Proceedings of the 10th International Chemical and
Biological Engineering Conference - CHEMPOR 2008
Braga, Portugal, September 4-6, 2008
E.C. Ferreira and M. Mota (Eds.)

A chemometric tool to monitor high-rate anaerobic granular sludge reactors during load and toxic disturbances

José Carlos Costa, M. Madalena Alves, Eugénio C. Ferreira*
IBB – Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710–057 Braga, Portugal.

Keywords: Control, Organic Loading Disturbance, Principal Components Analysis, Quantitative Image Analysis, Toxic Shock Load
Topic: Systematic methods and tools for managing the complexity.

Abstract

The wide fluctuations in flow rate and the presence of toxic compounds can damage the high efficiency of high-rate anaerobic granular sludge reactors. As earlier these disturbances are detected, more accurate would be the corrective actions, and less damage will be caused to the microorganisms involved in the process. The use of Principal Component Analysis (PCA) determined a latent variable, combining a weighted sum of operational, physiological, and morphological data, which showed high sensitivity to recognize the operational problems occurred when four organic loading disturbances and three toxic shock loads were applied to Expanded Granular Sludge Bed reactors. The high loadings/weights linked with the morphological parameters, specially the aggregates size distribution (>0.1, >1) and filaments length dynamics (TL/VSS), obtained using quantitative image analysis techniques, illustrate the usefulness of monitor the structural changes of the anaerobic granular sludge. The application of PCA chemometric tool to dataset gathering information from all disturbances allowed the differentiation between organic loading and toxic shock disturbances, as well as the main effects caused by each class of disturbance.

1 Introduction

High-rate anaerobic digestion reactors gained significance during the last decades in the field of wastewater treatment processes. However, these systems are designed with reference to a nominal operating condition, in which the organic loading rate is assumed to be constant in time. Also, some compounds can have inhibitory or toxic effects to the microbial populations, such as detergent and solvents. These facts, coupled with the long start-up periods imply the need to monitor the anaerobic granular sludge stability in order to achieve an appropriate control and sustainability of the process.

The recognition of parameters that could be used for monitoring the process is important to efficient control of those processes. It is equally feasible to obtain values of parameters measured in solid, liquid or gaseous phases. However, parameters involved in reactors control had been limited to indicators of the liquid and the gaseous phases (van Lier et al., 2001), due to difficulties in obtain and inaccuracy associated with morphological parameters.

With the rapid development of instrumental methods the amount of diverse data generated in an environmental process monitoring and/or control is increasingly drastically (Bourgeois et al., 2001; Schügerl, 2001; Spanjers and van Lier, 2006). This advance guide analysts and researchers to gathering further more multivariate data. Concurrently, with computer science and technology developments apply computers

* Corresponding author. Tel + 351-253-604407. E-mail:ecferreira@deb.uminho.pt

and advanced statistical and mathematical methods to analyse this data became easier.

In this framework, image analysis techniques appear as a promising tool to provide quantitative parameters of the solid phase evolution. And, chemometrics based techniques, such as Principal Components Analysis (PCA), can be useful to detect groups, trends, correlations, and outliers in datasets gathering vast amounts of information. This method allows identifying patterns in data, and expressing them in order to highlight their similarities and differences. PCA is a projection method for analyze data and reduce it from an n-dimensional space to few latent/hidden variables, while keeping information on its variability.

A multivariate statistical analysis has been used together with image analysis techniques to pattern recognition, such as discriminant analysis, neural networks, and decision trees (Ginoris et al., 2007). The relationships between morphological parameters and biomass properties in aerobic wastewater treatment processes were also assessed by partial least squares regression (Amaral et al., 2005) and principal components analysis (Jenné et al., 2006).

The objective of this work was to apply the chemometric technique Principal Components Analysis in order to recognize fluctuations, and respective effects, in high-rate anaerobic granular sludge reactors performance caused by organic loading and toxic disturbances.

2 Material and Methods

2.1 Datasets

Four organic loading disturbances (OLD) (Table 1) were applied to an Expanded Granular Sludge Bed (EGSB) reactor fed with 1.5$g_{COD\text{-}ethanol}$/L and hydraulic retention time (HRT) of 8 hours, in steady-state conditions. Also, three toxic shock loads (TSL) were applied in EGSB reactors operating in similar conditions (Table 1).

