

## ARTICLE

# Advanced Monitoring of High-Rate Anaerobic Reactors Through Quantitative Image Analysis of Granular Sludge and Multivariate Statistical Analysis

J.C. Costa, I. Moita, A.A. Abreu, E.C. Ferreira, M.M. Alves

Institute for Biotechnology and Bioengineering (IBB), Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal; telephone: +351-253-604-400; fax: +351-253-678-986; e-mail: madalena.alves@deb.uminho.pt

Received 2 May 2008; revision received 17 July 2008; accepted 21 July 2008

Published online 5 August 2008 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bit.22071

**ABSTRACT:** Four organic loading disturbances were performed in lab-scale EGSB reactors fed with ethanol. In load disturbance 1 (LD1) and 2 (LD2), the organic loading rate (OLR) was increased between 5 and 18.5 kg COD m<sup>-3</sup> day<sup>-1</sup>, through the influent ethanol concentration increase, and the hydraulic retention time decrease from 7.8 to 2.5 h, respectively. Load disturbances 3 (LD3) and 4 (LD4) were applied by increasing the OLR to 50 kg COD m<sup>-3</sup> day<sup>-1</sup> during 3 days and 16 days, respectively. The granular sludge morphology was quantified by image analysis and was related to the reactor performance, including effluent volatile suspended solids, indicator of washout events. In general, it was observed the selective washout of filamentous forms associated to granules erosion/fragmentation and to a decrease in the specific acetoclastic activity. These phenomena induced the transitory deterioration of reactor performance in LD2, LD3, and LD4, but not in LD1. Extending the exposure time in LD4 promoted acetogenesis inhibition after 144 h. The application of Principal Components Analysis determined a latent variable that encompasses a weighted sum of performance, physiological and morphological information. This new variable was highly sensitive to reactor efficiency deterioration, enclosing variations between 27% and 268% in the first hours of disturbances. The high loadings raised by image analysis parameters, especially filaments length per aggregates area (LfA), revealed that morphological changes of granular sludge, should be considered to monitor and control load disturbances in high rate anaerobic (granular) sludge bed digesters.

Biotechnol. Bioeng. 2009;102: 445–456.

© 2008 Wiley Periodicals, Inc.

**KEYWORDS:** anaerobic granular sludge; quantitative image analysis; organic loading disturbances; principal component analysis

## Introduction

Anaerobic granules are particulate biofilms, formed spontaneously by self-immobilization of anaerobic bacteria in the absence of a support material (Lettinga, 1995). Hence, each granule is a functional unit comprising of all the different microorganisms necessary for methanogenic degradation of organic matter (Sekiguchi et al., 1998). Over the last three decades, numerous works, theories and models have been developed to explain and/or stimulate the anaerobic granulation process (Liu et al., 2003).

Biological wastewater treatment plants are normally designed with reference to a nominal operating condition, in which the loading rate is assumed to be constant in time. However, in practice this steady-state assumption is seldom met and fluctuations, both in flow and influent concentration occur, which often result in performance degradation or even digester failure.

Anaerobic digesters run frequently at organic loading rates (OLR) below the maximum process capacity. Stable operation of high-rate anaerobic processes is in practice, a difficult task. Due to the slow growth of syntrophic and methanogenic bacteria, anaerobic digestion is quite often unstable during influent fluctuations with volatile fatty acids accumulation and pH decrease, leading sometimes to process failure (Voolapalli and Stuckey, 1998). Depending on the monitoring plan and operator's expertise, the influent disturbances can exhibit different configurations: either as a short/transient overload, which only lasts few hours, or as a longer step change of days to weeks.

Parameters involved in reactor monitoring and control, have been limited to indicators of the liquid and the gaseous phases and have mostly been limited to the assessment of operational performance measurements. The impact on microbial community structure has rarely been investigated (McMahon et al., 2004), although the combination of physical and biological methods for process monitoring

Correspondence to: M.M. Alves

Contract grant sponsor: Fundação para a Ciência e a Tecnologia (Portugal)

Contract grant number: SFRH/BD/13317/2003; POCI/AMB/60141/2004; POCTI/BIO/37934/2001

and control, has been proposed as a way to more effectively minimize problems of process imbalance in anaerobic digesters during load disturbance events (Voolapalli and Stuckey, 1998). There are evidences that microbial and structural changes of biological aggregates can be detected before the detection of significant process performance deterioration after a disturbance (Abreu et al., 2007; Amaral et al., 2004).

The link between digester performance, and the physiological and structural characteristics of the anaerobic sludge is particularly relevant in up-flow anaerobic sludge blanket (UASB), expanded granular sludge bed (EGSB), or internal circulation (IC) reactors (McMahon et al., 2004). Notwithstanding, limited knowledge about physiological and morphological changes of anaerobic granular sludge under unsteady state operating conditions is available in the literature.

The use of quantitative image analysis techniques for monitoring wastewater treatment processes has been initiated by Sezgin et al. (1978), who found a well-defined relationship between the sludge volume index (SVI) and the filament quantity in activated sludge. Recently, Li et al. (2007) compared several characteristics of activated sludge flocs and granules by using different techniques including image analysis. The statistical tool principal component analysis (PCA) was used to relate image analysis information with the sludge settleability in terms of SVI (Jenné et al., 2006). The first attempts to use digital image analysis in high-rate anaerobic digestion processes were limited to size measurements and number counting (Dudley et al., 1993; Jeison and Chamy, 1998). Several subsequent works demonstrated the application of quantitative image analysis tools to monitor granulation and granules deterioration processes (Amaral et al., 2004; Araya-Kroff et al., 2004), to evaluate the effect of detergents in the structure and microbial activity of granular sludge (Costa et al., 2007) or in the detection of structural and microbial changes of granular sludge during a process of acetoclastic activity recovery (Abreu et al., 2007).

Because the experimental approach of integrating reactor performance, physiological and morphological data may produce correlated and redundant data, a statistical instrument should be applied in order to extract the essential information for process monitoring and fault detection applications. Data reduction and interpretation can be approached through the application of multivariate statistical methods, such as PCA (Wise and Gallagher, 1996). 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 (Lee et al., 2006), while keeping information on its variability. It has been successfully applied to the monitoring of industrial processes (Li et al., 2000) and wastewater treatment processes (Lee et al., 2004).

This work aims at studying the effects of organic loading disturbances on EGSB reactors performance, and on anaerobic granular sludge morphology and physiology

using quantitative image analysis techniques and specific methanogenic activity measurements. Then, multivariate statistical tools were applied to integrate the operational, physiological, and morphological information in order to achieve a timely monitoring of the process, with early recognition of operational problems and recovery states.

## Material and Methods

### Experimental Set-Up

Four Plexiglas EGSB reactors (Fig. 1), 1.95 m height and 21 mm internal diameter were used in the experiment. The working volume was 1.15 L and the superficial velocity was  $4.0 \text{ m h}^{-1}$ . Temperature was kept at  $37 \pm 1^\circ\text{C}$  by means of an external jacket for water circulation.

