Chemometrics and Intelligent Laboratory Systems 159 (2016) 20–30

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

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

Transfer of a static PCA-MSPC model from a steady-state anaerobic reactor to an independent anaerobic reactor exposed to organic overload

Sébastien Lemaigre[a,c,*], Gilles Adam[b], Xavier Goux[a], Anaïs Noo[a], Bénédicte De Vos[a], Patrick A. Gerin[c], 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 de Liège, Arlon Campus Environnement, A venue de Longwy 185, B-6700 Arlon, Belgium
[c] Université catholique de Louvain, Earth & Life Institute, Bioengineering, Croix du Sud 2, box L7.05.19, B-1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

Keywords:
Anaerobic digestion
Organic overload
Multivariate process monitoring
Hotelling's T²

ABSTRACT

A static multivariate statistical process control model based on principal component analysis (PCA-MSPC) was developed for an anaerobic reactor maintained in steady-state, joining the biogas composition (CH_4, CO_2, H_2) to the total solids (TS), volatile solids (VS), total inorganic carbon (TIC) and total ammonia nitrogen (TAN) contents of the slurry. The principal component analysis (PCA) highlighted a lack of correlation between the individual process variables (IPVs) measured in the slurry and the gas phase. The application of this model to the data set collected for an independent anaerobic reactor (fed with the same substrate) progressively led from steady-state to critical volatile fatty acids (VFA) intoxication did not allow evaluating its process status. A second static PCA-MSPC model was built for the steady-state reactor in excluding the TS and VS content of the slurry and was successfully transferred to the overfed reactor. The T_A^2 and SPE control charts built for the overfed reactor using this second model closely reflected its process status and delivered valuable warning signals approaching the VFA intoxication. The gas phase composition (CH_4, CO_2, H_2) brought the main contribution to these warnings.

1. Introduction

In recent decades, anaerobic digestion (AD) of organic substrates for biomethane production has become one of the most mature technologies to produce renewable energy from wet biomass [1]. The life cycle assessment of the process shows an interesting environmental performance [2] and allows the valorisation of organic co-products that would otherwise undergo elimination processes, leading to nutrient losses in the agricultural systems.

One of the main obstacles for further development of the AD sector is the difficulty in keeping the biological process in optimal and stable conditions of activity, especially in a context of organic waste that varies in price, composition, and availability. So far, various methods have been evaluated to monitor the AD process but none seem to be ideal [3]. These methods usually consist in measuring a set of variables expected to be characteristic of the process status (i.e. pH of the slurry, CH_4/CO_2 ratio of the biogas,...) and interpreting the collected data for each parameter individually [4,5]. However, since these variables reflect the behaviour of the microbial community of the reactor, they probably present a certain degree of correlation. An efficient tool for AD process monitoring should therefore benefit from the integration of information on how the measured parameters interact when the process is in control. A method to satisfy this condition is to monitor the reactors using multivariate statistical process control (MSPC) as an alternative to usual univariate approaches. MSPC techniques reduce the information contained within a potentially high number of individual process variables (IPVs) down to a low number of composite indexes through the application of statistical modelling.

Abbreviation: AD, anaerobic digestion; CSTR, completely stirred tank reactor; HRT, hydraulic retention time; IPV, individual process variable; L_N, normalized liter of gas; MSPC, multivariate statistical process control; MOS, metal oxide semiconductor sensor; NIR, near infrared spectroscopy; OLR, organic loading rate; OR, overfed reactor; PCA, principal component analysis; PC, principal component; SPE, squared prediction error index; SPE_{lim}, control limit of the SPE index; SSR, steady-state reactor; T², hotteling's T² index; T_A^2, hotteling's T² index for the A first PCs; TAN, total ammonia nitrogen; TIC, total inorganic carbon; TS, total solids; UCL, upper control limit of the T² (or T_A^2) index; VFA, volatile fatty acids; 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@ulg.ac.be (G. Adam), xavier.goux@list.lu (X. Goux), patrick.gerin@uclouvain.be (P.A. Gerin), philippe.delfosse@list.lu (P. Delfosse).

http://dx.doi.org/10.1016/j.chemolab.2016.09.010
Received 22 February 2016; Received in revised form 27 September 2016; Accepted 29 September 2016
Available online 30 September 2016

Several studies assessed MSPC for monitoring of anaerobic biological processes such as biological anaerobic filters exploited for wastewater treatment [6] or batch AD reactors [7,8]. Nevertheless, very few studies focussed on (semi-)continuously fed AD reactors targeting biomethane production. Attempts to exploit multivariate statistics for process monitoring of semi-continuous AD reactors mainly use near infrared spectroscopy (NIR) and aim to correlate the NIR response with common AD process indicators measured in the slurry [9–11]. However, the main purpose of these studies is not to assess the potential of these parameters as IPVs exploited for MSPC but to find an alternative to their expensive and time-consuming traditional measurement methods (and to allow their on-line measurement). Reed et al. [12] applied MSPC to evaluate the process status of a semi-continuous AD reactor on the basis of IPVs measured in the slurry, the feeding substrate, and the effluents of the reactors but neglected to exploit the composition of the biogas produced. However, analysing the composition of the biogas offers multiple advantages compared to slurry analysis in terms of AD process monitoring, such as a potentially higher measurement frequency, the immediate availability of the results or the lower clogging risk inside the tubing. In recent years, electronic nose (e-nose) was the main method studied exploiting the gas phase for AD process monitoring through MSPC [13,14]. In process monitoring using e-nose, the responses of low-selectivity gas sensors are used as IPVs to perform MSPC. If the technology is promising, the metal oxide semiconductor sensors (MOS) exploited in e-noses are subject to drift and need frequent replacement [15]. Nevertheless, specific sensors that are more robust and less subject to drift (if regularly calibrated) can be used to measure the concentration of the biogas in its major compounds (CH_4, CO_2, H_2 and H_2S). To our knowledge, exploiting these parameters for AD process monitoring through MSPC was never attempted. However, a MSPC model joining parameters commonly measured in the slurry to the biogas composition in its major compounds could present interesting potential for AD process monitoring and deserves investigation.

