Biochemical Engineering Journal 140 (2018) 17–28

Contents lists available at ScienceDirect

Biochemical Engineering Journal

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

Regular article

Potential of multivariate statistical process monitoring based on the biogas composition to detect free ammonia intoxication in anaerobic reactors

Sébastien Lemaigre[a,b,*], Gilles Adam[a], Patrick A. Gerin[b], Anaïs Noo[a], Bénédicte De Vos[a], Dominika Klimek[a], Xavier Goux[a], Magdalena Calusinska[a], Philippe Delfosse[a]

[a] Environmental Research and Innovation Department (ERIN), Luxembourg Institute of Science and Technology (LIST), Rue du Brill 41, L-4422 Belvaux, Luxembourg
[b] Université catholique de Louvain, Earth & Life Institute, Bioengineering, Croix du Sud 2, box L7.05.19, B-1348 Louvain-la-Neuve, Belgium

HIGHLIGHTS

- Free ammonia nitrogen (FAN) extreme intoxication was reached at $\sim 5\ g_{FAN}\ L_{slurry}^{-1}$.
- Multivariate statistical process control (MSPC) models were built using only the biogas composition.
- Recursively updated MSPC models detected a decrease of reactor performance early.
- Early signs of process perturbation coincided to a drastic change in the reactor microbiome.
- These models performance was negatively affected by a drastic shift in the reactor microbiome.

ARTICLE INFO

Keywords:
Anaerobic digestion
Ammonia intoxication
Multivariate statistical process control
Monitoring model
Principal components analysis
Biogas composition

ABSTRACT

Three anaerobic digestion reactors were initially maintained at steady-state with low nitrogen input, then progressively exposed to increasing nitrogen input until extreme ammonia intoxication, leading to interruption of biogas production. Such interruption occurred for a free ammonia nitrogen (FAN) concentration of about 5 g L_{slurry}^{-1}. A fourth reactor used as a control was maintained at low nitrogen steady-state. The biogas composition (CH_4, CO_2, H_2 and H_2S) was recorded over time and data were fed into a multivariate statistical process control model based on principal components analysis (PCA-MSPC). Static PCA-MSPC failed to distinguish between the actual reactor performance decrease and the drift of the in-control process baseline. However, recursively updated PCA-MSPC (effective memory length of 14 days, i.e. 0.25 hydraulic retention time) allowed detection of the early signs of FAN induced process dysfunction, which were simultaneous with drastic changes in the microbiome of the replicate reactors. However, the process control model was not able to adapt to a drastic modification of the microbiome. Propionate continuously accumulated since the detection of the first signs of process dysfunction.

1. Introduction

In 2014, the European Union leaders agreed on a new climate and energy framework, establishing challenging objectives for the member states for the year 2030. These objectives include a 40% reduction of the greenhouse gas emissions and an increase of the renewable energy

Abbreviations: AD, anaerobic digestion; BMP, biochemical methane potential; E-nose, electronic nose; DNA, deoxyribonucleic acid; FAN, free ammonia nitrogen; FM, fresh matter; GC, gas chromatography; HDS, historical data set; HNR, high nitrogen input reactor; HRT, hydraulic retention time; IPV, individual process variable; L_N, litre of gas normalized to normal conditions of temperature and pressure (0 °C;1013 hPa); LNR, low nitrogen input reactor; MSPC, multivariate statistical process control; OLR, organic loading rate; PC, principal component; PCA, principal component analysis; OTU, operational taxonomic unit; PCA-MSPC, multivariate process control performed on PCA basis; rRNA, ribosomal ribonucleic acid; RU-PCA-MSPC, recursively updated multivariate process control performed on PCA basis; SPE, squared prediction error index; SPE_{lim}, control limit of the SPE index; T^2, Hotelling's T^2 index; T_A^2, T^2 index for the number *A* of PCs retained in the model; TAN, total ammonia nitrogen; TIC, total inorganic carbon; TS, total solids; UCL, upper control limit of the T^2 (or T_A^2) index; VFA, volatile fatty acid; VS, volatile solids

* Corresponding author at: Environmental Research and Innovation Department (ERIN), Luxembourg Institute of Science and Technology (LIST), Rue du Brill 41, L-4422 Belvaux, Luxembourg.

E-mail addresses: sebastien.lemaigre@list.lu (S. Lemaigre), gilles.adam@list.lu (G. Adam), patrick.gerin@uclouvain.be (P.A. Gerin), anais.noo@list.lu (A. Noo), benedicte.devos@list.lu (B. De Vos), dominika.klimek@list.lu (D. Klimek), xavier.goux@list.lu (X. Goux), magdalena.calusinska@list.lu (M. Calusinska), philippe.delfosse@list.lu (P. Delfosse).

https://doi.org/10.1016/j.bej.2018.08.018
Received 12 July 2017; Received in revised form 24 August 2018; Accepted 30 August 2018
Available online 31 August 2018
1369-703X/

contribution to 27% of the global energy mix [1]. To meet these objectives, on-farm anaerobic digestion (AD) can bring a valuable contribution since it is an efficient technology for renewable energy production from wet fermentable organic substrates and for recovery of the essential nutrients, especially nitrogen and phosphorus [2].

Animal manure should be digested in AD reactors to take advantage of its carbon content for energy production, while avoiding the emission of methane during its storage [3]. However, the high nitrogen load present in manure in the form of urea and proteins leads to the accumulation of ammonia as a final product of AD [4]. In addition, the carbon present in manure is mostly recalcitrant to AD because the animal has already assimilated the most available part of the organic matter present in its feed. Consequently, AD plants exploiting manure as the principal substrate must adopt long hydraulic retention times (HRT) to efficiently convert recalcitrant carbon into methane. This exacerbates ammonia accumulation in the reactor slurry.

Excessive concentration of ammonia is a major cause of AD reactor process inhibition [5]. Indeed, the free ammonia nitrogen (FAN: NH_3) is toxic for the AD microbial flora due to its high diffusion rate through cell membranes [3,6]. The FAN proportion in the total ammonia nitrogen (TAN) content of the slurry (TAN = FAN + N-NH_4^+) is driven by the pH value and the operating temperature [6]. FAN inhibition can strongly affect the performance of AD reactors by reducing the organic carbon to CH_4 conversion ratio, causing important economic losses for the plant manager. In this context, process monitoring that allows early detection of FAN inhibition is a major challenge for further development of the AD sector, especially in agriculture.

Classical AD process monitoring involves measuring indicators in the slurry and comparing these measurements to threshold values. However, literature on the FAN inhibition in AD reactors shows a wide range of threshold values [7,8]. The main cause is that the FAN toxicity for a given AD reactor depends of the degree of microbial acclimation to nitrogen-rich substrates [3,9].

Compared to classical methods used for AD process monitoring, multivariate statistical process control (MSPC) methods exploit statistically defined control limits and are thus independent of uncertain toxicity thresholds. MSPC based on principal component analysis (PCA-MSPC) is nowadays the basic tool to monitor multivariate processes [10–12]. In PCA-MSPC, the information contained within the measured individual process variables (IPVs) is reduced to two monitoring indexes through statistical modelling [10]: (1) the T_A^2 index describes the variation amplitude of the IPVs in the "plane" defined by a limited number *A* of principal components (PCs) retained in the model; (2) the squared prediction error (SPE) index quantifies the residuals, i.e. the part of variability unexplained by the model. For each time point, the value of each monitoring index is plotted in a control chart and compared to a control limit above which the process is considered as out-of-control. The *static* PCA-MSPC is the basic way to implement PCA-MSPC. This method consists of calculating the model parameters on the basis of data collected for an initial period during which the process stability is verified. Static PCA-MSPC is easy to implement and requires only a few computations once the model is defined, but is limited in the case of processes that demonstrate a drift of their in-control baseline. To handle such situations, *recursively updated* PCA-MSPC methods (RU-PCA-MSPC) were proposed [13]. RU-PCA-MSPC methods take into account the normal dynamics exhibited by a process through continuous updating of the model parameters.

In AD process, analysing the biogas composition offers multiple advantages as compared to slurry analysis. Biogas analyses allow a high measurement frequency and deliver immediate results [14]. In recent years, the electronic nose (e-nose) technology was the main technology using the biogas composition investigated for AD process monitoring [15,16]. In process monitoring using e-noses, PCA-MSPC techniques are used to interpret the response of an array of low-selectivity gas sensors. Although the e-nose technology is promising, it presents two major limitations. On the one hand, the metal oxide semiconductor sensors (MOS) used in e-noses are subject to drift when exposed to harsh environmental conditions [17,18]. On the other hand, MOS sensors need frequent replacement due to their high sensitivity to poisoning [17]. As an alternative, the major components of the biogas (CH_4, CO_2, H_2 and H_2S) can be analysed using specific sensors that are less subject to drift (if regularly calibrated) and more robust.

Lemaigre et al. [14] recently assessed a *static* PCA-MSPC model for process monitoring of an AD reactor progressively exposed to intoxication by volatile fatty acids (VFAs) using the CH_4, CO_2 and H_2 concentrations in the biogas as IPVs, as well as parameters measured in the slurry. The model delivered valuable warning signals announcing the dysfunction of the process in the reactor, which were mostly derived from the dynamics of the biogas composition. On the other hand, the dynamics of the parameters measured in the slurry contributed poorly to the warning.

The main objective of the study presented below was thus to assess the potential of PCA-MSPC using the CH_4, CO_2, H_2 and H_2S concentrations in the biogas as exclusive IPVs to allow early detection of FAN inhibition in AD reactors. In particular, static and recursively updated PCA-MSPC methods were compared regarding their ability to discriminate the periods that demonstrate stable and altered reactor performance. The experiment described in this study focussed on 100-L semi-continuous laboratory reactors that were successively submitted to substrate acclimation and progressive FAN intoxication until complete process failure. To assess the reliability of the PCA-MSPC model as a process status monitoring tool, a chemical and microbial monitoring of the reactor slurry was conducted.

2. Materials and methods

2.1. Anaerobic digestion

The experiment was performed using four stainless steel tank reactors of 100-L working volume with an additional 25-L headspace volume. The reactor working volume was completely and continuously stirred at 60 rpm with a central anchor-type stirrer.

Each reactor was inoculated with 100 kg of anaerobic sludge [total solids (TS): 0.04 gTS gFM^{-1}; volatile solids (VS): 0.56 gVS gTS^{-1}] collected on April 21, 2015 in the mesophilic anaerobic reactor of an urban waste water treatment plant (Schifflange, Luxembourg).