### Inoculum and Substrate

Four hundred milliliters of granular sludge from a lab-scale EGSB reactor, treating a synthetic effluent with ethanol as sole organic carbon source, was used as the inoculum of the four EGSB reactors used in these experiments. The biomass was characterized in terms of specific methanogenic activity (SMA) with acetate, and  $\text{H}_2/\text{CO}_2$  as substrates, morphology by quantitative image analysis, and volatile suspended solids (VSS) (Table I). The reactors were fed with ethanol at a

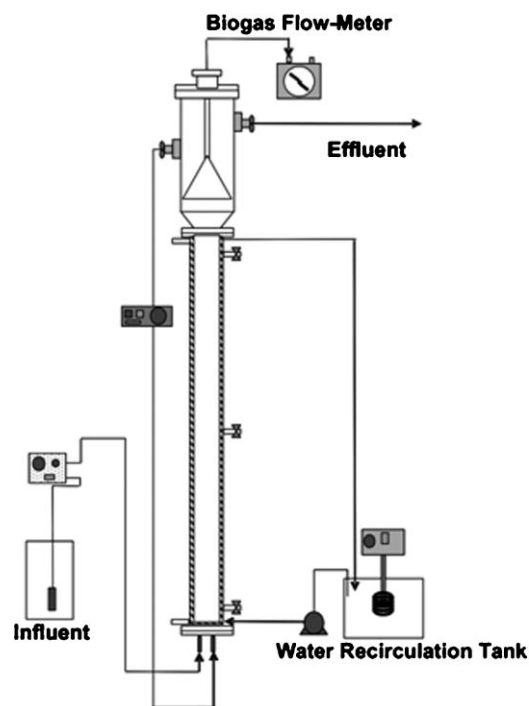


Figure 1. Expanded granular sludge bed reactor.

**Table I.** Inocula characterization and load disturbance conditions.

	LD1	LD2	LD3	LD4
Inocula characterization				
Specific methanogenic activity with individual substrates (mLCH <sub>4@STP</sub> g VSS <sup>-1</sup> day <sup>-1</sup> )				
Acetate	234 ± 15	367 ± 20	234 ± 15	328 ± 12
H <sub>2</sub> /CO <sub>2</sub>	1,427 ± 26	1,686 ± 52	1,427 ± 26	1,720 ± 44
Morphology				
LfA (mm <sup>-1</sup> )	20	18	20	18
TL/VSS (mgVSS <sup>-1</sup> )	1,099	1,238	1,099	1,238
VSS/TA (g m <sup>-2</sup> )	18.4	14.5	18.4	14.5
D <sub>eq</sub> macroflocs (mm)	1.06 ± 0.58	0.91 ± 0.67	1.06 ± 0.58	0.91 ± 0.67
VSS (g L <sup>-1</sup> )	14.78	33.57	14.78	33.57
Disturbance load conditions				
Ethanol (g COD L <sup>-1</sup> )	5	1.6	15	15
HRT (h)	8	2.5	8	8
OLR (kg COD m <sup>-3</sup> day <sup>-1</sup> )	15	15	50	50
Exposure time (h)	72	72	72	384
Recovery phase (h)	163	160	163	168

concentration of 1.5 g COD L<sup>-1</sup>. Sodium bicarbonate was added as the alkalinity source (3 g L<sup>-1</sup>) and macro- and micronutrients were added according to Zehnder et al. (1980). When the reactors were operating in steady-state, four organic load disturbances were applied (Table I). The organic overload was performed by increasing the substrate concentration (LD1, LD3, and LD4), and by decreasing the HRT (LD2).

### Routine Analysis

COD and VSS were determined according to Standard Methods (APHA, 1989). Biogas flow rate was measured by a Ritter Milligascounter (Dr. Ing. Ritter Apparatebau GmbH, Bochum, Germany). Methane content of biogas was determined by gas chromatography using a Porapak Q (100–180 mesh) column, with Helium as the carrier gas at 30 mL min<sup>-1</sup> and thermal conductivity detector. Temperatures of the detector, injector and oven were 110, 110, and 35°C, respectively. Volatile fatty acids (VFA) and ethanol were determined by high performance liquid chromatography using an HPLC (Jasco, Japan) with a Chrompack column (6.5 × 30 mm<sup>2</sup>); sulfuric acid (0.01 N) at a flow rate of 0.7 mL min<sup>-1</sup> was used as mobile phase. Column temperature was set at 60°C. Detection of VFA and ethanol was made sequentially with an UV detector at 210 nm and an RI detector, respectively.

### Sludge Sampling and Dilution

A sampler specially designed to minimize the turbulence inside the reactor was used to remove sludge samples. It consisted of a wide bore tube connected to a 100 mL syringe. It was introduced at the top of the reactor and biomass was collected to the tube, along the expanded granular sludge bed, avoiding mechanical stress. All the sludge samples were characterized by image analysis, SMA assays, and VSS content.

Biomass samples must be diluted for image analysis, but an optimal value should be found. When there are no objects in the image due to an excessive dilution, that image should be registered and considered for calculations. However the observer may unconsciously search objects, over estimating them, and, if the dilution is insufficient, the objects will be overlaid. The optimal dilution value was determined as the lowest dilution that enabled the maximum percentage of objects to be recognized. The percentage of recognition is the ratio between the area of objects that are completely inside the image and the total area of objects in the image including those that are cut-off by the borders of the images and cannot be completely recognized. The dilution 1:5 was determined as the optimal one.

### Specific Methanogenic Activity Assays

The specific acetoclastic activity (SAA) was measured in presence of acetate (30 mM) and the specific hydrogenotrophic methanogenic activity (SHMA) was measured in the presence of H<sub>2</sub>/CO<sub>2</sub> 80:20 (v/v), at 1 bar, according to the previously reported pressure transducer method (Alves et al., 2001). No nutrients were added. Methane was measured by gas chromatography with helium as the carrier gas and a TCD detector. All assays were performed in triplicate.

### Image Acquisition, Processing, and Analysis

For the acquisition of filaments and micro-aggregates ( $D_{eq} < 0.2$  mm) images, a volume of 35 μL from the diluted sample was distributed on a slide and covered with a 20 mm × 20 mm cover slip for visualization and image. This volume was exactly covered by the cover slip. Each image corresponded to a volume of 0.0445 μL. Then, more than 120 images were acquired. Image acquisition was obtained by dividing the cover slip into 42 identical fields and taking a

photo in each square. At least three slides were examined to minimize sampling errors.

Concerning to macro-aggregates ( $D_{eq} \geq 0.2$  mm) images, an arbitrary volume was transferred to a Petri dish for visualization and image acquisition. All the aggregates present in that volume were captured in more than 120 images. Then, the VSS content in the Petri dish was measured.

Images used to quantify filaments and micro-aggregates were acquired through phase contrast and bright field, respectively, on a Nikon Diaphot 300 microscope (Nikon Corporation, Tokyo, Japan) with  $100\times$  magnification. Images used to quantify macro-aggregates were acquired through visualization on a Olympus SZ 40 stereo microscope (Olympus, Tokyo, Japan) with  $15\times$  magnification. All the images were digitized and saved with the help of a CCD AVC D5CE Sony gray scale video camera (Sony, Tokyo, Japan) and a DT 3155 Data Translation frame grabber (Data Translation, Marlboro, MA) with  $768 \times 576$  pixel size in 8 bits (256 gray levels) by Image Pro Plus (Media Cybernetics, Silver Spring, MD) software package.

The metric unit dimensions were further calibrated to pixel using a micrometer slide, for each magnification.

Image processing and analysis was accomplished by means of three programs developed in Matlab (The Mathworks, Inc., Natick, MA), for filaments, micro, and macro-aggregates (Amaral, 2003). In the next paragraphs a brief description of the programs are presented.