Table 1. Organic Loading Disturbances and Toxic Shock Loads conditions.

Disturbance	OLD1	OLD2	OLD3	OLD4	TSL1	TSL2	TSL3
Ethanol (g_{COD}/L)	5	1.5	15	15	1.5	1.5	1.5
HRT (h)	8	2.5	8	8	8	8	8
[Toxic] (mg/L)	-	-	-	-	1.6	3.1	40
Exposure phase (h)	72	72	72	384	56	222	222
Recovery phase (d)	7	7	7	7	14	12	7

Three programmes previously developed (Amaral, 2003) were used as the final step of a procedure (Costa et al., 2007) to obtain quantitative morphological information from anaerobic granular sludge.

Three datasets were created gathering morphological, physiological, and reactors performance information. Datasets 1 and 2 included observations of OLD and TSL, respectively. The objective consisted in examine the sensitivity of the latent variables to recognize the disturbances. Dataset 3 encompassed all observations and was created to study the differentiation of the OLD from the TSL, and respective effects. The variables used are described in Table 2.

2.2 Principal Component analysis

Principal components analysis aims at finding and interpreting hidden complex, and possibly causally determined, relationships between features in a dataset. Correlating features are converted to the so-called factors which are themselves noncorrelated (Einax et al., 1997).

SIMCA-P (Umetrics AB) software package was used to perform the Principal Components Analysis. The first step of the analysis consists in the pre-treatment of data by standardization of the variables, i.e., guarantee that each individual variable has about the same range, avoiding that some variables would be more important than others because of scale effects. During this work each variable was autoscaled so that each variable has mean zero and unit standard deviation.

Subsequently, the software iteratively computes one principal component at a time, comprising a score vector t_a and a loading vector p_a. The score vectors contain information on how the samples relate to each other. Otherwise, the loading vectors define the reduced dimension space and contain information on how the variables relate to each other. Usually, few principal components (2 or 3) can express most of the variability in the dataset when there is a high degree of correlation among data.

The criterion used to determine the model dimensionality (number of significant components) was cross validation (CV). Part of data is kept out of the model development, and then are predicted by the model and compared with the actual values. The prediction error sum of squares (PRESS) is the squared differences between observed and predicted values for the data kept out of the model fitting. This procedure is repeated several times until data element has been kept out once and only once. Therefore, the final PRESS has contributions from all data. For every dimension, SIMCA computes the overall PRESS/SS, where SS is the residual sum of squares of the previous dimension. A component is considered significant if PRESS/SS is statistically smaller than 1.0.

3. Results and Discussion

3.1 Recognition of organic load and toxic disturbances

The PCA expressed the importance of the proposed morphological parameters to recognize, possible problematic, disturbances to high-rate anaerobic granular sludge reactors. The latent variables t[1] and t[2] showed high percentages of variation in the first hours of exposure in every disturbances applied (Figs. 1 and 2).

As showed in table 1, LD3 and 4 were the most severe organic loading disturbances. Watching at t[1] (Fig. 1) is visible its high variation in this two OLD, essentially caused by the granules fragmentation, biomass washout, and decrease in specific acetoclastic activity (SAA) and COD removal efficiency. These variables had the higher loadings in Principal Component (PC) 1 (Table 2). In LD1, although the COD removal efficiency was unaffected, granule fragmentation was observed. In LD2 (hydraulic shock load) there was no changes in granules distribution sizes, but a severe release of filaments. For that reason the TL/VSS and LfA parameters had high loadings in PC2 (Table 2), and t[2] show the higher variation in the LD2 (data not shown).

Relatively to toxic shock loads was observed that the detergent SL1 had no significant effects on reactor performance and biomass characteristics. These was detected by the PCA, since the variable t[1] do not show large variation during this TSL (Fig. 2). The detergent SL2 caused the biggest decrease in reactor performance, and, solvent SL3 caused granules fragmentation and continuous deterioration of biomass and reactor performance. Observing the evolution of t[1] in figure 2, is clearly visible an immediate deviation between the inoculum observation in SL2 and the first observation during shock load (after 8h), recognizing the operational problems caused by the detergent. In the same figure, is visible a constant decrease of t[1] during SL3. During toxic shock loads, constant filaments release was visible. This morphological change was responsible by the variability detected in PC2 (data not shown) as showed by the high loadings of TL/VSS associated with the latent variable t[2] (Table 2).