The digestion was performed in the mesophilic temperature range with the operating temperature set at 37 ± 0.5 °C. The reactors were fed as real digesters with unsterilized substrates. The main fed substrate was dry sugar beet pulp pellets. This substrate was selected for its low nitrogen content (0.012 g N gVS^{-1}; C/N = 39.6) and its complex composition (Table 1), offering a growth environment similar to a real-scale co-digestion reactor to the microbial flora [19]. The biochemical methane potential (BMP) of the sugar beet pulp pellets was evaluated as 0.38 L_N CH_4 gVS^{-1} according to the methods described by Mayer et al. [20]. The reactors were manually fed every legal working day, usually five consecutive days per week, following a semi-continuous scheme [19]. The HRT was adjusted at 56 days during the whole experiment by completing the feed with tap water at 37 °C.

The experimental design stretched over 23 weeks (Fig. 1). The progress over time of the feeding regime of the reactors, expressed in terms of sugar beet pulp, urea and nitrogen input, is presented in Fig. 2a and b. The four reactors were initially acclimatized for 8 weeks (1 HRT), during which their organic loading rate (OLR) was progressively increased to reach a value of 2 gVS L^{-1} per feeding day (10 gVS L^{-1} $week^{-1}$), using sugar beet pulp as the sole substrate and equivalent to a nitrogen supply of 0.024 g N L^{-1} per feeding day (0.12 g N L^{-1} $week^{-1}$).

From week 8, urea powder (Sigma-Aldrich, USA) was progressively added to the feed of three reactors (high nitrogen input reactors 1, 2 and 3; HNRs) while the feed of the remaining reactor (low nitrogen input reactor; LNR) was left unchanged (Fig. 2a and b). Urea was

Table 1
Composition of the sugar beet pulp pellets used to feed the anaerobic reactors, as measured by Adam et al. [15].

	Dry basis (%)	As fed (%)
Total solids		86.9
Moisture		13.1
Volatile solids	92.1	80.0
Ash	7.9	6.9
Protein	8.2	7.2
Crude fibre	17.0	14.7
Acid detergent fibre (ADF)	26.7	23.2
Total digestible nutrients (TDN)	69.4	60.3
Fat	1.1	1.0
Nitrogen free extract (NFE)	66.1	57.4
Calcium	1.0	0.9
Phosphorus	0.1	0.1
Potassium	0.6	0.5
Reducing sugars	2.6	2.3
Sucrose	9.0	7.8
Total sugars as invert (TSI)	8.2	7.2

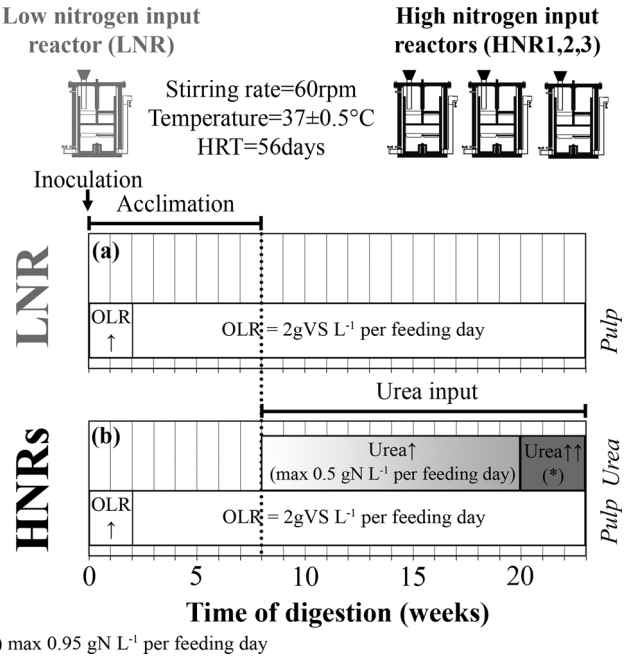

Fig. 1. Schematics of the experimental design. Organic loading rate (OLR) as sugar beet pulp and nitrogen load as urea for the low nitrogen input reactor LNR (a), and for the high nitrogen input reactors HNRs (b). The hydraulic retention time (HRT) was adjusted at 56 days with tap water at 37 °C in the feed.

selected as an additional nitrogen source for the high nitrogen input reactors due to its ease of degradation to ammonia by a large diversity of micro-organisms [21] and because it is a major form of nitrogen in farm manure. For each feeding day, the high nitrogen input reactors were fed with the same mass of sugar beet pulp as the low nitrogen input reactor (Fig. 2a). From week 20 onwards, large amounts of urea (maximal nitrogen supply of 0.95 g_N L_{slurry}^{-1} per feeding day) were added to the feed of the high nitrogen input reactors until almost complete interruption of their biogas production was observed (end of week 22).

2.2. Analytical monitoring

2.2.1. Biogas monitoring

The biogas produced by each reactor passed through a cooling unit at ~8 °C to remove excess water vapour and was collected in an 80-L gas bag (Tecobag, Tesseraux, Germany).

Every 2 h, the content of the bag was automatically pumped through a recirculation loop for 4 min to allow proper mixing in the sampling tubing (Tygon, VWR International, USA). About 50 mL of gas was then pumped from the recirculation loop to a three-channel gas chromatograph (CompactGC, global analyser solutions™, Interscience, Belgium) for analysis of the CH_4, CO_2, H_2, H_2S, O_2 and N_2 concentrations. The gas chromatograph was equipped with a thermal conductivity detector on each channel. The detectors were heated at 80 °C and the filaments at 110 °C. The first channel was equipped with a RI-QBond column (10 m × 0.32 mm) and allowed separation and analysis of CO_2. The second channel was equipped with a Rtx-1 column (30 m × 0.32 mm) and allowed separation and analysis of H_2S. The third channel was equipped with a RI-QBond pre-column (3 m × 0.32 mm) followed by a Molsieve 5 A column (7 m × 0.32 mm), the latter allowing separation of the H_2, O_2, N_2 and CH_4 gases. The RI-QBond pre-column allowed separation of CO_2 from the other gases to protect the Molsieve column from CO_2 poisoning; then the H_2, O_2, N_2 and CH_4 gases were separated in the Molsieve column while the CO_2 was back-flushed through the RI-QBond pre-column and eliminated through the GC exhaust. Helium was used as the carrier gas for the first two channels. Argon was used as the carrier gas for the third channel. For the three channels, the elution was performed under isothermal conditions at 50 °C and with a flow of 10 mL min^{-1}.

The remaining content of the gas bag was pumped through a drum-type wet gas meter (TG-5, Ritter, Germany) for volume measurement, which was normalized to normal conditions of temperature and pressure (0 °C; 1013 hPa). Between each analysis, the recirculation loop was flushed with ambient air for 2 min. After each measurement cycle, ambient air was pumped through the GC to prevent potential corrosion due to H_2S.

During weeks 3, 8 and 17, the CH_4 and CO_2 contents of the biogas were measured using specific non-dispersive infrared sensors (Dynament, UK) as an alternative during compact GC maintenance.

2.2.2. Slurry monitoring

The total ammonia nitrogen (TAN, g L_{slurry}^{-1}), total inorganic carbon (TIC, $gCaCO_3$ L_{slurry}^{-1}), individual VFA concentrations (C2 to C6, g kg_{slurry}^{-1}) and pH were measured in the slurry using the analytical methods described in Goux et al. [19]. The pH was measured on a daily basis (feeding days) while the individual VFA, TIC and TAN contents were measured weekly (every Monday). The FAN content in the slurry (g L_{slurry}^{-1}) was calculated on the basis of the TAN content and the pH using the formulae provided by Calli et al. [6], with an operating temperature fixed at 37 °C (pKa = 8.89). The FAN content was only calculated when the pH was measured (feeding days), with interpolated TAN values being used if TAN was not measured on the same day. Over the experiment, aliquots of around 200 µL of slurry were sampled every week (on Wednesdays), directly frozen in liquid nitrogen and stored at -80 °C prior to a 16 s rRNA gene amplicon high throughput sequencing targeting a monitoring of both bacteria and archaea. Based on the reactor performances toward progressive FAN intoxication and PCA-MSPC results, samples for weeks 0, 6, 7, 8, 10, 11, 12, 14, 18, 20, 21 and 22 were selected for the microbial community analysis. For that purpose, DNA from the selected slurry samples was extracted using the DNeasy PowerSoil Kit (Qiagen) according to the manufacturer's protocol. The 16S rRNA amplicon libraries were then prepared, sequenced, de-multiplexed and quality trimmed as detailed in [22] and [23].

2.2.3. Individual process variables (IPVs) and process stability indicators

The CH_4, CO_2, H_2S and H_2 concentrations in the biogas were used as IPVs to build the PCA-MSPC models [14]. The pulp to CH_4 conversion ratio as well as the TAN, TIC, FAN and VFA individual contents of the slurry were used as process stability indicators, but were not used for building the PCA-MSPC models. The sugar beet pulp to CH_4 conversion ratio (L_N CH_4 gVS^{-1}) was chosen as the principal process efficiency