The first step of all programs consists in divide the gray-scale image by the background image in order to remove background light differences.

### Filaments Program

A bottom hat filter (Russ, 1995) was applied to enhance the filaments and small aggregates that have low gray levels. The larger aggregates, which have high gray levels, were subsequently identified on the image resulting from the background elimination step by applying a 10-order closing (to enhance the aggregates), a segmentation at a fixed threshold value, a filling of the resulting binary image (to remove the inner holes in the aggregates) and a erosion-reconstruction step to eliminate the debris. Filaments and small aggregates were then isolated by segmentation at a fixed threshold and by logic subtraction of the mask binary image containing the large aggregates. Then, the small aggregates were eliminated by deleting all the objects smaller than 32 pixels (in area) or with a gyration radius below 1.2 (Pons and Vivier, 1999). The final image contained only filaments that were characterized in terms of their length and number.

### Micro- and Macro-Aggregates Programs

Firstly, histogram equalization was performed to enhance the micro-aggregates (this step only occurs in the micro-aggregates program). Afterwards, the image was smoothed by a Wiener filtering (Glasbey and Horgan, 1995). Then, it

segmented in black (background) and white (aggregates) by the simultaneous use of a boundary based segmentation and a user chosen or automatically determined threshold segmentation. The objects smaller than  $3 \times 3$  pixels (small debris) were then removed and small gaps ( $6 \times 6$  pixels or less) were filled on the remaining objects. Subsequently, in order to remove filaments, all the objects smaller than 2,000 pixels in area and with a gyration radius above 1.2 (Pons and Vivier, 1999) were deleted (micro-aggregates program). Finally, all the objects cut off by the image boundaries were removed, and the morphological characterization of the micro- and macro-aggregates were performed.

*Morphological parameters.* The filaments image was skeletonized and pruned (Russ, 1995). The filaments length was determined by:

$$L = N \times 1.122 \times F_{cal} \quad (1)$$

where  $N$  is the number of pixels of the skeletonized filament and  $F_{cal}$  is the calibration factor ( $\mu\text{m pixel}^{-1}$ ). The factor 1.122 is used in order to homogenize the different angles of the filaments (Walsby and Avery, 1996).

The specific total filament length was calculated as:

$$L_{spec} = \frac{L}{V_{field}} \quad (2)$$

where  $V_{field}$  is the volume (in  $\mu\text{m}^3$ ) corresponding to the field of view (i.e., the image).

Filaments are not only the dispersed bulk filaments, but include also those that are attached to an aggregate and still have one free extremity (protruding filaments).

The aggregates program is used to determine:

The aggregate Area from which the Equivalent Diameter ( $D_{eq}$ ) is calculated by:

$$D_{eq} = 2F_{cal} \sqrt{\frac{\text{Area}}{\pi}} \quad (3)$$

And, the total area (AT) occupied by aggregates in each image from which the specific area occupied by aggregates ( $A_{spec}$ ) is calculated:

$$A_{spec} = \frac{A_T}{V_{field}} \quad (4)$$

Finally, morphological parameters representing the dynamic evolution of filaments and aggregates inside the reactor were calculated as:

$$\frac{TL}{VSS} = \frac{L_{spec}}{VSS} \quad (5)$$

$$\frac{VSS}{TA} = \frac{VSS}{A_{spec(<0.2\text{ mm})} + A_{spec(\geq 0.2\text{ mm})}} \quad (6)$$

where VSS are the volatile suspended solids present in each sample, and  $A_{\text{spec}(<0.2 \text{ mm})}$  and  $A_{\text{spec}(\geq 0.2 \text{ mm})}$  are the specific aggregate area ratio for aggregates of equivalent diameter  $<0.2$  and  $\geq 0.2$  mm, respectively.

A morphological parameter based on the ratio of specific total filament length ( $L_{\text{spec}}$ ) to total aggregates projected area (LfA) is determined by:

$$\text{LfA} = \frac{L_{\text{spec}}}{A_{\text{spec}(<0.2 \text{ mm})} + A_{\text{spec}(\geq 0.2 \text{ mm})}} \quad (7)$$

## Principal Components Analysis

Variables summarizing the morphological, physiological and performance data obtained during load disturbances were grouped to create the dataset used to perform the PCA. Four datasets were created, one for each load disturbance.

Principal components analysis aims at finding and interpreting hidden complex, and possibly causally determined, relationships between features in a data set. Correlating features are converted to the so-called factors which are themselves noncorrelated (Einax et al., 1997). PCA modeling, that is, the approximation of a matrix by a model, defined by variables and a relatively small number of outer vector products, shows the correlation structure of a data matrix  $X$ , approximating it by a matrix product of lower dimension ( $T \times P'$ ), called the principal components (PC), plus a matrix of residuals ( $E$ ):

$$X = 1 \times \bar{x}' + T \times P' + E \quad (8)$$

where the term  $1 \times \bar{x}'$  represents the variable averages. The second term, the matrix product  $T \times P'$ , models the structure, and the third term,  $E$ , is a matrix of residuals, containing the deviations between the original values and the projections, that is, contains the noise.  $T$  is a matrix of scores that summarizes the  $X$ -variables (Scores), and  $P$  is a matrix of loadings showing the influence of the variables on each score. Geometrically, it corresponds to fitting a line, a plane or a hyper plane to the data in the multidimensional space, with the variables as axes. The scaling of the variables specifies the length of the axes of this space.

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, that is, 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 by:

$$Z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (9)$$

where  $x_{ij}$  is the value of the variable  $j$  in the sample  $i$ ,  $\bar{x}_j$  and  $s_j$  are the mean and the standard deviation of the variable  $j$ , respectively, and,  $z_{ij}$  is the autoscaled value of  $x_{ij}$ . At the end of this standardization, each variable has mean zero and unit standard deviation.

Subsequently, SIMCA-P (Umetrics AB) 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) is 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.