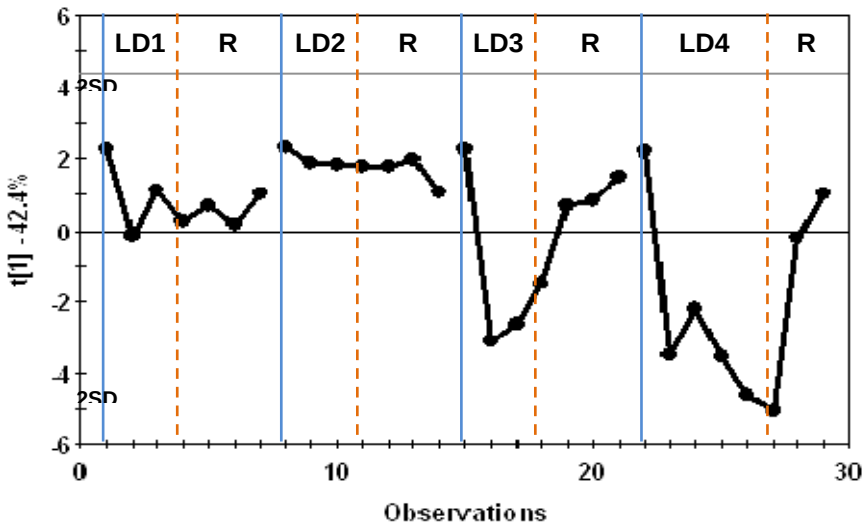


Figure 1. First latent variable (t[1]) evolution during organic loading disturbances. LD – Load Disturbance Phase; R – Recovery Phase.

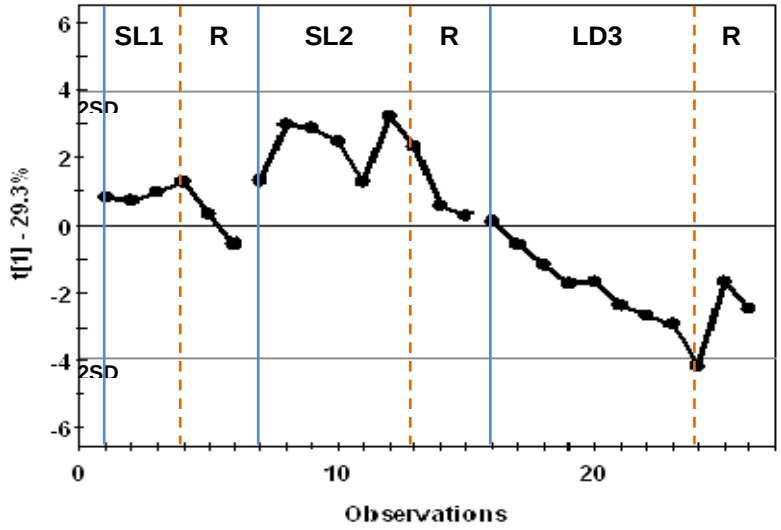


Figure 2. First latent variable (t[1]) evolution during toxic shock loads. SL – Toxic Shock Load Phase; R – Recovery Phase.

Table 2. Loadings/weights associated with the first (t[1]) and second (t[2]) latent variable in organic loading disturbances (LD) and toxic shock loads (SL).

Variable	LD		SL		Name
	t[1]	t[2]	t[1]	t[2]	
OLR	**-0.40**	0.06	0.12	-0.20	Organic Loading Rate
Csd	-	-	**-0.34**	-0.07	Toxic (detergent or solvent) concentration
Eff	**0.40**	0.02	-0.20	**0.30**	Chemical Oxygen Demand (COD) Removal Efficiency
pH	**0.32**	-0.12	0.00	**0.38**	pH
VSS	**-0.40**	-0.12	-0.05	-0.28	Effluent Volatile Suspended Solids
<0.1	-0.12	-0.29	-0.20	**0.30**	% of Aggregates Projected Area with Deq < 0.1 mm
>0.1	**-0.35**	0.19	**-0.43**	-0.09	% of Aggregates Projected Area with 0.1 ≤ Deq (mm) < 1
>1	**0.36**	-0.16	**0.44**	0.07	% of Aggregates Projected Area with Deq ≥ 1 mm
SAA	**0.33**	0.13	-0.26	**0.30**	Specific Acetoclastic Activity
SHMA	0.04	**-0.41**	0.18	-0.04	Specific Hydrogenotrophic Methanogenic Activity
LfA	0.13	**0.55**	0.24	**0.38**	Total Filaments Length / Total Aggregates Projected Area
VSS/TA	0.16	-0.18	0.17	**-0.33**	VSS / Total Aggregates Projected Area
TL/VSS	0.05	**0.55**	0.13	**0.44**	Total Filaments Length / Volatile Suspended Solids
vsed	-	-	**0.46**	0.01	Settling Velocity