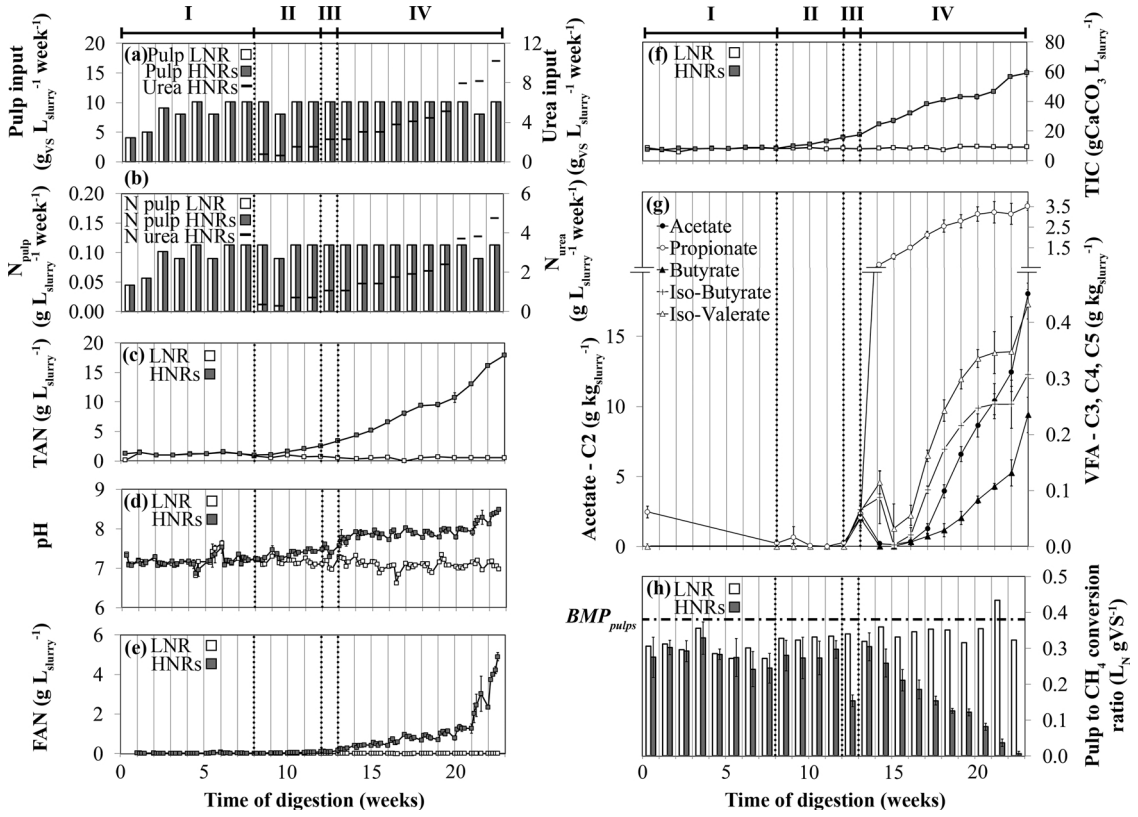

Fig. 2. Progress over time of: the loading rate in term of sugar beet pulp and urea (a), the load of nitrogen introduced weekly in the reactors (nitrogen from both sugar beet pulp and urea) (b); the total ammonia nitrogen (TAN) (c); the pH (d); the calculated free ammonia nitrogen (FAN) (e); the total inorganic carbon (TIC) (f); the volatile fatty acids (VFAs) individual concentrations in the slurry (g); the pulp to CH_4 conversion ratio (g). For the high nitrogen input reactors (HNRs), the data are expressed as the average of the 3 reactors ± the standard error (error bars). LNR: low nitrogen input reactor.

indicator. This ratio was calculated for each week using the normalized total volume of CH_4 produced divided by the total mass of sugar beet pulp (expressed in VS) introduced during the week.

2.3. PCA-MSPC using the biogas composition

For each reactor, the biogas composition measurements collected from week 2 to the end of week 6 were used as a historical data set (HDS). The model parameters were calculated from this initial stable process period. PCA-MSPC was then performed for the next biogas composition measurements (12 measurements per day) using: (1) static PCA-MSPC and (2) RU-PCA-MSPC. The calculations were performed using the Scilab open-source software [26].

2.3.1. Initial model building on the basis of the HDS

The outliers were first removed from the HDS using the iterative method described by Lemaigre et al. [14]. The parameters characterizing the initial stable process baseline were calculated on the basis of the cleaned HDS: average value and standard deviation of each IPV; loading vectors and eigenvalues associated with a limited number A of PCs retained in the model on the basis of the eigenvalue > 1 criterion.

2.3.2. Static PCA-MSPC

For each new biogas composition measurement, static PCA-MSPC was performed using the parameters calculated based on the outlier-cleaned HDS, without any update of the model parameters. For that purpose, the T_A^2 index and its corresponding SPE index were calculated and plotted in a control chart in which they were compared to their respective upper control limits.

The upper control limit of the T_A^2 control chart (UCL) was calculated using:

$$UCL = \left[\frac{A(m+1)(m-1)}{m(m-A)}\right] F(\alpha; A; m-A) \quad (1)$$

where A is the number of PCs retained in the model, m is the number of biogas composition measurements included in the outlier-cleaned HDS and $F(\alpha; A; m-A)$ is the upper α^{th} percentile of the Fisher distribution with parameters A and $(m-A)$. The upper control limit of the SPE control chart (SPE_{lim}) was calculated according to Kourti and MacGregor [11]. The upper control limits of the T_A^2 and the SPE indexes were calculated using a 99% confidence level.

The values of the T_A^2 and SPE indexes corresponding to each new biogas composition measurement were calculated using the following method:

(a) Each IPV was centred and scaled using the average value and the standard deviation calculated on the basis of the outlier-cleaned HDS.

(b) The T_A^2 index was calculated on the basis of the scores obtained for the A PCs retained in the model using:

$$T_A^2 = \sum_{i=1}^{A} \frac{t_i^2}{s_{t_i}^2} \quad (2)$$

where t_i and s_{t_i} are, respectively, the score calculated on the i^{th} PC and the standard deviation related to this PC.

(c) The SPE index was calculated for each sample using:

$$SPE = \sum_{i=1}^{p} (\hat{x}_i - x_i)^2 \quad (3)$$

where p is the number of IPVs, x_i is the observed value for the i^{th} IPV and $\hat{x}_i$ is the corresponding value computed using the PCA model.

For each out-of-control situation detected by the T_A^2 and the SPE control charts, the contribution of each IPV was computed and plotted in a contribution plot. We refer to Lemaigre et al. [14] for an exhaustive

description of the methods and formulae used to build the T_A^2 and the SPE contribution plots.

2.3.3. RU-PCA-MSPC

For each new biogas composition measurement, the model parameters defined on the basis of the outlier-cleaned HDS were recursively updated using the method proposed by Li et al. [13]:

(a) Recursive update of the mean and the standard deviation.
(b) Recursive update of the correlation matrix.
(c) PCA on the basis of the updated correlation matrix.
(d) Selection of PCs using the eigenvalue > 1 criterion.
(e) T_A^2 and SPE calculation using the updated loading vectors and eigenvalues.
(f) Update of the control limits of the T_A^2 and SPE indexes (99% confidence level).

The previous biogas composition measurements were exponentially ignored, giving more weight to the new ones. For that purpose, a forgetting factor β ($0 \leq \beta \leq 1$) was introduced into the recursive calculation of the PCA model. The effective memory length N_{eff}, giving an idea of the number of data that are accounted for in the model, is given by $N_{eff} = \frac{1}{1-\beta}$

To avoid biasing the updated process baseline with new biogas composition measurements corresponding to out-of-control process situations, the recursive update of the PCA model was only performed for new biogas composition measurements corresponding to in-control values of the T_A^2 and SPE indexes ($T_A^2 \leq$ UCL and SPE $\leq SPE_{lim}$). For new biogas composition measurements corresponding to out-of-control values of the T_A^2 or the SPE index, the values of the T_A^2 and the SPE indexes were therefore calculated using the model parameters resulting from the previous model update.

We refer to Li et al. [13] for an exhaustive description of the methods and formulae exploited for the recursive update of PCA-MSPC models using a forgetting factor.

3. Results and discussion

3.1. Anaerobic digestion process efficiency and evolution of the microbial community composition

The evolution over time of the CH_4, CO_2, H_2 and H_2S concentrations in the biogas, exploited as IPVs to build the process monitoring models, is presented in Fig. 3 for the four reactors. In relation to these parameters, the changes over the whole experiment of the bacterial and archaeal communities highlighted by high-throughput 16S rRNA amplicon sequencing are also presented in Fig. 3.

For the low nitrogen input reactor, the H_2 concentration remained very low during the whole experiment (Fig. 3b) and the $[CH_4]/[CO_2]$ ratio showed a weekly pattern related to the feeding mode but CH_4 remained the dominating gas component (Fig. 3a). The reactor efficiency remained above 0.3 L_N CH_4 gVS^{-1} (Fig. 2h). The H_2S concentration in the biogas increased during the experiment (Fig. 3b) probably due to sulfur accumulation in the slurry associated with the feeding of pulp over a 56-day HRT (the sulfur content of the sugar beet pulp is 0.4 mgS gVS^{-1}).

The three high nitrogen input reactors showed well synchronized behaviour regarding the efficiency of conversion of pulp into CH_4 (Fig. 3c,e and g). The pulp to CH_4 conversion ratio was used as the main process efficiency indicator since it reflects well the main purpose of the AD process, which is to convert organic matter into CH_4.

The evolution over time of the pulp to CH_4 conversion ratio (Fig. 2h), the biogas composition and the process stability indicators measured in the slurry allowed the definition of 4 successive distinct phases characterizing the reactor process status (Table 2).

During phase I (weeks 0–7), no urea was fed and the pulp was efficiently converted into CH_4 (Fig. 2h). Short stirrer blocking events occurred for the high nitrogen input reactors 1 and 2 and led to clear increase of the H_2 concentration in the biogas (Fig. 3d and f). During phase I, no drastic change in the bacterial and archaeal communities of the four studied reactors could be identified, with respectively *Bacteroidetes*, *Cloacimonetes*, *Firmicutes* and *Proteobacteria* being the dominant bacterial phyla and *Methanosaeta* and "*Candidatus* Methanofastidiosum" being the dominant archaeal genera (Fig. 3). These different groups of organisms are commonly found in anaerobic mesophilic reactors [19,23].

During phase II (weeks 8 to 11, start of urea supply), the three high nitrogen input reactors maintained stable performance while urea was introduced in the feed. The pulp to CH_4 conversion ratio was not affected (Fig. 2h) and FAN remained at low concentration in the slurry (Fig. 2e). No important change occurred in the microbial community, except that the relative abundance of *Bacteroidetes* (for bacteria) and *Methanomassiliicoccus* (for archaea) increased over time while at the same time the relative abundance of the archaeal "*Candidatus* Methanofastidiosum" decreased from around 40% to 20% in all studied reactors (Fig. 3). The supply of urea contributed to increase the TIC content in the slurry (supplementary material, Fig. S1), but did not affect the CH_4 production (Fig. 3). This is consistent with a previous report [24] and with urea hydrolysis that releases CO_2 ($H_2N\text{-}CO\text{-}NH_2 + H_2O \rightarrow CO_2 + 2\ NH_3$ [25]). CO_2 accumulates in the slurry due to the increasing pH, as a consequence of ammonia accumulation. Urea does not generate H_2 or reducing equivalents and does not contribute stoichiometrically to hydrogenotrophic methane production.

