactor processes and conditions, the variation of the influent flow-rate and variations in ammonium and DO concentrations in the carrousel reactor. The latter can be indirectly connected with the control strategy of the carrousel reactor, since the surface aerators were manipulated based on the effluent ammonium concentration. PC2 linked ammonium concentration in the plug-flow reactor, nitrate concentration in the carrousel reactor and temperature (loadings higher than 0.47). In PC3 ammonium concentration in the carrousel reactor had high negative loading, while DO2 and DO3 concentrations had positive loadings that was not expected considering the control strategy of the system. Investigation of the variables’ profiles, though, showed an increasing trend of DO2 and DO3, whereas the ammonium profile did not present a similar trend.

The biplot of the first 2 PCs is used to visualize the combined behavior of significant variables that affect the system. Data points assigned to cluster 6 (Fig. 7), had negative scores in PC2 and PC1. Therefore, ammonium concentration in the carrousel reactor and influent flow rate were higher than average, while the nitrate concentration in the system was lower than average. Fig. 8 shows the profile of N2O emissions and NH4-N in the carrousel reactor for sub-period 2. The colored points in the diagram represent the identified clusters. Peaks in emissions coincided with peaks in the NH4-N C profile, whereas peaks in NH4-N C coincided with precipitation events (cluster 6).

The scores of the data-points in cluster 5 were mainly positive in PC1 and negative in PC2 (Fig. 7). PC2 increased with the increase of NH4-N concentration in the plug-flow reactor (Table 4). Given that PC2 had an average equal to 0 (data are standardized), data-points with negative scores in PC2 represent occasions with lower than average NH4-N concentration in the plug-flow reactor. This is supported by the correlation plot (Fig. 7), where the arrow of NH4-N concentration in the plug-flow reactor points to the direction of increasing concentrations of NH4-N. Therefore, data-points belonging to cluster 5 were characterized by higher than average ammonium concentration in the plug-flow reactor. Similarly, NO3-N concentration in the plug-flow reactor had relatively significant positive loading in PC1 (0.36 – Table 4). The latter indicates that NH4-N and DO concentrations (measured by three probes) in the carrousel reactor (that had negative loadings in PC1 – Table 5) tended to decrease when NO3-N concentration in the plug-flow reactor increased. Given that all data-points in cluster 5 had positive scores in PC1, it can be concluded that they are characterized by lower than average NH4-N concentration in the carrousel reactor and higher than average NO3-N concentration in the plug-flow reactor. According to the clustering results the latter can be an indication of the high nitrogen loadings of the normal diurnal variability in the reactor. This finding is supported from the results presented in Fig. 8, where the data-points of cluster 5 correspond to the daily low range of ammonium concentrations in both reactors.

Fig. 9 summarizes scores of the PC2 and the respective clusters (colored points in the diagram) indicating strong diurnal cyclic fluctuations of the water quality during this sub-period. It also shows that after each precipitation event, a sudden temperature drop occurred; the system was disturbed and cannot recover immediately. Spearman’s rank correlation coefficient between PC2 and N2O emissions is equal to 0.72.

In sub-period 4, mechanisms triggering high N2O emissions in the carrousel reactor prevailed (average = 5.6 kg/h). The PCA loadings were similar to sub-period 2, while the clustering results indicated 3 clusters; clusters 10 and 11 were affected by the diurnal
variability and cluster 12 was affected by the precipitation events (Table 3). Again, the DO data obtained from the 3 sensors in the carrousel reactor had significant negative loadings in PC1. However, ammonium concentration in the carrousel reactor was not identified as a significant variable affecting the system in the first two PCs. This can be attributed to the fact that less NH4-N concentration peaks were observed in the effluent of the carrousel reactor (17 data points belong to cluster 12). The correlation coefficient of PC1 with NH4-N concentration in the carrousel reactor was −0.75. Therefore, PCA analysis shows that PC1 is a good indicator of the ammonium concentration in the carrousel reactor. The DO concentrations in this sub-period especially for cluster 10 (with average NH4-N concentration in the carrousel reactor equal to 1.26 mg/L) was the highest observed in all the clusters with similar NH4-N concentrations in the carrousel effluent. The alternation of aerobic and anaerobic conditions observed in this reactor, combined with high NH4-N and DO concentrations has been identified as a significant cause of nitrification sourced emissions (Yu et al., 2010).