MSPC involves the construction of control charts that represent the progress over time of the process status. In these control charts, the value of the index is compared to a control limit that defines the index value above which the process is considered as out-of-control. Control charts based on the Hotelling's T^2 index are nowadays the basic tools exploited for MSPC and numerous descriptions of the technique can be found in the literature [16,17]. For processes characterized by two or more highly correlated IPVs and/or by a large set of IPVs, a common procedure to reduce the dimensionality of the problem is to perform a principal components analysis (PCA). In PCA-MSPC, a limited number *A* of principal components is used to compute the T^2 index. The T^2 index calculated on this basis (T_A^2) can express the deviation in the IPVs that are the most meaningful in describing a process [18]. Since the T_A^2 index only describes the variation amplitude of the IPVs in the "plane" defined by the *A* PCs retained in the model, a complementary index, squared prediction error (SPE), is combined to the T_A^2 to quantify the residuals. The pair of control charts based on the T_A^2 and SPE indexes is an effective tool to monitor multivariate processes [18]. The basic way to implement PCA-MSPC, called *static* PCA-MSPC, consists in calculating the model parameters on the basis of data collected for a period during which the process is judged stable. PCA-MSPC is easy to implement and requires only a few computations once the model is defined (the model is not updated after it has been built). In addition, static PCA-MSPC allows the easy computation of contribution plots that provide valuable information about the relationship between each process dysfunction detected by the T_A^2 and SPE control charts and the IPVs [19–21]. Nevertheless static PCA-MSPC is limited in the case of processes that demonstrate a drift of their stable behaviour such as AD, which is affected by the permanent evolution of the microbial flora or accumulation of chemical compounds in the slurry. However, a reactor maintained in a steady-state over a long period could allow measuring of the relationship existing between the IPVs for a large diversity of in-control process situations. On this basis, a static PCA-MSPC model integrating exhaustive information on the in-control process could be built and transferred to independent reactors to evaluate their process status.

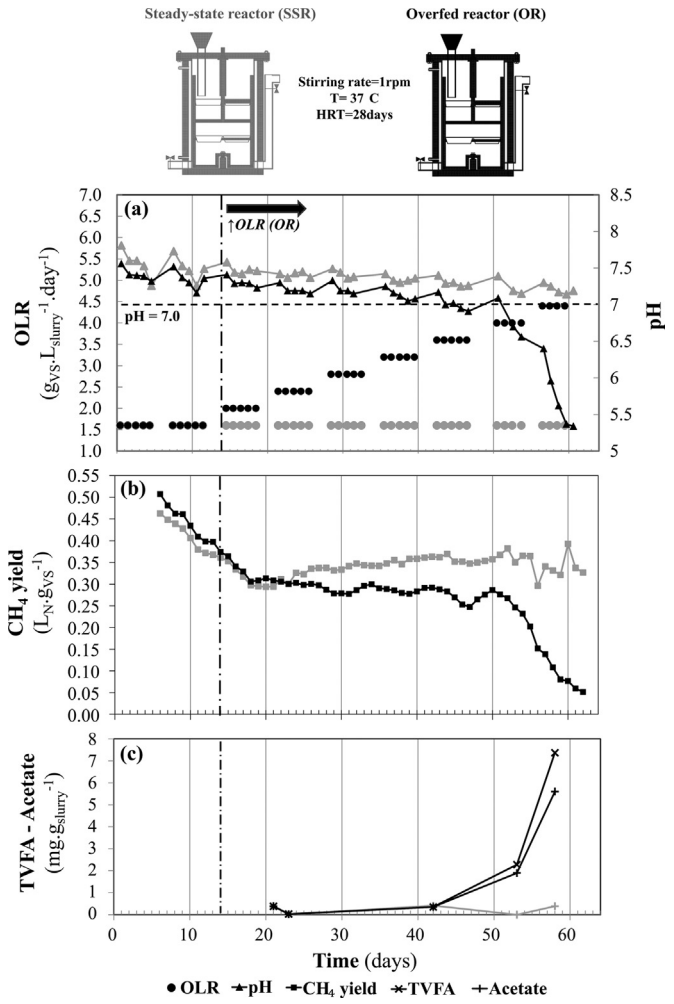

Fig. 1. Progress over time of the 3 selected process stability indicators (pH, CH_4 yield, volatile fatty acids content in the slurry) for the steady-state reactor (SSR, grey) and the overfed reactor (OR, black): (a) organic loading rate (OLR) and pH; (b) methane yield; (c) total volatile fatty acids (TVFA) and acetate concentrations in the slurry. HRT: hydraulic retention time.

The main objective of the present paper was to assess whether a static PCA-MSPC model joining the biogas composition to parameters measured in the slurry and built using an AD reactor maintained in a steady-state could be transferred to an independent AD reactor progressively led to an organic overload and reflect its process status. The experiment described in this study focussed on a large scale semi-continuous lab reactors expected to be comparable to real-scale co-digestion digesters regarding the complexity of the feeding substrate and the hydraulic retention time (HRT).

2. Materials and methods

2.1. Anaerobic digestion

The experiment was performed using two stainless steel tank reactors of 100 L working volume (Fig. 1), continuously stirred at 1 rpm, with 25 L headspace volume connected to a 10 L gas bag (Tecobag, Tesseraux, Germany) to equilibrate the pressure (feeding, gas sampling). Tygon tubing (VWR International, USA) was used for all the gas connections.