During phase III (week 12), an acute process disturbance was observed for the three high nitrogen input reactors, marked by a sudden drop of the pulp to CH_4 conversion ratio (Fig. 2h). Simultaneously, the $[CH_4]/[CO_2]$ ratio showed values < 1 (Fig. 3c,e and g) and all the measured VFAs accumulated in the slurry, especially acetate (Fig. 2g). A short stirrer blocking event occurred for the high nitrogen input reactor 1 and led to clear increase of the H_2 concentration in the biogas (Fig. 3d).

During phase IV (weeks 13 to 22), FAN accumulation in the slurry became clearly detectable (Fig. 2e). After week 13, a drastic change appeared in the microbial community profile for the three HNRs while the microbial community remained stable for the LNR (Fig. 3). In the three HNRs, *Firmicutes* replaced *Bacteroidetes* in the bacterial community while *Methanosarcina* replaced the two previously dominant archaeal genera *Methanomassiliicoccus* and "*Candidatus* Methanofastidiosum". This total reshaping of the microbial community clearly indicates that the anaerobic digestion process suffered from an acute disturbance between week 12 and 13. From week 13 onwards the AD process was no longer run by the microbes that prevailed during weeks 0 to 12. In consequence, even if the three high nitrogen input reactors showed first transient signs of process recovery, a progressive process performance decline followed thereafter (from week 14 onwards). Indeed, the pulp to CH_4 conversion ratio of the three high nitrogen input reactors initially increased (week 13), but then continuously decreased until the end of the experiment (Fig. 2h). During week 13, the concentrations of acetate and butyrate in the slurry initially decreased, while iso-butyrate, iso-valerate and, especially, propionate continued to accumulate (Fig. 2g). During week 14, the iso-valerate and iso-butyrate concentrations in the slurry temporarily decreased. From week 15 onwards, all the measured individual VFAs accumulated in the slurry until the end of the experiment (Fig. 2g). The consumption of most VFAs and their likely conversion into CH_4 during weeks 13 and 14 might be correlated with the transient improvement of the process efficiency as evaluated by the pulp to CH_4 conversion ratio observed in the three high nitrogen input reactors. Indeed, the microbial conversion of the acetate that accumulated in the slurry during week 12 could explain ~27% of the CH_4 production observed during week 13. This suggests that the ability of the high nitrogen input reactors to convert the fed

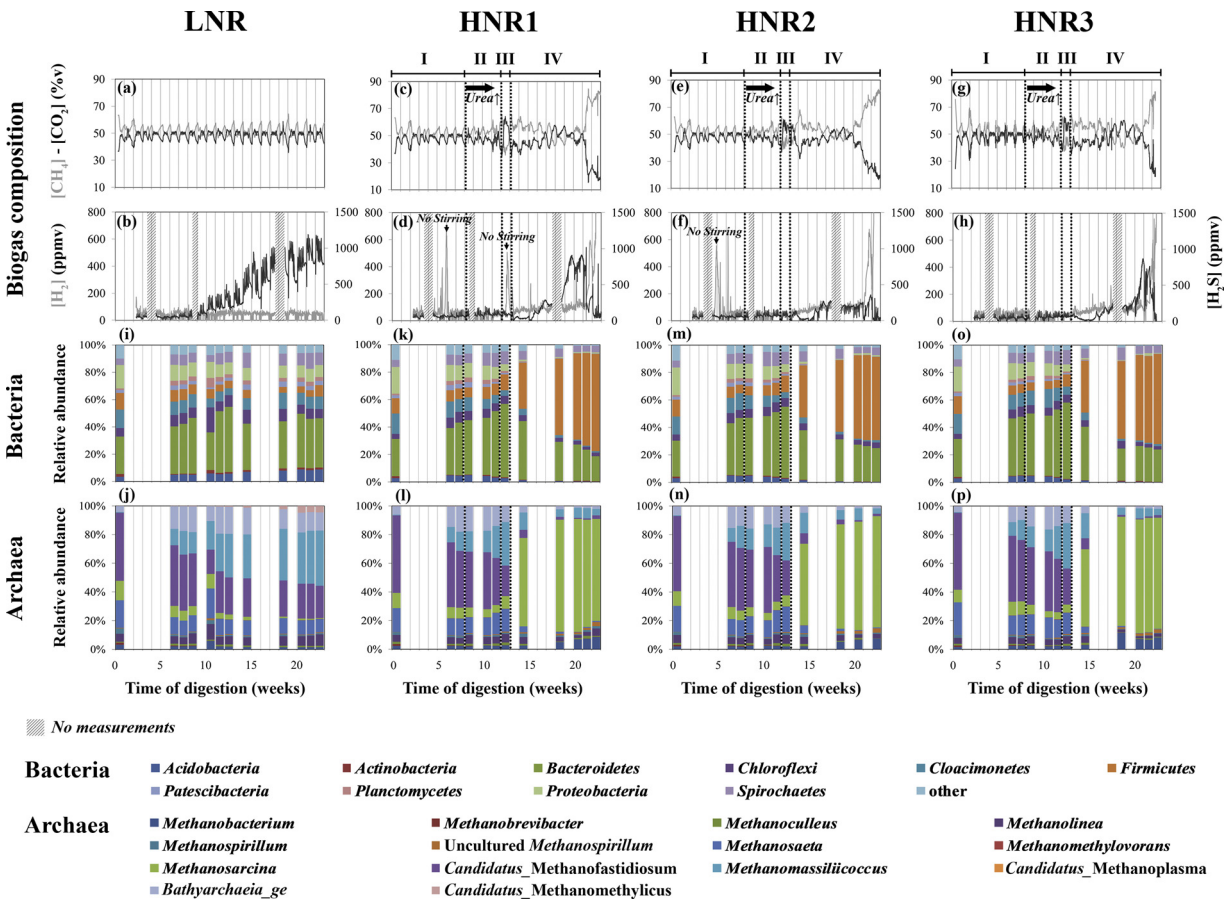

Fig. 3. Progress over time of the biogas concentrations in CH_4, CO_2, H_2 and H_2S for: the low nitrogen input reactor LNR (a, b), the high nitrogen input reactor HNR1 (c, d), the HNR2 (e, f) and the HNR3 (g, h). Biogas composition measurements could not be performed during three short periods for H_2 and H_2S due to gas chromatograph maintenance (hatched grey area). Short stirrer blockings events occurred for HNR1 and HNR2, leading to transient increase of the H_2 production. Bacterial (phylum level) and archaeal (genus level) diversity dynamics over time as assessed by high-throughput 16S rRNA amplicon sequencing for the LNR (i, j), the HNR1 (k, l), the HNR2 (m, n) and the HNR3 (o, p). The presented data correspond to slurry samples collected on Wednesdays. I, II, III and IV correspond to the phases described in Table 2.

Table 2
Process status of the high nitrogen input reactors.

Phase	Time period	Process status of the high nitrogen input reactors
I	Week 0 to week 7	Stable process (no urea provided)
II	Week 8 to week 11	Process status unaffected by urea
III	Week 12	Acute process disturbance
IV	Week 13 to week 22	Apparent process status improvement followed by progressive establishment of complete process failure

sugar beet pulp into CH_4 remained altered during week 13, and that the process status improvement observed during that period was only transient. This assumption is supported by the fact that volatile solids accumulated in the slurry of the high nitrogen input reactors from week 12 onwards (supplementary material, Fig. S2). This also indicates a limitation of the pulp to CH_4 conversion ratio as a direct indicator of the process status since accumulation of methanogenic substrates in the reactor, in this case acetate, can allow continued release of CH_4 when the microbial community adapts to or recovers from harsh environmental conditions.

For the three high nitrogen input reactors, propionate abruptly accumulated in the slurry during week 12 (phase III) and its concentration continued to increase until the end of the experiment, while the concentrations of the other VFAs decreased during weeks 13 and 14 (Fig. 2g). Zhou and Qiu [27] and Calli et al. [6] have already observed that the acetogenic bacteria that use propionate as a substrate particularly suffer from inhibition caused by FAN. Thus, the accumulation of propionate could be a useful factor for FAN intoxication diagnosis as the acetotrophic methane production is the first methane production pathway hampered by high FAN content [28].

During early phase IV, the alkaline pH values due to FAN accumulation may have contributed to a significant sulphide sequestration in the slurry of the high nitrogen input reactors while lower pH in the low nitrogen input reactor allowed the process to eliminate this toxic compound via the biogas extracted from the system. The large release of gaseous H_2S observed during the last weeks of phase IV for the high nitrogen input reactors 1 and 3 (Fig. 3d and h) was not associated with a pH drop (Fig. 2d) and therefore can only be explained by a large increase of the total sulphide content in the slurry. Extreme FAN intoxication may have caused a high microbial death rate, causing cellular lysis and release of sulphur-containing microbial proteins in the slurry. Indeed, several operational taxonomic units (OTUs) were no longer detected during the microbial monitoring (data not presented). At the alkaline pH observed in phase IV, sulphide is mostly in the form of HS^- (~90%) while the main toxic form H_2S is not dominant (~10%) [29]. Nevertheless, the presence of the H_2S toxic form in the slurry might have added to the FAN toxicity encountered by the microbial flora, causing cumulated inhibition by NH_3 and H_2S.

After the urea addition was started (week 8), the TIC content in the slurry continuously increased for the three high nitrogen input reactors (Fig. 2f), indicating that a major part of the CO_2 released by the

Table 3
Outliers cleaned from the raw historical data sets (HDS) collected for the four reactors from week 2 to week 6 (concentrations of CH_4, CO_2, H_2 and H_2S in the biogas used as individual process variables).

Reactor	Biogas composition measurements (raw HDS)	Outliers cleaned from the raw HDS	
		Number of outliers	Percentage of outliers in raw HDS (%)
LNR	311	9	3
HNR1	309	58	19
HNR2	308	50	16
HNR3	297	27	9

LNR: Low nitrogen input reactor; HNR: High nitrogen input reactor.

degradation of the pulp and the urea remained trapped in the slurry as bicarbonate due to the pH increase (Fig. 2d).

3.2. Static PCA-MSPC based on the biogas composition

The potential of static PCA-MSPC to deliver valuable warning signals of process dysfunction was assessed by analysing the dynamics of the biogas composition. The concentrations of CH_4, CO_2, H_2 and H_2S in the biogas were used as IPVs.