In PC2, the NO3-N concentration and temperature had significant positive loadings (Table 5). The score plot of PC2 (Fig. 10a) presented an increasing trend and therefore, showed that nitrate and temperature increased. The latter was verified by the profiles of NO3-N concentrations in the carrousel reactor (Fig. 10b) and NO3-N concentration and temperature in the plug-flow reactor (Supplementary material S30). In the beginning of the sub-period 4 very low concentrations of nitrate were observed in the system and they gradually increased especially after the 28th of March. The Spearman’s correlation coefficient between N2O emissions and PC2 scores were relatively high and equal to 0.62. However, contrary to sub-period 2, the clustering analysis showed that there is no nitrate accumulation (Table 3). The average nitrate concentration in the plug-flow reactor was equal to 0.2 mg/L until the 28th of March and increased up to 1.6 mg/L until the end of the sub-period. Therefore, the observations in section 3.3 are supported by the PCA results (low nitrate in the plug flow resulted in increased loadings in the subsequent carrousel reactor and the denitrification activity in the carrousel reactor is affected by the low temperature resulting in nitrite accumulation).

In this section, the combination of hierarchical k-means clustering and PCA was used in order to link the different emission ranges with all the online monitored variables (i.e. Fig. 7). Even though, the online dynamics of significant variables that can trigger N2O emissions in biological processes (i.e. COD, pH) were not available, the applied methodology enabled the identification of a set of variables that are connected with N2O emissions in each sub-period (i.e. Fig. 8). Considering that online data were not available for the influent of the carrousel reactor, higher NH4-N loadings in the carrousel reactor were linked with clusters characterized by higher than average influent flow-rates and ammonium concentration and lower than average NO3-N concentration in the plug-flow reactor. The latter can be supported by the fact that the behavior of variables in the carrousel reactor was significantly dependent on the nutrient concentrations in the plug-flow reactor (Table S4 – clustering results). Additionally, more intense aeration in the carrousel reactor (that can affect the stripping of dissolved N2O) was linked with clusters characterized by higher than average NH4-N concentration in the carrousel reactor (since the surface aerators were manipulated by the effluent ammonium concentration).

3.5. N2O generation pathways

In line with Daelman et al. (2015) findings, both AOB pathways can be considered responsible for the N2O emissions observed in the carrousel reactor. The combination of nitrite accumulation and low oxygen concentrations can be linked with the nitrifier denitrification pathway, whereas higher AOR (ammonia oxidation rate), correlation of NH4-N concentration in the carrousel reactor with N2O emissions and higher DO concentrations can be linked with the hydroxylamine oxidation pathway (Law et al., 2012). N2O generation via heterotrophic denitrification can be also significant especially in periods with nitrate accumulation, suggesting insufficient anoxic conditions (Daelman et al., 2015).

In terms of the offline monitored variables, low pH, accompanied with nitrite accumulation, as observed in sub-period 4 has been identified as a significant factor inhibiting N2O reduction during denitrification (Pan et al., 2012). Zhou et al. (2008) reported that under these conditions the production of free nitrous acid (FNA) in a denitrifying-Enhanced Biological Phosphorus Removal culture was the main contributor to N2O emissions production even at low concentrations equal to 0.0007–0.001 mg HNO2-N/L (nitrite concentration 3–4 mg/L at pH 7). Additionally, high pH values (>7) combined low DO concentration (~0.55 mg/L) have been reported to be responsible for nitrification driven N2O emissions via the nitrifier denitrification pathway (Law et al., 2011). The latter is
attributed to increasing ammonium oxidation rate (due to the pH increase), enhancing the nitrifier denitrification pathway through electrons provision. On the other hand, lower pH (<7) has been linked with elevated nitrification driven N2O emissions at higher DO concentrations (~3 mg/L) (Li et al., 2015). The authors argued, that at higher pH the electrons available from the ammonium oxidation rate are mainly used to form water from molecular oxygen and H⁺. In the current study, the pH in the effluent of the reactor was steady during the monitoring campaign (~8 ± 0.2). However, online pH data showing the exact dynamics of the pH in the carrousel reactor were not available.

Low COD/N ratios have been reported to be responsible for denitrification induced N2O emissions (Schulthess and Gujer, 1996).