Summarising, besides the imposed variables, organic loading rate (OLR) and toxic concentration (Csd), the morphological variables, >1 and >0.1 in t[1], and LfA and TL/VSS in [t2], had the higher loadings/weights, enhancing the need to monitor the biomass morphology to control the reactors. It may be suggested that, from the liquid and gaseous phase, the pH can be an important parameter to monitor/control the process, as reported by others (Liu et al., 2004). The results show the adequacy of use the chemometric technique Principal Component Analysis to recognize disturbances in high-rate anaerobic reactors and detect the respective effects.

3.2 Distinguish organic load disturbances from toxic shock loads

The two firsts Principal Components gathered 51.2% of the total variability in dataset 3 (encompassing observations from organic loading and toxic disturbances). However, analysing the score and loading plots of PC1-PC2 plane (Fig. 3) is possible to visualize the variables with higher influence to distinguish load disturbances from shock load observations and respective effects, i.e. the variables more affected by the disturbances. The organic loading disturbances displaced the observations in direction of negative scores in PC1 and PC2 (line 1 in Fig. 3a). Concerning to the toxic shock load, it was observed that observations were dislocated in the direction of positive scores in PC1 and negative scores in PC2 (line 2 in Fig. 3a).

Watching at the direction lines of the exposure phase observations (Fig. 3a) and the loading plot (Fig. 3b), is visible that the TSL affects mostly the morphological variables LfA and TL/VSS. As stated before a severe release of filaments was observed in these disturbances, being these variables responsible for the detection of operational problems even before the COD removal efficiency decrease. The OLD caused increases in the VSS, and, decreases in the pH and COD removal efficiency. A severe fragmentation phenomenon was also observed, as showed by the increase in the percentage of aggregates projected area with equivalent diameter (D_{eq}) between 0.1–1 mm (>0.1), and consequent decrease in the % of aggregates area with D_{eq} > 1 mm (>1).

The recovery phases observations tend to return to close the inocula observations (Fig. 3a). Effectively, the reactors regain its pre-disturbances performance few hours after their stop, indicating that only temporary inhibitions occurred.