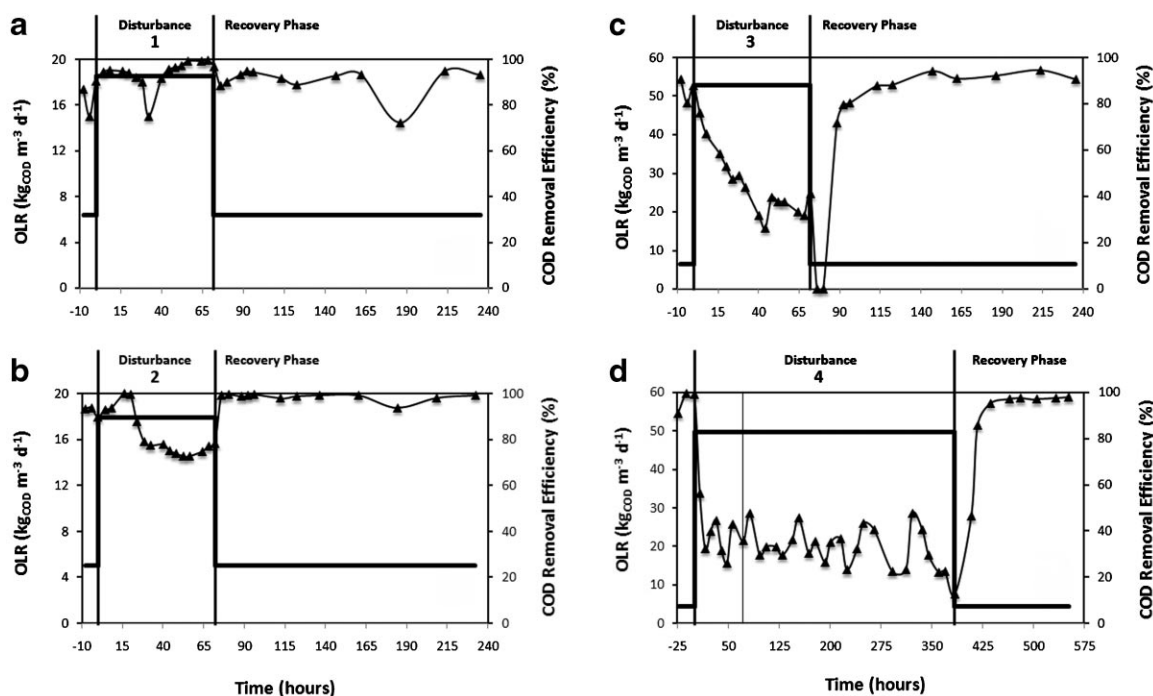
## Results and Discussion

### Reactor Performance

Four load disturbances (LD) were applied when the EGSB reactors were operating in steady state with organic loading rates (OLR) of roughly  $5 \text{ kg COD m}^{-3} \text{ day}^{-1}$  and hydraulic retention time (HRT) of 8 h. The COD removal efficiencies were in the range 80–90% (Fig. 2). In LD1 and LD2 the organic loading rates (OLR) were increased to  $18.5 \text{ kg COD m}^{-3} \text{ day}^{-1}$ , through the substrate (ethanol) concentration increase and the hydraulic retention time (HRT) decrease to 2.5 h, respectively. The LD3 and LD4 were performed by increasing the substrate concentration (OLR of  $50 \text{ kg COD m}^{-3} \text{ day}^{-1}$ ) during 3 and 16 days, respectively.

The load disturbance LD1 did not cause deterioration in the reactor performance. After a temporary decrease, the COD removal efficiency increased to more than 95%, suggesting that the reactor was operating, in the pre-disturbance conditions, far from its maximum capacity. In contrast, when the flow rate increased in LD2, the COD removal efficiency decreased and stabilized around 73%, until the end of the disturbance. This led to a temporary accumulation of acetate in the effluent that peaked to values of  $300 \text{ mg COD L}^{-1}$  (data not shown).

When the OLR increased to  $50 \text{ kg COD m}^{-3} \text{ day}^{-1}$  (LD3 and LD4), the COD removal efficiency decreased linearly



**Figure 2.** Time course of organic loading rate (OLR, —), and COD removal Efficiency (▲). a: LD1, (b) LD2, (c) LD3, (d) LD4.

from 90% to 30% (Fig. 2c) in the first 40 h of disturbance. After this period, and extending the exposure time (LD4, Fig. 2d), the COD removal efficiency was recurrently around 35%. Few hours after the end of the shock exposure period, all the reactors regained the pre-disturbance efficiencies (Fig. 2).

Accumulation of acetate was observed in LD3 and LD 4, achieving a temporary peak of 1,400 and 1,000 mg COD L<sup>-1</sup>, respectively. A fast decrease of pH from 7.5 to <6 was observed in the first hours of disturbances (data not shown). In LD3, the pH remained near 6, during all the period of shock exposure, whereas in LD4, a gradual increase in the pH from 5.8 to 7 was observed in the last 150 h of exposure. The pH increase was simultaneous with a decrease in acetate concentration and an increase in the ethanol concentration up to 5,000 mg COD L<sup>-1</sup> in LD3 and 9,000 mg COD L<sup>-1</sup> in LD4, suggesting an inhibition of acetogenic reaction between ethanol and acetate.

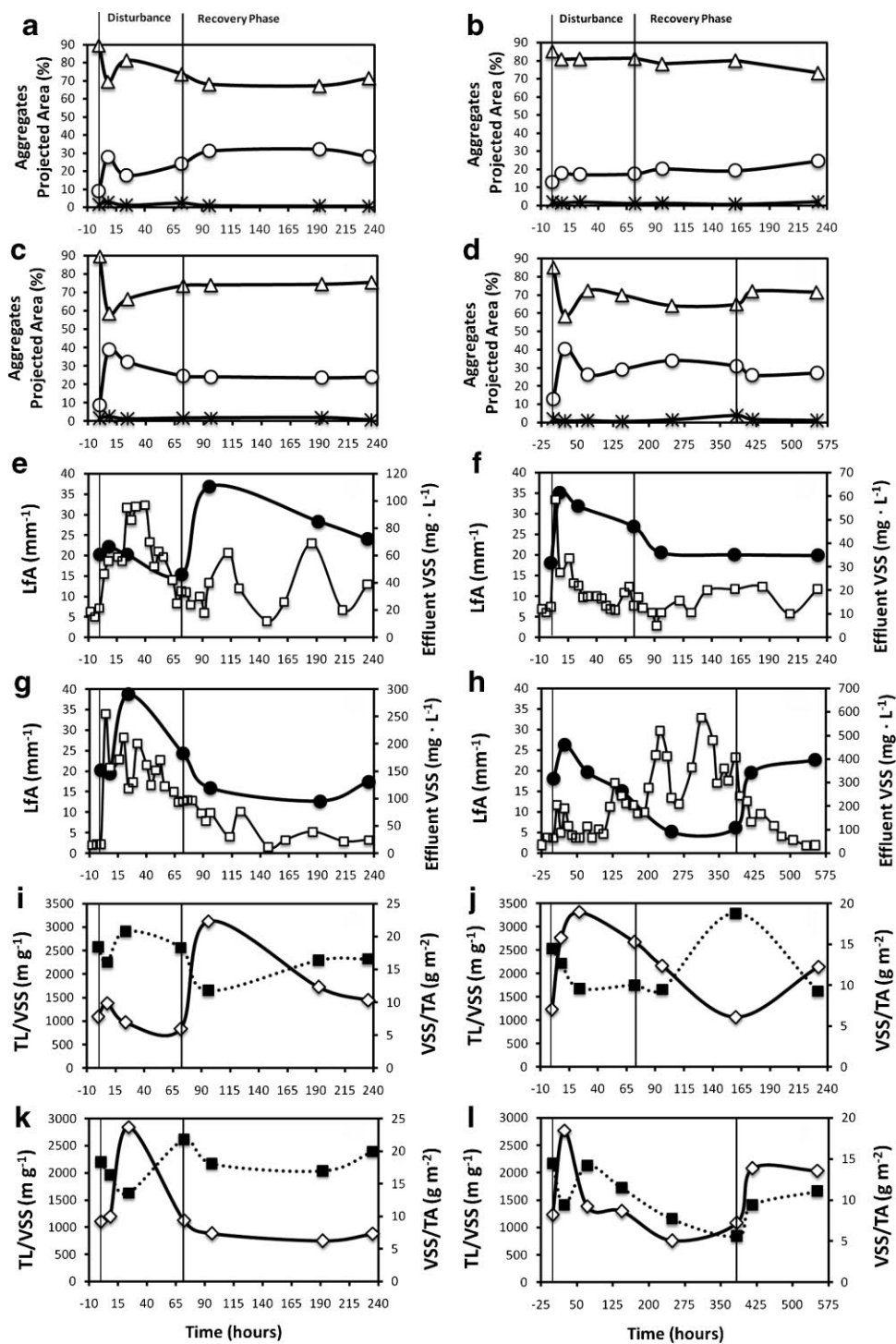
### Morphological, Structural and Physiological Data

Quantitative morphological information retrieved by image analysis techniques can be related with the operational and physiological information in order to interpret the impact of load disturbances in the structure of granular sludge and consequently its impact in the sludge bed stability.