Table 3 shows the results of the outlier cleaning performed on the raw historical data sets (HDSs) collected for the four reactors from week 2 to week 6 (during the acclimation period). The outlier-cleaned HDSs were used to build the static PCA-MSPC models of the four reactors. The results of the PCA are presented in Table 4 and Fig. 4.

The loadings of the IPVs regarding the two first PCs were similar for the four PCA models (Fig. 4). The CH_4, CO_2 and H_2 concentrations in the biogas mainly contributed to PC1 while the H_2S concentration in the biogas mainly contributed to PC2. One PC was retained in the model built for the low nitrogen input reactor and the high nitrogen input reactor 2. These PCs explained, respectively, 68.3% and 63.1% of the variance exhibited by the process during the initial model building period (Table 4). Two PCs were retained in the models built for the high nitrogen input reactors 1 and 3. These PCs explained, respectively, 87.3% and 88.0% of the variance exhibited by the process during the initial model building period. For all the reactors, PC1 was highly significant (eigenvalue > 2) but PC2 was close to the significance limit (eigenvalue ~ 1). This explains why PC2 was integrated in the model for the high nitrogen input reactors 1 and 3, while it was not for the low nitrogen input reactor and the high nitrogen input reactor 2 (PC2 eigenvalue < 1, Table 4).

Table 4
Contribution of each principal component (PC) in the PCA performed for each reactor on the basis of the historical data sets cleaned from outliers (concentrations of CH_4, CO_2, H_2 and H_2S in the biogas used as individual process variables).

		PC1	PC2	PC3	PC4
LNR	Eigenvalues	2.732*	0.857	0.411	0.000
	Proportion of Variance	0.683	0.214	0.103	0.000
	Cumulative Proportion	0.683	0.897	1.000	1.000
HNR1	Eigenvalues	2.477*	1.016*	0.506	0.000
	Proportion of Variance	0.619	0.254	0.127	0.000
	Cumulative Proportion	0.619	0.873	1.000	1.000
HNR2	Eigenvalues	2.522*	0.972	0.506	0.000
	Proportion of Variance	0.631	0.243	0.126	0.000
	Cumulative Proportion	0.631	0.874	1.000	1.000
HNR3	Eigenvalues	2.513*	1.008*	0.479	0.000
	Proportion of Variance	0.628	0.252	0.120	0.000
	Cumulative Proportion	0.628	0.880	1.000	1.000

LNR: Low nitrogen input reactor; HNR: High nitrogen input reactor. * PCs retained to build the static PCA models on the basis of the eigenvalue > 1 criterion.

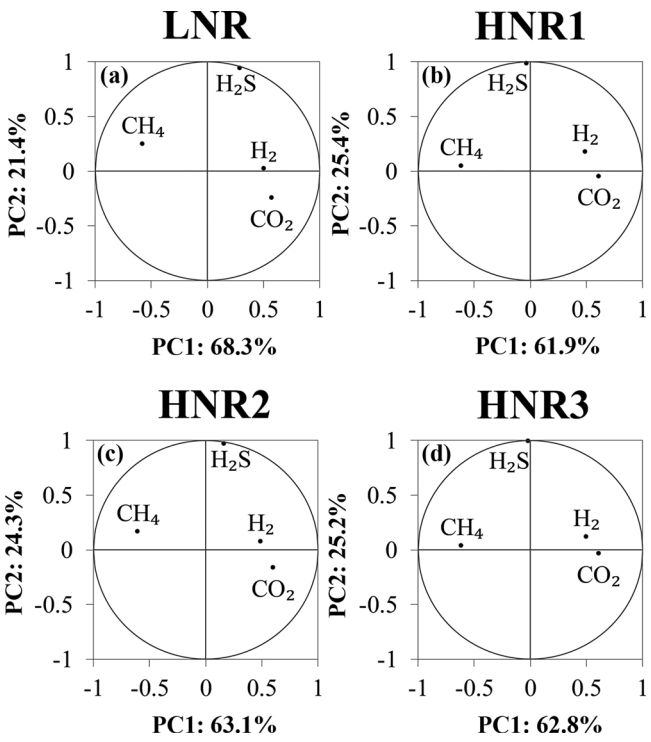

Fig. 4. Principal component analysis (PCA) performed for the four reactors using the biogas concentrations in CH_4, CO_2, H_2 and H_2S measured during their acclimation period (week 2 to week 6) as individual process variables. Loading plots of the 2 first principal components (PCs) for: the low nitrogen input reactor LNR (a), the high nitrogen input reactor HNR1 (b), the HNR2 (c) and the HNR3 (d).

The resulting T_A^2 and SPE control charts are presented in Fig. 5 for the four reactors, together with the pulp to CH_4 conversion ratio as an indicator of process efficiency.

The increases of the H_2 concentration related to the short stirrer blocking events that occurred for the high nitrogen input reactors 1 (Fig. 3d) and 2 (Fig. 3f) were detected by the T_A^2 control chart for the high nitrogen input reactor 1 (Fig. 5c) and by the SPE control chart for the high nitrogen input 2 (Fig. 5f). This shows that the number of PCs that are retained in the model (2 and 1 PCs for the high nitrogen input reactors 1 and 2, respectively) is not crucial for the detection of the out-of-control process situations. Indeed, the SPE control chart complements the T_A^2 control chart by controlling the amplitude of the model residuals.

For the low nitrogen input reactor, both the T_A^2 and the SPE indexes showed a drift in their out-of-control range, starting at week 9 (Fig. 5a and b). According to the individual contributions, the H_2S concentration in the biogas was the main IPV responsible for these alarms. The cause was probably the increase of the H_2S concentration observed for the low nitrogen input reactor from week 9 till the end of the experiment (Fig. 3b). However, no reduction of the reactor performance was observed while the H_2S concentration increased. The T_A^2 and SPE out-of-control status is therefore interpreted as a drift of the in-control process baseline that causes false alarms.

These results show the limitations of static PCA-MSPC in distinguishing between an actual performance decrease and a drift of the reactor in-control process baseline. Adequate update of the model parameters should therefore be performed to take into account the evolution over time of the in-control process baseline. Such an approach is evaluated below using recursively updated PCA-MSPC.

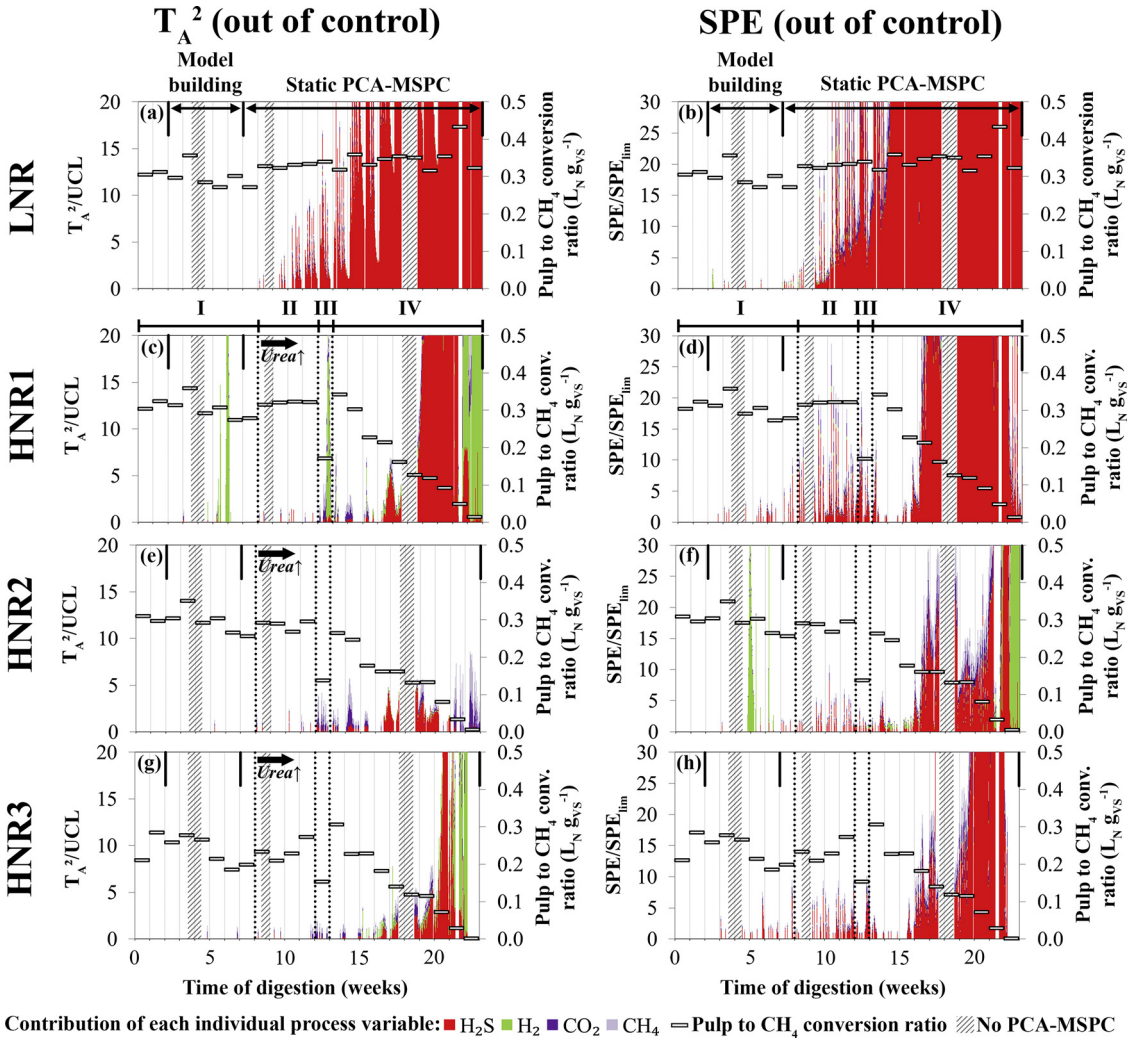

Fig. 5. Static PCA-MSPC performed for the four reactors on the basis of the biogas concentrations in CH_4, CO_2, H_2 and H_2S collected during their acclimation period (week 2 to week 6) as individual process variables (IPVs). According to the constraint set to eigenvalue > 1, one PC was retained in the monitoring models built for the LNR and the HNR2 while two PCs were retained in the monitoring model built for the HNRs 1 and 3 (Table 3). T_A^2 and squared prediction error (SPE) control charts built for the LNR (a, b), the HNR1 (c, d), the HNR2 (e, f) and the HNR3 (g, h). The T_A^2 and the SPE are expressed as a ratio between the index value and its control limit (UCL and SPE, respectively). Only T_A^2/UCL and SPE/SPE_{lim} > 1 are plotted, indicating out-of-control process (coloured columns). The relative contribution of each IPV for each out-of-control process value is identified by a different colour. When the T_A^2/UCL or SPE/SPE_{lim} are out of the graphic scale, the colour indicates the main IPV contributor while the others have negligible contribution. The pulp to CH_4 conversion ratio is presented for each control chart as the process efficiency indicator. PCA-MSPC was not performed during three periods due to gas chromatograph maintenance (hatched grey area). I, II, III and IV correspond to the phases described in Table 2.