Table 5
<table>
<thead>
<tr>
<th></th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₄-N PF</td>
<td>-0.48</td>
<td>0.04</td>
<td>-0.11</td>
<td>0.25</td>
</tr>
<tr>
<td>NO₃-N PF</td>
<td>0.26</td>
<td>0.56</td>
<td>-0.04</td>
<td>-0.35</td>
</tr>
<tr>
<td>Influent</td>
<td>-0.33</td>
<td>-0.07</td>
<td>-0.52</td>
<td>-0.17</td>
</tr>
<tr>
<td>NH₄-N C</td>
<td>-0.28</td>
<td>0.14</td>
<td>-0.50</td>
<td>-0.46</td>
</tr>
<tr>
<td>NO₃-N C</td>
<td>-0.17</td>
<td>0.59</td>
<td>0.32</td>
<td>0.04</td>
</tr>
<tr>
<td>DO1</td>
<td>-0.37</td>
<td>0.24</td>
<td>-0.13</td>
<td>0.59</td>
</tr>
<tr>
<td>DO2</td>
<td>-0.40</td>
<td>0.08</td>
<td>0.41</td>
<td>-0.14</td>
</tr>
<tr>
<td>DO3</td>
<td>-0.37</td>
<td>0.01</td>
<td>0.33</td>
<td>-0.40</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.23</td>
<td>0.51</td>
<td>-0.27</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Fig. 8. Profile of (a) N₂O emissions, (b) NH₄-N concentration in the Carrousel reactor and (c) NH₄-N concentration in the plug-flow reactor for sub-period 2; coloured points indicate the respective clusters.
The offline data showed that COD/TKN ratio in the influent remained relatively steady during the monitoring campaign with a slight decrease in sub-periods 4 and 5 (<5) where emissions were higher (5.6 and 2.6 kg/h respectively). However, low COD/TKN (<5) was also observed in other sub-periods and did not result into high N\textsubscript{2}O emissions (Fig. 4). The frequency of the offline data (~6 days) did not enable the identification of the exact contribution of COD loading to the system. Fig. 4 shows that COD limitation is not the sole contributor to the increased N\textsubscript{2}O emissions in sub-period 4. Therefore, the results indicate that heterotrophic denitrification induced by COD/TN limitation was not the main N\textsubscript{2}O emissions source in sub-periods 4 and 5.

The results from the application of multivariate statistical techniques can be used for the identification and explanation of potential pathways for N\textsubscript{2}O generation. In sub-periods with lower average N\textsubscript{2}O emission fluxes (1, 6, and 7), emission peaks coincided with ammonium peaks in the plug-flow reactor and therefore in the influent carrousel reactor. In that case, average emission fluxes ranged from 0.05 kg/h (sub-period 1) to 2.54 kg/h (sub-period 6). Wunderlin et al. (2012) demonstrated that N\textsubscript{2}O production through hydroxylamine oxidation is accompanied by excess ammonia, low nitrite concentration and high ammonia oxidation rate.
Additionally, in these sub-periods, N₂O emissions were higher at higher temperatures and DO concentrations. The high DO concentrations coincided with peaks in nitrate and nitrite concentrations indicating also insufficient denitrification zones in the reactor. AOB can use nitrite instead of oxygen as electron acceptor (Kampschreur et al., 2009a) especially in oxygen limiting conditions (low DO zones exist even when all surface aerators are under operation); thus, nitrifier denitrification by AOB could potentially contribute in N₂O emissions. Burgess et al. (2002) found strong dependency between nitrite accumulation and N₂O emissions, especially at sudden increase of ammonia loading.

Overall, N₂O emissions increased significantly and peaked at low nitrate concentrations in both reactors (i.e., sub-periods 3 and 4) and high nitrite concentrations in the carrousel reactor (i.e., sub-period 4). Under aerobic conditions, nitrite accumulates in the system when the ammonia oxidation rate to nitrite exceeds the nitrite oxidation rate to nitrate (Guisasola et al., 2005) inducing the system when the ammonia oxidation rate to nitrite exceeds the nitrite oxidation rate to nitrate (Guisasola et al., 2005) inducing the system.

The proposed methodological approach can detect and visualize disturbances in the system (i.e., precipitation events, high NH₄-N concentrations, etc.) and their effect on N₂O emissions.

**Acknowledgements**

This paper is supported by the Horizon 2020 research and innovation programme, SMART-Plant under grant agreement No 690323. The authors acknowledge Alex Sengers and David Philo from Hooghoomaardsgen van Schieland en de Krimpenerwaard, the Water Board of Schieland and Krimpenerwaard, for sharing their knowledge regarding the Kralingseveer WWTP operation.

**Appendix A. Supplementary data**

Supplementary data related to this article can be found at https://doi.org/10.1016/j.watres.2018.04.052.

**References**