In the organic shocks imposed by increasing the influent concentration (LD1, LD3, and LD4) it was clearly observed a decrease in aggregates size. The percentage of aggregates

projected area with equivalent diameter ( $D_{eq}$ ) > 1 mm decreased to 70% in LD1, and to 58% in LD3 and LD4. Concurrently, the % of projected area of aggregates inside the range of  $D_{eq}$  0.1–1 mm increased to 28% and 40%, respectively, in LD1 and LD3/4 (Fig. 2a, c, and d). A higher flow of biogas production through the granular structure, induced by the increase in the substrate concentration, is likely responsible for the observed fragmentation, being expectable that the respective impact could be dependent on the magnitude of the load disturbance. In the hydraulic shock LD2, a negligible change in size distribution was observed (Fig. 3b). However, in this experiment, an increase in the total filaments length per volatile suspended solids (TL/VSS) was observed immediately (8 h) after starting the disturbance period (Fig. 3j), likely induced by the increasing shear stress inside the reactor, due to the higher up-flow velocity applied. In LD1 a negligible increase of free/protruding filaments was observed (Fig. 3i), and in LD3/4 a peak of filaments occurred only 24 h after increasing the load (Fig. 3k and l).

It is interesting to observe that, although the granules were fragmented, LD1 disturbance did not impact the reactor performance, but LD2 affected the reactor efficiency, likely due to the erosion of filamentous bacteria from the granules, followed by detachment and washout. Therefore, by evaluating reactor's performance (Fig. 2a and b), it can be concluded that stability in the granules size distribution is of minor importance when compared to the capacity of filaments retention in the granular microbial structures.



**Figure 3.** Time course of morphological parameters: (a–d) percentage of aggregates projected area by equivalent diameter ( $D_{eq}$ ) ranges ( $\Delta$ ,  $D_{eq} \geq 1$  mm;  $\circ$ ,  $0.1 \text{ mm} \leq D_{eq} < 1$  mm;  $\times$ ,  $D_{eq} < 0.1$  mm); (e–h) dynamic between total filament length and total aggregates projected area (LfA,  $\bullet$ ), and, effluent volatile suspended solids (effluent VSS,  $\square$ ); and (i–l) total filament length per volatile suspended solids (TL/VSS,  $\diamond$ ), and volatile suspended solids per total aggregates projected area (VSS/TA,  $\blacksquare$ ). a, e, i: LD1; (b, f, j) LD2; (c, g, k) LD3; and (d, h, l) LD4.

Reactor performance deterioration was also subsequent to a high release of filaments in LD3 and LD4. Although granular sludge is a complex structure, it is accepted that, in general, acetoclastic bacteria are located mainly in the

core of a stratified microbial structure (Harmsen et al., 1996; Liu et al., 2002a). This fact allows them to be more protected against operational problems such as overloads, pH oscillations or presence of toxic compounds (Liu

et al., 2002b). In result of the observed fragmentation, the inner core of the granules becomes exposed to the bulk conditions, promoting the release and the selective washout of methanogenic bacteria. Because filamentous forms increased in the bulk, in LD2, LD3, and LD4 (Fig. 3j, k, and l), it is worth to suggest that filamentous forms could have been selectively washed-out during those load disturbances, causing a decrease in the COD removal efficiency (Fig. 2b–d). The observed decrease in the specific acetoclastic activity (SAA), especially during LD3 and LD4 (Fig. 4), the increase of effluent VSS (Fig. 3f, g, and h) and the kinetic characteristics of the *Methanosaeta* genera, which is a very slow growing archaea may allow to speculate that *Methanosaeta* genera could have been selectively washed out during LD2, LD3, and LD4. Apparently, extending the exposure time in LD4, the deterioration of reactor performance was persistent and simultaneous with the increase in effluent VSS concentration. This was likely due to filaments release, detachment and washout, inducing minimal values, both for filaments (Fig. 3l) and SAA (Fig. 4).

It was observed in different previous works (Abreu et al., 2007; Araya-Kroff et al., 2004) that aggregation and densification of aggregates was simultaneous with an increase in the SHMA. In this respect, it is remarkable that, particularly during the recovery phase, the SHMA (Fig. 4) and the VSS/TA (Fig. 3) presented similar trends in LD4.

The LfA parameter quantifies the level of biomass aggregation by measuring the ratio between the free filaments length and the projected area of the aggregates. A high correlation with the total filaments length per VSS was observed. Amaral (2003) and Costa et al. (2007) proposed that LfA could be an early alert of washout during influent disturbances in EGSB reactors, particularly when

toxic substances were accidentally fed. In the present study, however the effects on the morphology caused by the overloads were simultaneous with the increase in the effluent VSS. Therefore, the LfA parameter followed the trend of effluent VSS without advance (Fig. 3e–h), forbidding an earlier detection of the washout phenomenon.

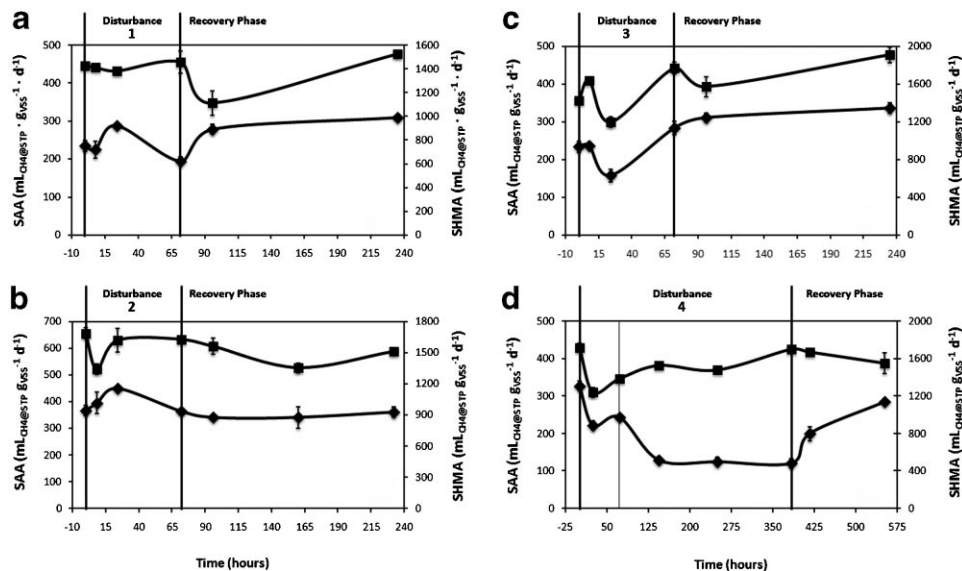
By the end of the exposure periods, the turbulence and instability caused by the raise in biogas production inside the reactors, decreased. The effluent VSS decreased to values similar to the pre-disturbance conditions (Fig. 3e–h). During the recovery phase the morphological parameters stabilized in values close to the initial ones, with the exception of aggregates size distribution (Fig. 3).