3.3. Recursively updated PCA-MSPC based on the biogas composition

PCA-MSPC with recursive update of the model parameters (RU-PCA-MSPC) was assessed for its ability to distinguish between drift of the in-control process baseline and real signs of decreasing reactor performance.

First, the effective memory length N_{eff}, describing the rate of update, was selected on the basis of the data set from the low nitrogen input reactor. The aim was to maintain the false alarm rate as low as possible, given the stable pulp to CH_4 conversion ratio exhibited by this control reactor during the experiment (Fig. 2h). Three N_{eff} values were compared, corresponding to: *fast* model update (1 day: 12 biogas composition measurements), *moderately fast* model update (14 days: 168 biogas composition measurements corresponding to 0.25 HRT) and *slow* model update (28 days: 336 biogas composition measurements corresponding to 0.5 HRT). The T_A^2 and SPE control charts obtained for each new biogas composition measurement collected following the initial model building period (week 2 to week 6), as well as the rate of the model update and the number of PCs retained in the model, are presented in Fig. 6 for the three N_{eff} values.

Fig. 6c and f show that each model update can induce the modification of the number of PCs retained in the model. However, the control limit (SPE_{lim}) of the SPE control chart is recalculated for each model update. As for the static models, the out-of-control process situations detected by the SPE control chart are thus complementary to those detected by the T_A^2 control chart. It can therefore be assumed that the change in the number of PCs retained in the model has only a minor impact on the detection of the out-of-control process situations detected by the recursively updated PCA-MSPC models.

The T_A^2 and SPE control charts built using the *fast* model update (Fig. 6a and b) and the *slow* model update (Fig. 6g and h) are similar to those obtained using static PCA-MSPC. Indeed, the model was updated only for new biogas composition measurements corresponding to in-control values of the T_A^2 and SPE indexes. For the *fast* model update, much more weight was attributed to the most recent individual measurements than the oldest ones. Therefore, the model parameters were

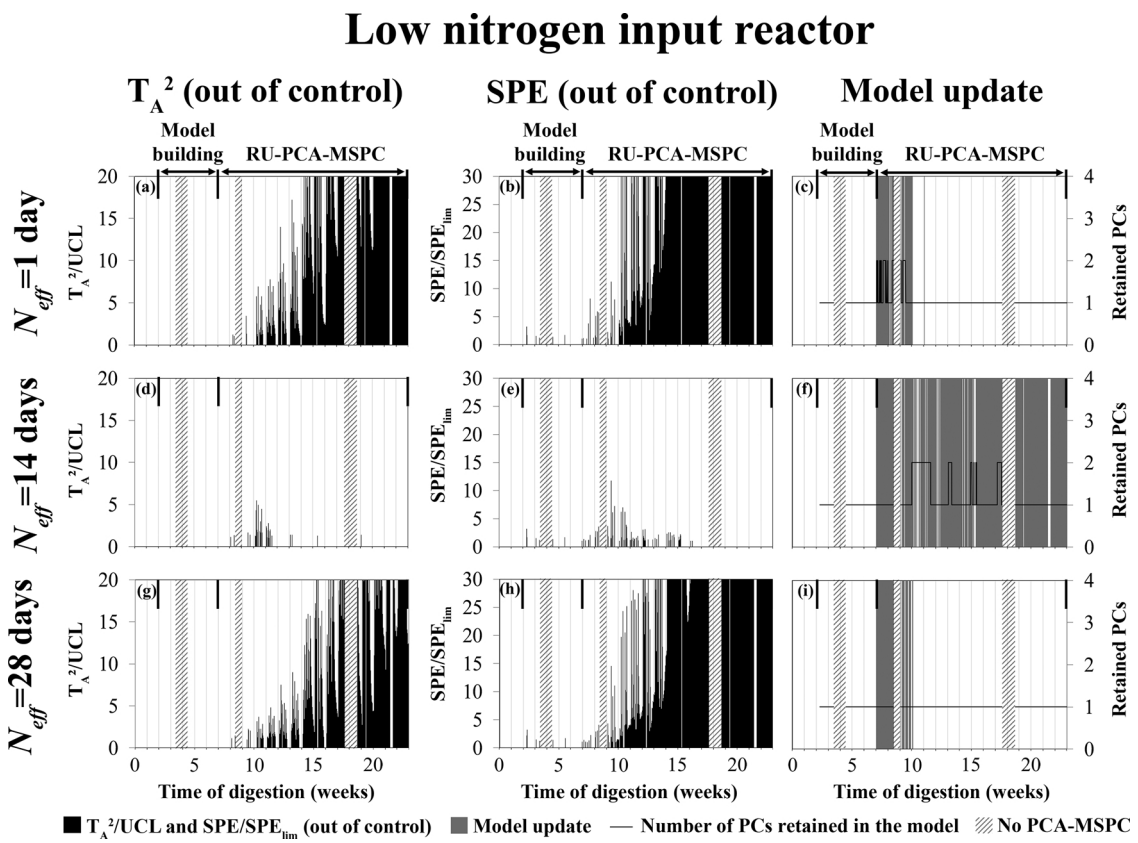

Fig. 6. Influence of the model effective memory length N_{eff} on the T_A^2 and SPE of recursively updated PCA-MSPC (RU-PCA-MSPC) for the low nitrogen input reactor (LNR). The PCA-MSPC model was initially built using the biogas concentrations in CH_4, CO_2, H_2 and H_2S collected from week 2 to week 6 (during the acclimation period) as individual process variables (IPVs). Then the model was recursively updated with the new biogas composition measurements using 3 different values: N_{eff} = 1 day (fast model update, a,b,c); N_{eff} = 14 days (moderately fast model update, d,e,f) and N_{eff} = 28 days (fast model update, g,h,i). The resulting T_A^2 and SPE control charts as well as the rate of model update and the impact of the model update on the number of principal components (PCs) retained in the model are presented. The T_A^2 and the SPE are expressed as a ratio between the index value and its control limit (UCL and SPE_{lim}, respectively). Only T_A^2/UCL and SPE/SPE_{lim} > 1 are plotted (i.e. out-of-control process). PCA-MSPC was not performed during three periods due to gas chromatograph maintenance (hatched grey area).

updated on the basis of too few data to capture the variability associated with the daily and weekly dynamics of the biogas composition. The amplitude of these dynamics clearly increased from week 10 in the case of the H_2S concentration (Fig. 3b). As we imposed the condition of not updating the model parameters for biogas composition measurements corresponding to out-of-control process situation, the model parameters were not updated from week 10 onwards (Fig. 6c). For *slow* model update, more weight was attributed to the oldest individual measurements than to the most recent ones. Therefore, new data brought only a minor contribution to each re-estimation of the model parameters. In consequence, the model did not adapt fast enough to the progressive increase of the H_2S concentration observed for the low nitrogen input reactor. This caused a progressive drift of the T_A^2 and SPE indexes in the out-of-control range, which also stopped the model update (Fig. 6i).

More interestingly, the *moderately fast* model update (N_{eff} = 14 days; 0.25 HRT) allowed continuous model adaptation until the end of the experiment. The T_A^2 and SPE control charts only delivered a few short alarm signals (Fig. 6d and e), which is in accordance with the stable pulp to CH_4 conversion ratio measured for the low nitrogen input reactor.

3.4. RU-PCA-MSPC: Prediction quality assessment for process disturbance detection

RU-PCA-MSPC using an effective memory length of 14 days was applied to the biogas samples collected after the initial model building period for the three high nitrogen input reactors. The resulting T_A^2 and SPE control charts are presented in Fig. 7 for the four reactors, together with the pulp to CH_4 conversion ratio as the process efficiency indicator.

For the low nitrogen input reactor, only minor and transient out-of-control values were observed for both the T_A^2 and SPE indexes (Fig. 7a and b), while the reactor exhibited a stable pulp to CH_4 conversion ratio throughout the experiment. Therefore, these transient alarm signals, probably due to minor and reversible process perturbations, did not announce a risk of process dysfunction that would require adaptation of the reactor feeding regime.

For the high nitrogen input reactors, the monitoring model initially delivered only minor and transient alarm signals during phase II, similar to those observed for the low nitrogen input reactor. The process status was not affected by urea.