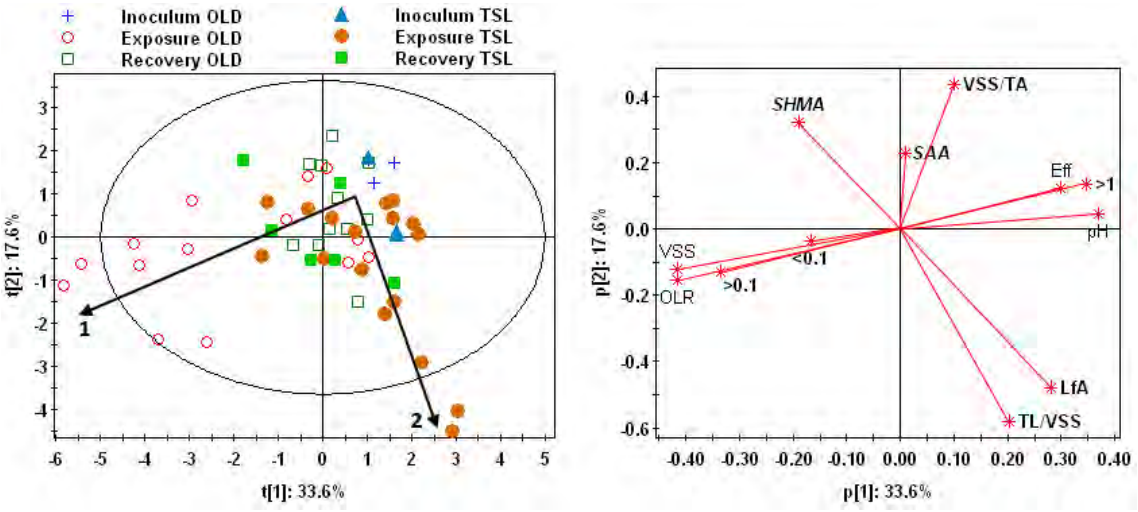


Figure 4. PCA of the first principal component versus the second principal component: (**a**) score plot t[1] *vs.* t[2], and, (**b**) loading plot p[1] *vs.* p[2].

4. Conclusions

The morphological parameters >0.1 and >1, indicator of granules fragmentation, experience high variation in the first hours of disturbances, either when was an organic or a toxic shock load. Their high weights were relevant for the immediate recognition of the deviations to the normal operation of the EGSB reactors detected by the latent variable determined by the PCA. The results enhance the important role that biomass morphology monitoring can have in the recognition of organic loading and toxic disturbances in anaerobic granular sludge reactors. Indeed, the proposed morphological parameters were sensitive enough to detect the disturbances before COD removal efficiency decreased. The PCA allow to distinguish organic loading from the toxic disturbances, and the main effects caused by each other. Increase the OLR caused granules fragmentation and washout (increase in effluent VSS) and decrease in COD removal efficiency. The exposure to detergent and solvent caused the filaments release and consequent increase in LfA parameter, and decrease in the SHMA.

Acknowledgements

We grateful acknowledge the financial support to J.C. Costa through the grant SFRH/BD/13317/2003 and the project POCTI/AMB/60141/2001 from the Fundação para a Ciência e a Tecnologia (Portugal).

References

Amaral, A.L. (2003). *Image Analysis in Biotechnological Processes: Application to Wastewater Treatment.* PhD thesis, University of Minho, Portugal (http://hdl.handle.net/1822/4506).

Amaral, A.L., and Ferreira, E.C. (2005). Activated sludge monitoring of a wastewater treatment plant using image analysis and partial least squares regression. *Analytica Chimica Acta*, 544, 246-253.

Bourgeois, W., Burgess, J.E., and Stuetz, R.M. (2001). On-line monitoring of wastewater quality: a review. *Journal of Chemical Technology and Biotechnology*, 76, 337-348.

Costa, J.C., Abreu, A.A., Ferreira, E.C., Alves M.M. (2007). Quantitative image analysis as a diagnostic tool for monitoring structural changes of anaerobic granular sludge during detergent shock loads. *Biotechnology and Bioengineering,* 98(1), 60-68.

Einax, J.W., Zwanziger, H.W, and Geiss, S. (1997). *Chemometrics in Environmental Analysis*. Weinheim: VCH.).

Ginoris, Y.P., Amaral, A.L., Nicolau, A., Coelho, M.A.Z., and Ferreira, E.C. (2007). Recognition of protozoa and metazoa using image analysis tools, discriminant analysis, neural networks and decision trees. *Analytica Chimica Acta*, 595, 160-169.

Liu, J., Olsson, G., and Mattiasson, B. (2004). Monitoring and control of an anaerobic upflow fixed-bed reactor for high-loading-rate operation and rejection of disturbances. *Biotechnology and Bioengineering*, 87(1), 43-53.

Jenné, R., Banadda, E.N., Gins, G., Deurinck, J., Smets, I.Y., Geeraerd, A.H., and Van Impe, J.F. (2006). Use of image analysis for sludge characterisation: studying the relation between floc shape and sludge settleability. *Water Science and Technology*, 54 (1), 167-174.

Schügerl, K. (2001). Progress in monitoring, modeling and control of bioprocesses during the last 20 years. *Journal of Biotechnology*, 85, 149-173.

Spanjers, H., and van Lier, J.B. (2006). Instrumentation in anaerobic treatment – research and practice. *Water Science & Technology*, 53 (4-5), 63-76.

Van lier, J.B., Tilche, A., Ahring, B.K., Macarie, H., Moletta, R., Dohanyos, M., Hulshoff-Pol, L.W., Lens, P., and Verstraete, W. (2001). New perspectives in anaerobic digestion. *Water Science and Technology*, 43 (1), 1-18.