### Principal Components Analysis

The extraction of three principal components gathered more than 80% of the total variability in all four datasets. The use of other PCs did not enhance significantly the correlation factors.

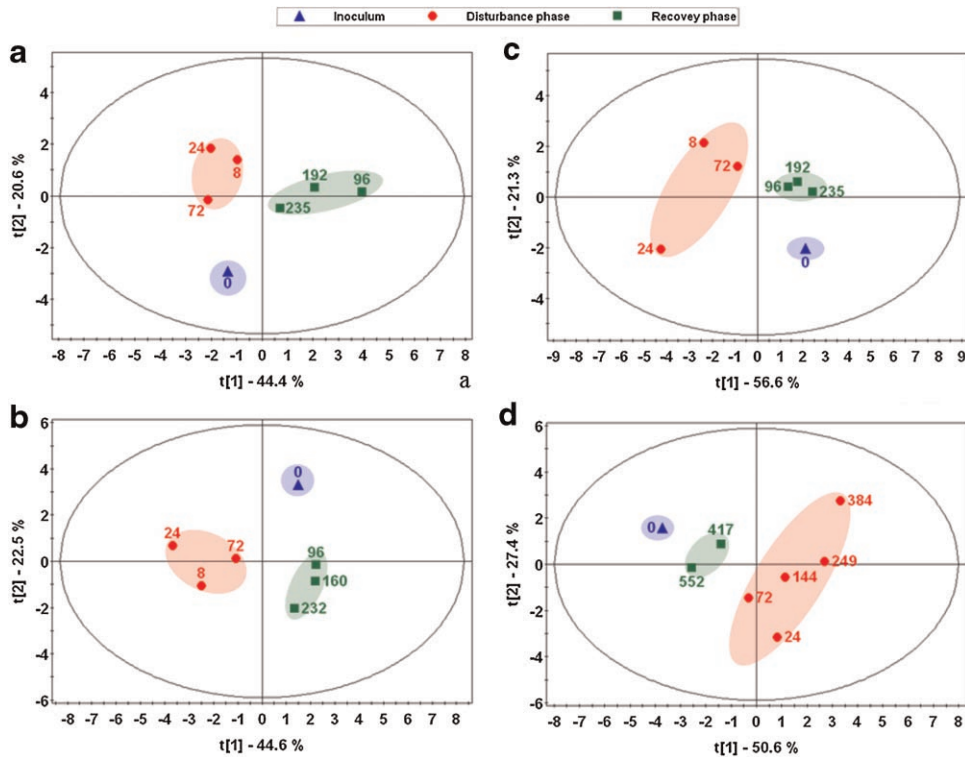
The score maps (Fig. 5) can be seen as windows in the X space, displaying the observations as situated on the projection planes or hyperplanes, and revealing groups, trends, outliers and similarities between samples. Two clusters clearly isolate the observations during the load disturbances and the recovery period, besides the inoculum sample, which emerge as an isolated observation. This shows that immediately after the load disturbance, a deviation occurred.

The information that can be obtained about the variables and the role they play in the principal component (PC), resides in the so-called loadings and respective loading maps (Fig. 6). It allows deciding which variables are most



**Figure 4.** Time course of specific acetoclastic activity (SAA, ◆), and specific hydrogenotrophic methanogenic activity (SHMA, ■). a: LD1, (b) LD2, (c) LD3, (d) LD4.





**Figure 5.** PCA score plot of the first PC ( $t[1]$ ) versus the second PC ( $t[2]$ ), in dataset of: (a) LD1; (b) LD2; (c) LD3; and (d) LD4. [Color figure can be seen in the online version of this article, available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

important for the differences observed between the samples. The loading plots can be very useful to detect correlations between variables. Variables close to each other, that is, with similar coordinates, represent strongly correlated variables. Concurrently, variables with symmetric coordinates are correlated, although inversely proportional. Combination of score and loading maps (Figs. 5 and 6) allows for the visualization of the main effects/problems occurred during EGSB reactors operation. For instance, the inoculum sample is located in the upper part of the plot, with high score in PC2 (Fig. 5b,  $t[2]$ ). The variables that most influence this PC are  $>1$ ,  $>0.1$ , SHMA, Eff, and VSS (Fig. 6b,  $p[2]$ ) that evidenced immediate changes when the disturbance was applied. The main clusters refer to observations with negative (disturbance) and positive (recovery) scores in PC1 (Fig. 5b,  $t[1]$ ). The variables with higher loadings in PC1 (Fig. 6b,  $p[1]$ ), that is, responsible for grouping disturbance and recovery phase observations are HRT, pH, SAA, OLR, LfA, and TL/VSS. The LfA showed the highest variation caused by the decrease of HRT, that is, the highest loading (Fig. 6b,  $p[1]$ ).

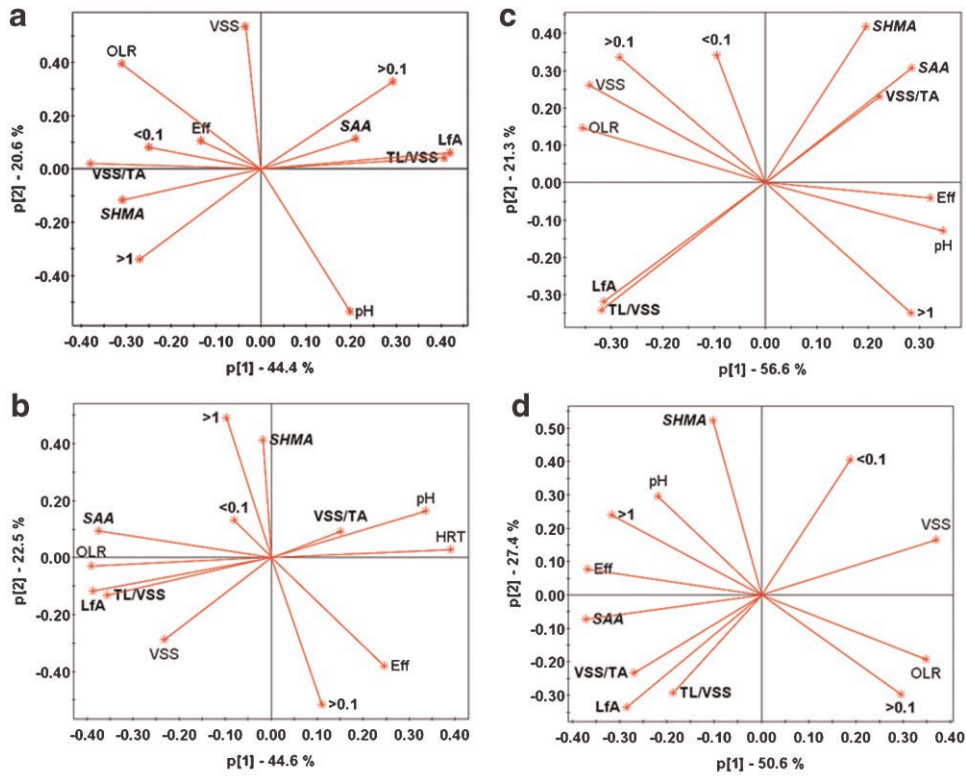
Some correlations should be distinguished from the loading maps. It was observed a high positive correlation between LfA and TL/VSS parameters in all datasets. Previously some hypotheses were postulated, regarding correlations between reactor performance deterioration,

biomass activity, and erosion/fragmentation of granules and consequent filaments release and washout.