Interestingly, during phase III, the monitoring models delivered a clear and persistent alarm signal (T_A^2 > UCL) for the three high nitrogen input reactors (Fig. 7c, e and g) during which an acute process disturbance occurred (a sudden drop of the pulp to CH_4 conversion ratio and VFA accumulation in the slurry). At that time, the microbial community was completely reshaped (Fig. 3). A skilled reactor manager would have taken into account these early and persistent alarm signals seriously and would have verified the value of slurry indicators to identify the cause of these alarms, in this case a probable FAN intoxication characterized by an increase in VFA concentration (Fig. 2g) which was contradictory to a pH above neutrality (Fig. 2d) and increasing TIC and TAN contents in the slurry (Fig. 2c and f). In consequence, the reactor manager would probably have lowered the nitrogen supply to the reactors to reduce the risk of further deterioration

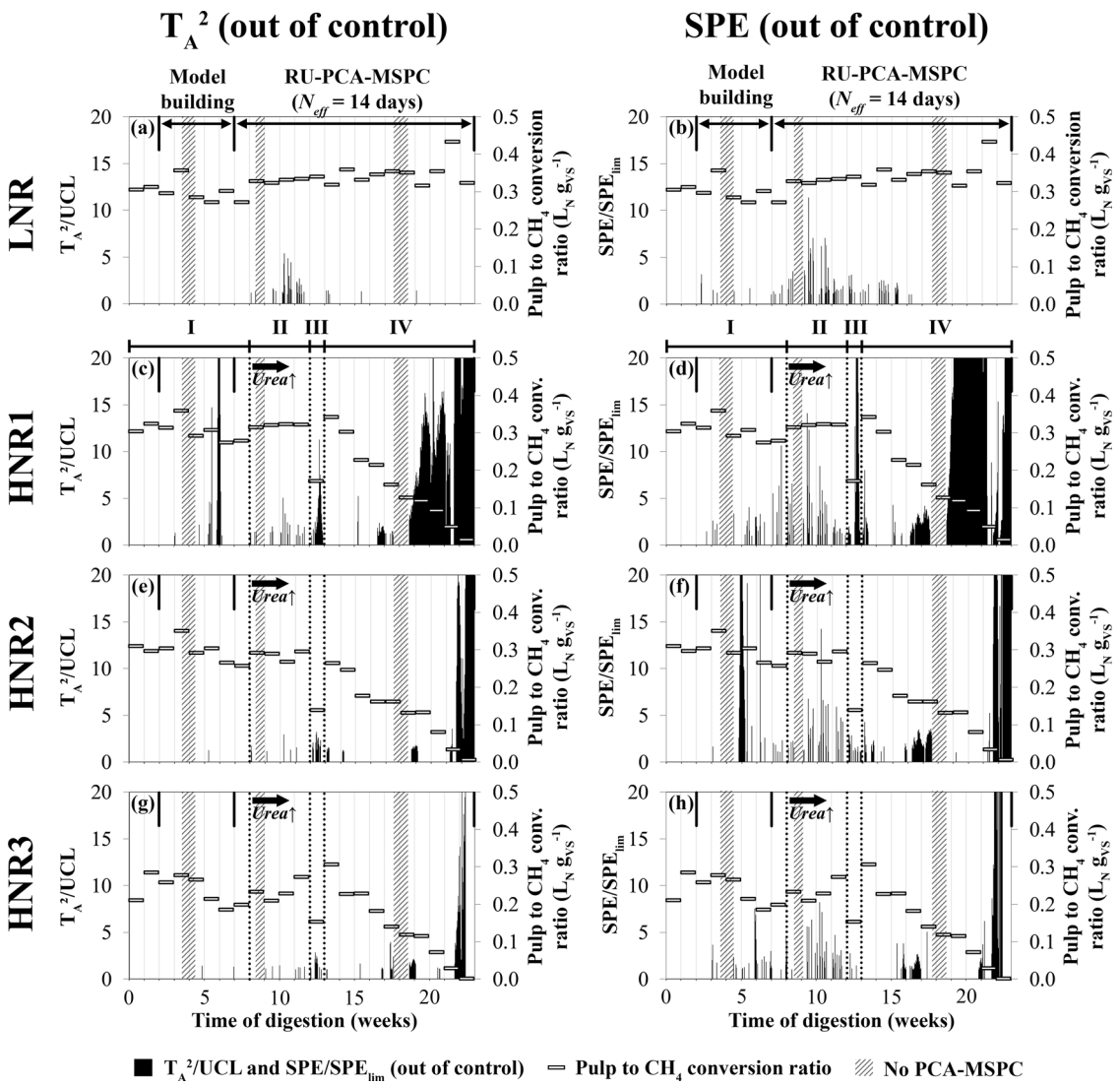

Fig. 7. Recursively updated PCA-MSPC (RU-PCA-MSPC) performed for the four reactors using the biogas concentrations in CH_4, CO_2, H_2 and H_2S as individual process variables (IPVs) and an effective memory length N_{eff} of 14 days for the model update. The initial model building considered the biogas composition measurements collected from week 2 to week 6. T_A^2 and squared prediction error (SPE) control charts built for: the low nitrogen input reactor LNR (a, b), the high nitrogen input reactor HNR1 (c, d), the HNR2 (e, f) and the HNR3 (g, h). The T_A^2 and the SPE are expressed as a ratio between the index value and its control limit (UCL and SPE_{lim}, respectively). Only T_A^2/UCL and $SPE/SPE_{lim} > 1$ are plotted (i.e. out-of-control process). The pulp to CH_4 conversion ratio is presented for each control chart as the process efficiency indicator. PCA-MSPC was not performed during three short periods due to gas chromatograph maintenance operations (hatched grey area). I, II, III and IV correspond to the phases described in Table 2.

of the process.

For the experimental purpose of assessing the robustness of RU-PCA-MSPC model to predict the AD process performance status under a severe FAN intoxication, the nitrogen input was further increased in the high nitrogen input reactors during phase IV. During this phase, characterized by an initial transient improvement of the process status followed by a clear drop of the pulp to CH_4 conversion ratio and a drastic change in the microbiome composition, the three high nitrogen input reactors behaved differently. As the decline of process performance in early phase IV was progressive, the RU-PCA-MSPC model did not detect major out-of-control process values (Fig. 7). Later in phase IV, the process progressively declined further in efficiency while the model provided inconsistent alarms between the replicate HNRs until the final process collapse characterized by extreme and frequent alarms (weeks 19–22). During this period, the alarm rate was clearly higher for the high nitrogen input reactor 1 than for the other high nitrogen input reactors. Differences between the reactors regarding the dynamics of the biogas composition explain this different alarm rate. As reported elsewhere [19], these differences in the behaviour of replicated AD reactors can be explained by a distinctive change in the reactor microbiome composition. Indeed, under sub-optimal conditions (i.e. severe stress) the microbial populations driving the AD process for each reactor progressed differently at the OTU level (data not presented). For the high nitrogen input reactor 1, the H_2S concentration abruptly increased during week 19 (Fig. 3d), thus preventing adaptation of the model parameters using the selected N_{eff} value (14 days). This caused permanent out-of-control values of the T_A^2 and SPE indexes (Fig. 7c and d) for the following biogas composition measurements. As compared to the high nitrogen input reactor 1, the biogas composition was more stable for the high nitrogen input reactors 2 and 3 during that period, which allowed adaptation of the model and explains the lower alarm rate observed for these reactors. This suggests that there is room for further optimisation of the N_{eff} value to be selected. Data generated during this critical process performance decline under severe FAN intoxication showed clearly the limitation of the RU-PCA-MSPC to adapt to a new process driven by a drastically modified microbiome.

Nevertheless, the monitoring models also delivered marked alarm signals during all three stirrer blocking events, demonstrating a good sensitivity of disturbance detection both during low and high nitrogen input in the feed (Fig. 7d, e and f).

3.5. Perspectives for process monitoring in full-scale AD reactors

The use of the concentrations of the major components in the biogas through PCA-MSPC models has already been assessed for early detection of VFA intoxication in AD reactors under organic overload condition [14]. However, this monitoring model used IPVs measured in the slurry in addition to the biogas composition. In the present study, RU-PCA-MSPC models using exclusively the biogas composition showed potential to predict the first signs of performance decrease of reactors exposed to FAN accumulation in the slurry (a sudden decrease of the pulp to CH_4 conversion ratio and VFA accumulation in the slurry). The RU-PCA-MSPC models correctly detected these first symptoms of process dysfunction for three independent high nitrogen input reactors. These symptoms were also correlated with a drastic change in the bacterial and archaeal anaerobic communities. The model capacity to detect early signs of process dysfunction suggests a potential applicability for a large diversity of AD reactors. These results are interesting since the concentrations of the major components of the biogas are parameters that are often neglected for AD process monitoring due to lagged response (CH_4 and CO_2) or oversensitivity (H_2) to process perturbations [30,31]. In comparison with univariate process monitoring techniques, the multivariate approach presented here considers the correlation structure of the IPVs, which could potentially compensate the lagged warning of IPVs such as CH_4 and CO_2. In addition, the PCA-MSPC models were trained on an initial period during which the process and its associate microbiome were stable. This could potentially resolve the oversensitivity to H_2 by taking into account its in-control variation in the model. These results are also promising because analysing the biogas composition allows high frequency on-line measurements involving low workload and reduced sampling issues, as compared to slurry analysis. Indeed, biogas analysers are often present in biogas production plants.

The RU-PCA-MSPC models showed nevertheless limitation to adapt to a process driven by a drastically modified microbiome. The models were trained on stable reactors characterized by similar microbial communities but when applied to the same reactors after they encountered a major shift in their bacterial and archaeal communities, the models provided inconsistent alarms that could not be directly related to the process performance. Thus, even if recursively updated, PCA-MSPC models show limitations when applied to a microbiome they were not trained on.

The results obtained here for the monitoring of a FAN intoxication suggest good versatility of the PCA-MSPC approach for the monitoring of the process in real-scale AD reactors, which are exposed to a large diversity of process perturbations. In addition, the control limits used in multivariate process monitoring are independent of toxicity thresholds [7], which are especially uncertain for FAN due to potential reactor acclimation to nitrogen-rich substrates [3,9].

In the study of Lemaigre et al. [14], data collected from an independent reactor maintained in a steady state were required to build the (static) PCA-MSPC model used to monitor the process of a reactor exposed to organic overload. In the present study, we attempted to build the monitoring models by using only data collected from the monitored reactors themselves. Recursively updated PCA-MSPC showed good results in distinguishing between the actual process perturbations and drift of the reactor in-control process baseline. However, it was demonstrated that the value of the effective memory length should be carefully selected in such recursively updated models. An optimal effective memory length value representing one quarter of the actual HRT (N_{eff} = 14 days) was selected among the tested values on the basis of the lowest alarm rate detected for a reactor known to show permanent and optimal process performance (pulp to CH_4 conversion ratio close to the BMP). An independent steady-state reactor remains therefore necessary to optimize the prediction quality of the recursively updated monitoring model. In a practical situation, the effective memory length of the monitoring model could be selected on the basis of the behaviour of post-reactors, which are known to be rarely exposed to process failure.

In the present study, MSPC techniques were demonstrated as an interesting tool to assess the process status of AD reactors by measuring the concentrations of the major components in the biogas. Currently, many biogas plant operators install biogas analysers, allowing the on-line measurement of the CH_4 and CO_2 concentrations in the biogas, as well as, in most cases, of its H_2S concentration [32]. The measurement of the H_2 concentration is also proposed as an option by many manufacturers of biogas analysers. The process monitoring method proposed here can be applied to those biogas plants that have installed such a biogas analyser, having the advantage of not requiring the installation of additional measuring equipment. This is a clear advantage as compared to the e-nose [15,16] or NIR-based monitoring technologies [33,34] that require additional investment for operational use.