Emphasis should be given to the correlation showed in LD3 (Fig. 6c) between the apparent granules density (VSS/TA) and the specific hydrogenotrophic methanogenic activity (SHMA), supporting the hypothesis presented earlier, where densification of aggregates was simultaneous with an increase in the SHMA (Abreu et al., 2007).

Special care should be put in the analysis of the loading plot of LD4 (Fig. 6d), because extending the period of disturbance, induced a new dynamic of filaments release and washout. Consequently, the correlations detected were poorer in this case. To elucidate this new dynamic, in Figure 6d, it is observed a negative correlation between the efficiency (Eff) and the % of aggregates with equivalent diameter within the range 0.1–1 mm ( $>0.1$ ). The increase of smaller aggregates suggests the occurrence of erosion and/or granules fragmentation. This phenomenon could have destroyed the granules structure and architecture, likely, exposing the inner core of acetoclastic bacteria to the bulk conditions. The decrease in the removal efficiency was probably associated to the selective washout of those bacteria, evidenced by the negative correlation between the SAA and the effluent VSS (Fig. 6d).

The increasing number of reports with methods for monitoring and control of anaerobic digestion processes



**Figure 6.** PCA loading plot of the first and second principal components ( $p[1]$  vs.  $p[2]$ ), from dataset of: (a) LD1; (b) LD2; (c) LD3; and (d) LD4. [Color figure can be seen in the online version of this article, available at [www.interscience.wiley.com](http://www.interscience.wiley.com).]

does not consider the integration of quantitative morphological indicators. In the present work, the application of PCA to datasets obtained during load disturbances, allowed the establishment of a new latent variable that encompasses a weighted sum of performance, physiological and morphological information. This new variable can be used as a warning indicator of operational problems during load disturbances as can be clearly observed in Table II—recognition of load disturbance. The variable  $t[1]$  was calculated in the first hours of operation in the load disturbances and the corresponding % of variation in  $t[1]$  ranged between 27% and 268%, evidencing the high sensitivity of the latent variable to detect the disturbance. In LD1, the percentage of variation after 8 h was the smallest one (27%) because this was the softer shock applied. Simultaneously, was observed that this variable,  $t[1]$ , was also sensible to detect the recovery of the reactors

(Table II—recognition of recovery state). This was possible due to the rapid recovery in reactor performance and morphological changes that occurred few hours after the end of the load disturbance.

For a better understanding of principal components analysis and the factors involved, is necessary to remember that the score ( $t_i$ ) of an observation ( $i$ ) on a principal component  $PC_j$  ( $t_i PC_j$ ) is the weighted sum of the original variables ( $x_i$ ). The weights ( $p_i PC_j$ ) are called the loadings of the variables on that  $PC_j$ . The loading of a variable is related to its variation (Massart and Vander Heyden, 2005)

$$t_i PC_j = \sum p_i PC_j x_i \quad (10)$$

Analyzing the loadings associated with the new latent variable  $t[1]$  described previously it is possible to distinguish

**Table II.** Coefficients of the first latent variable  $t[1]$  in PC1, at the beginning and after few hours of load disturbance (recognition of load disturbance), and, at the end of load disturbance and in the beginning of recovery phase (recognition of recovery state).

LD	Recognition of load disturbance				Recognition of recovery state				
	$t=0$ h	$t=8$ h	$t=24$ h	% of variation	$t=72$ h	$t=96$ h	$t=384$ h	$t=417$ h	% of variation
1	-1.343	-0.976	—	27	-2.119	3.938	—	—	286
2	+1.482	-2.493	—	268	-1.069	2.229	—	—	308
3	+2.132	-2.364	—	211	-0.875	1.344	—	—	254
4	-3.701	—	0.817	122	—	—	3.329	-1.386	142

**Table III.** Loadings/weights of the variables in datasets associated to PC1.

Variable	LD1	LD2	LD3	LD4	Observation
HRT	—	0.390	—	—	Controlled variable
OLR	-0.309	-0.390	-0.355	0.348	Controlled variable
Eff	-0.134	0.245	0.320	-0.368	
pH	0.197	0.336	0.345	-0.219	
VSS	-0.036	-0.232	-0.341	0.369	
<0.1	-0.250	-0.080	-0.094	0.189	Morphological variable
>0.1	0.293	0.111	-0.283	0.295	Morphological variable
>1	-0.270	-0.097	0.283	-0.316	Morphological variable
SAA	0.211	-0.375	0.284	-0.372	
SHMA	-0.308	-0.018	0.196	-0.102	
LfA	0.419	-0.387	-0.314	-0.284	Morphological variable
TL/VSS	-0.380	0.151	0.222	-0.270	Morphological variable
VSS/TA	0.407	-0.356	-0.318	-0.186	Morphological variable

the variables that mostly influence the early detection of reactor. The morphological parameters, with emphasis to the LfA parameter that relates the free filaments length per aggregates projected area emerges due to its high loadings, that is, weights (absolute value) in all datasets (Table III). The results clearly suggest that quantitative morphological parameters should be considered in monitoring high rate anaerobic reactors, especially those based on granular sludge. It is important to mention, in this respect, that granular sludge is a highly valuable resource used to start-up industrial UASB reactors, EGSB reactors, or IC reactors. If available, cost of granular sludge can attain about 300 Euro per ton, including 2,000 km transportation. Therefore, physiological quality but also morphological stability should be assessed when operating in a different reactor or in different conditions.

The integration of morphological information quantified by image analysis, reactor performance, specific activity, and chemometric methods constitutes a package for advanced monitoring of anaerobic granular sludge bed reactors, which in being tested at industrial scale with promising results.

## Conclusions

The effects of organic load disturbances on the operational, physiological and morphological properties of anaerobic granular sludge were studied by combining quantitative image analysis, methanogenic activity assays and performance data.

An increase of OLR from 5 to 18.5 kg COD m<sup>-3</sup> day<sup>-1</sup> caused by the increase in influent concentration did not affect the reactor efficiency. However, when the same increase was due to the HRT decrease or the OLR was increased to 50 kg COD m<sup>-3</sup> day<sup>-1</sup>, the COD removal efficiency decreased to 72% and 30%, respectively. In the last case, increasing the exposure time caused acetogenic inhibition. Concomitantly, granules fragmentation/erosion was observed in organic shock loads (LD1, LD3, LD4). However, in hydraulic shock (LD2), a negligible change in size distribution was observed although an increase in filaments was observed. Therefore, by evaluating reactor's

performance, it can be concluded that stability in the granules size distribution is of minor importance when compared to the capacity of filaments retention in the granular microbial structures.

It was demonstrated that the use of a multivariate statistical tool such as PCA was appropriate to visualize and isolate the main effects caused by the transient instabilities. The proposed morphological parameters proved to be more sensitive to detect promptly the problems than the normal operating parameters. The new latent variable  $t[1]$ , defined as an weighted sum of all variables included in the dataset, showed high variations in the first hours of disturbance and in the recognition of the recovery state, in all datasets. Concurrently, the high loadings/weights of the morphological parameters enhanced the need to monitor the anaerobic digestion process solid phase in order to achieve an effective and timely control of the process.

## Nomenclature

### Reactor Performance Data

OLR	organic loading rate
HRT	hydraulic retention time (only used in LD2)
Eff	chemical oxygen demand removal efficiency
pH	pH
VSS	effluent volatile suspended solids

### Physiological Data

SAA	specific acetoclastic activity
SHMA	specific hydrogenotrophic methanogenic activity

### Morphological Data

LfA	total filaments length per total aggregates projected area
TL/VSS	total filaments length per volatile suspended solids
VSS/TA	volatile suspended solids per total aggregates projected area (apparent granules density)
>1	percentage of aggregates projected area with equivalent diameter ( $D_{eq} \geq 1$ mm)
>0.1	percentage of aggregates projected area within the range $0.1 \leq D_{eq}$ (mm) < 1 mm
<0.1	percentage of aggregates projected area with $D_{eq} < 0.1$ mm

We thank the financial support to J.C. Costa, I. Moita, and A. Abreu through the grant SFRH/BD/13317/2003 and the projects POCI/AMB/60141/2004 and POCTI/BIO/37934/2001, respectively, from the Fundação para a Ciência e a Tecnologia (Portugal).

## References

- Abreu AA, Costa JC, Araya-Kroff P, Ferreira EC, Alves MM. 2007. Quantitative image analysis as a diagnostic tool for identifying structural changes during a revival process of anaerobic granular sludge. *Water Res* 41(7):1473–1480.
- Alves MM, Mota Vieira JA, Álvares Pereira RM, Pereira MA, Mota M. 2001. Effects of lipids and oleic acid on biomass development in anaerobic fixed bed reactors. Part I. Biofilm growth and activity. *Water Res* 35:255–263.

- Amaral AL. 2003. Image Analysis in Biotechnological Processes: Application to Wastewater Treatment. PhD dissertation, Braga: University of Minho, Portugal. (<http://hdl.handle.net/1822/4506>).
- Amaral AL, Pereira MA, da Motta M, Pons M-N, Mota M, Ferreira EC, Alves MM. 2004. Development of image analysis techniques as a tool to detect and quantify morphological changes in anaerobic sludge. II. Application to a granule deterioration process triggered by contact with oleic acid. *Biotechnol Bioeng* 87(2):194–199.
- APHA, AWWA, WPC. 1989. Standard methods for the examination of water and wastewater. 17th edn. Washington, DC: American Public Health Association.
- Araya-Kroff P, Amaral AL, Neves L, Ferreira EC, Pons M-N, Mota M, Alves MM. 2004. Development of image analysis techniques as a tool to detect and quantify morphological changes in anaerobic sludge. I. Application to granulation process. *Biotechnol Bioeng* 87(2):184–193.
- Costa JC, Abreu AA, Ferreira EC, Alves MM. 2007. Quantitative image analysis as a diagnostic tool for monitoring structural changes of anaerobic granular sludge during detergent shock loads. *Biotechnol Bioeng* 98(1):60–68.
- Dudley BT, Howgrave-Graham AR, Bruton AG, Wallis FM. 1993. Image analysis to quantify and measure UASB digester granules. *Biotechnol Bioeng* 42(3):279–283.
- Einax JW, Zwanziger HW, Geiss S. 1997. Chemometrics in environmental analysis. Weinheim: VCH.
- Glasbey CA, Horgan GW. 1995. Image analysis for the biological sciences. Chichester: John Wiley and Sons.
- Harmsen HJM, Kengen HMP, Akkermans ADL, Stams AJM, de Vos WM. 1996. Detection and localization of syntrophic propionate-oxidizing bacteria in granular sludge by in situ hybridization using 16S rRNA-based oligonucleotide probes. *Appl Environ Microbiol* 62(5):1656–1663.
- Jeison D, Chamy R. 1998. Novel technique for measuring the size distribution of granules from anaerobic reactors for wastewater treatment. *Biotechnol Bioeng* 12(9):659–662.
- Jenné R, Banadda EN, Gins G, Deurinck J, Smets IY, Geeraerd AH, Van Impe JF. 2006. Use of image analysis for sludge characterisation: Studying the relation between floc shape and sludge settleability. *Water Sci Technol* 54(1):167–174.
- Lee JM, Yoo CK, Choi SW, Vanrolleghem PA, Lee IB. 2004. Nonlinear process monitoring using kernel principal component analysis. *Chem Eng Sci* 59:223–234.
- Lee DS, Lee MW, Woo SH, Kim Y-J, Park JM. 2006. Multivariate online monitoring of a full-scale biological anaerobic filter process using kernel-based algorithms. *Ind Eng Chem Res* 45:4335–4344.
- Lettinga G. 1995. Anaerobic digestion and wastewater treatment systems. *Antonie van Leeuwenhoek* 67:3–28.
- Li W, Yue H, Valle-Cervantes S, Qin SJ. 2000. Recursive PCA for adaptive process monitoring. *J Process Control* 10:471–486.
- Li J, Garry K, Neu T, He M, Lindenblatt C, Horn H. 2007. Comparison of some characteristics of aerobic granules and sludge flocs from sequencing batch. *Water Sci Technol* 55(8–9):403–411.
- Liu WT, Chan O-C, Fang HHP. 2002a. Characterization of microbial community in granular sludge treating brewery wastewater. *Water Res* 36:1767–1775.
- Liu Y, Xu H-L, Show K-Y, Tay J-H. 2002b. Anaerobic granulation technology for wastewater treatment. *World J Microbiol Biotechnol* 18:99–113.
- Liu Y, Xu H, Yang S, Tay J. 2003. Mechanisms and models for anaerobic granulation in upflow anaerobic sludge blanket reactor. *Water Res* 37:661–673.
- Massart DL, Vander Heyden Y. 2005. From tables from visuals: Principal component analysis. Part 2. Practical data handling. *LC-GC* 18(2):84–89.
- McMahon KD, Zheng D, Stams AJM, Mackie RI, Raskin L. 2004. Microbial population dynamics during start-up and overload conditions of anaerobic digesters treating municipal solid waste and sewage sludge. *Biotechnol Bioeng* 87(7):823–834.
- Pons MN, Vivier H. 1999. Biomass quantification by image analysis. *Adv Biochem Eng/Biotechnol* 66:133–184.
- Russ CR. 1995. The image processing handbook. Boca Raton, FL: CRC Press.
- Sekiguchi Y, Kamagata Y, Nakamura K, Syutsubo K, Ohashi A, Harada H, Nakamura K. 1998. Diversity of mesophilic and thermophilic granular sludge determined by 16S rRNA gene analysis. *Microbiology* 22:2655–2665.
- Sezgin M, Jenkins D, Parker DS. 1978. A unified theory of filamentous activated sludge bulking. *J Water Pollut Control Fed* 50:362–381.
- Voolapalli RK, Stuckey DC. 1998. Stability enhancement of anaerobic digestion through membrane gas extraction under organic shock loads. *J Chem Technol Biotechnol* 73:153–161.
- Walsby AE, Avery A. 1996. Measurement of filamentous cyanobacteria by image analysis. *J Microbiol Methods* 26:11–20.
- Wise BM, Gallagher NB. 1996. The process chemometrics approach to process monitoring and fault detection. *J Proc Cont* 6(6):329–348.
- Zehnder AJB, Huser BA, Brock TD, Wuhrmann K. 1980. Characterization of an acetate-decarboxylating, non-hydrogen-oxidizing methane bacterium. *Arch Microbiol* 124:1–11.