4. Conclusions

Using the biogas composition through recursively updated PCA-MSPC showed potential for early warning of FAN intoxication in AD reactors, which is useful given the uncertainty related to FAN toxicity thresholds in individual reactors. The models also repeatedly detected mixing failures in the studied reactors both during low and high nitrogen input in the feed. An effective memory length of 14 days (0.25 HRT) allowed model adaptation to limit the false alarm rate. Early signs of FAN intoxication, coinciding to drastic changes in the reactor microbiomes were detected by the models for triplicate high nitrogen input reactors, confirming their potential for general application. However, the models presented here lost their prediction capability for process performance after a major shift in the microbiome composition of the reactor. Using only the biogas composition to detect the two main AD process dysfunctions (VFA and FAN intoxication) is promising, given the general use of biogas analysers in most biogas production plants and the limitations related to slurry analysis.

Acknowledgements

This work was supported by the Fond National de la Recherche, Luxembourg (FNR CORE 2011 project GASPOP, CO11/SR/1280949: Influence of the Reactor Design and the Operational Parameters on the Dynamics of the Microbial Consortia Involved in the Biomethanation Process) and the LIST core funding. The authors are most grateful to Jillian M. Lenné and Pau Ferrer Alegre for proofreading the manuscript and their useful comments.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.bej.2018.08.018.

References

[1] D. Helm, The European framework for energy and climate policies, Energy Policy 64 (2014) 29–35, https://doi.org/10.1016/j.enpol.2013.05.063.
[2] D.J. Batstone, B. Virdis, The role of anaerobic digestion in the emerging energy economy, Curr. Opin. Biotechnol. 27 (2016) 142–149, https://doi.org/10.1016/j.copbio.2014.01.013.
[3] R. Rajagopal, D.I. Massé, G. Singh, A critical review on inhibition of anaerobic digestion process by excess ammonia, Bioresour. Technol. 143 (2013) 632–641, https://doi.org/10.1016/j.biortech.2013.06.030.
[4] C. González-Fernández, P.A. García-Encina, Impact of substrate to inoculum ratio in anaerobic digestion of swine slurry, Biomass Bioenergy 33 (2009) 1065–1069, https://doi.org/10.1016/j.biombioe.2009.03.008.

[5] O. Yenigün, B. Demirel, Ammonia inhibition in anaerobic digestion: a review, Process Biochem. 48 (2013) 901–911, https://doi.org/10.1016/j.procbio.2013.04.012.
[6] B. Calli, B. Mertoglu, B. Inanc, O. Yenigun, Effects of high free ammonia concentrations on the performances of anaerobic bioreactors, Process Biochem. 40 (2005) 1285–1292, https://doi.org/10.1016/j.procbio.2004.05.008.
[7] L. Li, Q. He, Y. Wei, Q. He, X. Peng, Early warning indicators for monitoring the process failure of anaerobic digestion system of food waste, Bioresour. Technol. 171 (2014) 491–494, https://doi.org/10.1016/j.biortech.2014.08.089.
[8] C. Resch, A. Wörl, R. Waltenberger, R. Braun, R. Kirchmayr, Enhancement options for the utilisation of nitrogen rich animal by-products in anaerobic digestion, Bioresour. Technol. 102 (2011) 2503–2510, https://doi.org/10.1016/j.biortech.2010.11.044.
[9] Y. Chen, J.J. Cheng, K.S. Creamer, Inhibition of anaerobic digestion process: a review, Bioresour. Technol. 99 (2008) 4044–4064, https://doi.org/10.1016/j.biortech.2007.01.057.
[10] J. MacGregor, Statistical process control of multivariate processes, Control Eng. Pract. 3 (1995) 403–414, https://doi.org/10.1016/0967-0661(95)00014-L.
[11] T. Kourti, J.F. MacGregor, Process analysis, monitoring and diagnosis, using multivariate projection methods, Chemometr. Intell. Lab. Syst. 28 (1995) 3–21, https://doi.org/10.1016/0169-7439(95)80036-9.
[12] T. Kourti, Application of latent variable methods to process control and multivariate statistical process control in industry, Int. J. Adapt. Control Signal Process. 19 (2005) 213–246, https://doi.org/10.1002/acs.859.
[13] W. Li, H. Yue, S. Valle-Cervantes, S.J. Qin, Recursive PCA for adaptive process monitoring, J. Process Control 10 (2000) 471–486, https://doi.org/10.1016/S0959-1524(00)00022-6.
[14] S. Lemaigre, G. Adam, X. Goux, A. Noo, B. De Vos, P.A. Gerin, P. Delfosse, Transfer of a static PCA-MSPC model from a steady-state anaerobic reactor to an independent anaerobic reactor exposed to organic overload, Chemometr. Intell. Lab. Syst. 159 (2016) 20–30, https://doi.org/10.1016/j.chemolab.2016.09.010.
[15] G. Adam, S. Lemaigre, A.-C. Romain, J. Nicolas, P. Delfosse, Evaluation of an electronic nose for the early detection of organic overload of anaerobic digesters, Bioprocess Biosyst. Eng. 36 (2012) 23–33, https://doi.org/10.1007/s00449-012-0757-6.
[16] G. Adam, S. Lemaigre, X. Goux, P. Delfosse, A.C. Romain, Upscaling of an electronic nose for completely stirred tank reactor stability monitoring from pilot-scale to real-scale agricultural co-digestion biogas plant, Bioresour. Technol. 178 (2015) 285–296, https://doi.org/10.1016/j.biortech.2014.09.106.
[17] A.-C. Romain, J. Nicolas, Sensors and actuators B : chemical Long term stability of metal oxide-based gas sensors for e-nose environmental applications: an overview, Sens. Actuators B. Chem. 146 (2010) 502–506, https://doi.org/10.1016/j.snb.2009.12.027.
[18] L. Zhang, D. Zhang, Domain adaptation extreme learning machines for drift compensation in E-nose systems, IEEE Trans. Instrum. Meas. 64 (2015) 1790–1801, https://doi.org/10.1109/TIM.2014.2367775.
[19] X. Goux, M. Calusinska, S. Lemaigre, M. Marynowska, M. Klocke, T. Udelhoven, E. Benizri, P. Delfosse, Microbial community dynamics in replicate anaerobic digesters exposed sequentially to increasing organic loading rate, acidosis, and process recovery, Biotechnol. Biofuels 8 (2015) 122, https://doi.org/10.1186/s13068-015-0309-9.
[20] F. Mayer, P.A. Gerin, A. Noo, G. Foucart, J. Flammang, S. Lemaigre, G. Sinnaeve, P. Dardenne, P. Delfosse, Assessment of factors influencing the biomethane yield of maize silages, Bioresour. Technol. 153 (2014) 260–268, https://doi.org/10.1016/j.biortech.2013.11.081.
[21] M.C. Sterling, R.E. Lacey, C.R. Engler, S.C. Ricke, Effects of ammonia nitrogen on H2 and CH4 production during anaerobic digestion of dairy cattle manure, Bioresour. Technol. 77 (2001) 9–18, https://doi.org/10.1016/S0960-8524(00)00138-3.
[22] X. Goux, M. Calusinska, M. Fossépré, E. Benizri, P. Delfosse, Start-up phase of an anaerobic full-scale farm reactor – appearance of mesophilic anaerobic conditions and establishment of the methanogenic microbial community, Bioresour. Technol. 212 (2016) 217–226, https://doi.org/10.1016/J.BIORTECH.2016.04.040.
[23] M. Calusinska, X. Goux, M. Fossépré, E.E.L. Muller, P. Wilmes, P. Delfosse, A year of monitoring 20 mesophilic full-scale bioreactors reveals the existence of stable but different core microbiomes in bio-waste and wastewater anaerobic digestion systems, Biotechnol. Biofuels 11 (2018) 196, https://doi.org/10.1186/s13068-018-1195-8.
[24] J. Moestedt, B. Müller, M. Westerholm, A. Schnürer, Ammonia threshold for inhibition of anaerobic digestion of thin stillage and the importance of organic loading rate, Microb. Biotechnol. 9 (2016) 180–194, https://doi.org/10.1111/1751-7915.12330.
[25] T.-C. Huang, D.-H. Chen, Kinetic study of urease-catalysed urea hydrolysis, J. Chem. Technol. Biotechnol. 52 (2007) 433–444, https://doi.org/10.1002/jctb.280520402.
[26] Scilab Enterprises, Scilab: Free and Open Source Software for Numerical Computation, available from: (2012) http://www.scilab.org.
[27] H. Zhou, G. Qiu, Inhibitory effect of ammonia nitrogen on specific methanogenic activity of anaerobic granular sludge, J. Cent. South Univ. Technol. 13 (2006) 63–67, https://doi.org/10.1007/s11771-006-0108-3.
[28] I.A. Fotidis, D. Karakashev, T.A. Kotsopoulos, G.G. Martzopoulos, I. Angelidaki, Effect of ammonium and acetate on methanogenic pathway and methanogenic community composition, FEMS Microbiol. 83 (2013) 38–48, https://doi.org/10.1111/j.1574-6941.2012.01456.x.
[29] J.L. Chen, R. Ortiz, T.W.J. Steele, D.C. Stuckey, Toxicants inhibiting anaerobic digestion: a review, Biotechnol. Adv. 32 (2014) 1523–1534, https://doi.org/10.1016/j.biotechadv.2014.10.005.
[30] L. Björnsson, M. Murto, B. Mattiasson, Evaluation of parameters for monitoring an anaerobic co-digestion process, Appl. Microbiol. Biotechnol. 54 (2000) 844–849, https://doi.org/10.1007/s002530000471.
[31] K. Boe, D.J. Batstone, J.P. Steyer, I. Angelidaki, State indicators for monitoring the anaerobic digestion process, Water Res. 44 (2010) 5973–5980, https://doi.org/10.1016/J.WATRES.2010.07.043.
[32] B. Drosg, Process monitoring in biogas plants, IEA Bioenergy Task 37 (2013) 1–38.
[33] C. Lomborg, J. Holm-Nielsen, Near infrared and acoustic chemometrics monitoring of volatile fatty acids and dry matter during co-digestion of manure and maize silage, Bioresour. Technol. 100 (2009) 1711–1719, https://doi.org/10.1016/j.biortech.2008.09.043.
[34] H. Jacobi, C. Moschner, Use of near infrared spectroscopy in monitoring of volatile fatty acids in anaerobic digestion, Water Sci. Technol. 60 (2009) 339–346, https://doi.org/10.2166/wst.2009.345.