Multivariate monitoring of an activated sludge process for biological treatment of a synthetic wastewater effluent

A.M.A. Dias, A. Paulo, D.P. Mesquita, M.M. Alves, E.C. Ferreira*
IBB – Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710–057 Braga, Portugal

Keywords: activated sludge, monitoring, principal component analysis, spectroscopy

To face the steadily increasing global water problems, the search for alternative effective ways to monitor biological wastewater treatment processes is essential when changes should be detected as soon as possible so measures could be taken accordingly to avoid major damage to the biological process and to the environment. Besides the need for rapid monitoring, these techniques should have no addition of chemicals, avoiding its consumption, its emission to the wastewater and transportation for adequate treatment of the residues resulting from analysis.

Multivariate statistical tools and spectroscopic techniques are among the present valuable alternatives for these purposes. Besides the advantages already referred, spectroscopic techniques can be used without any sample pre-treatment, what makes them interesting for online monitoring. In this work they were applied to monitor a lab scale plant treating synthetic wastewater. The system was disturbed to promote changes that simulate real possible scenarios in order to verify if the technique can efficiently detect the induced variations.

Experiments were carried out in an activated sludge plant consisting of a 25 L tank with 17 L of suspended biomass and a cylindrical settler of 2.5 L, with recirculation of biomass from the settler to the tank using an air pump. The complete mixing inside the reactor is guaranteed by the diffusion of air from its bottom. During the monitoring period the concentration of dissolved oxygen was maintained in excess. The reactor was inoculated with biomass collected from a municipal WWTP designed to a population of 214.605 inhabitants. Synthetic wastewater was prepared with a mixture of peptone and meat extract as carbon source, urea, K_2HPO_4, NaCl, $CaCl_2$ and $MgSO_4$.

During the monitoring period, influent flow (Q_{in}), COD_{in}, COD_{out}, TSS and VSS in the reactor and in the effluent, N-NO_3^- in the effluent and pH in the reactor were measured.

N-NO_3^- was measured using HPLC and remaining parameters were analyzed according to Standard Methods. Each sample was analyzed in triplicate. An immersible probe and a USB4000 portable dispersive UV-Vis detector, both from Ocean Optics, were used to acquire spectra in the settler in the wavelength range from 230 to 700nm. In order to create system imbalances the process was disturbed by i) slight increase of the organic load (with simultaneous decrease of HRT and increase of COD_{in}); ii) a systematic removal of sludge from the reactor after a long period without purging biomass from the system; and iii) sudden decrease of HRT, by this order. Principal Component Analysis (PCA) was used to analyze all data. PCA is a linear mathematical method used to find patterns in large data matrices. The algorithm reads data as vectors in a multidimensional space and seeks the direction (vector) in which the variance is the largest. These vectors are designed principal component (PC) vectors and are usually numbered (PC1, PC2, ...) in order of importance. Typically, the variance in the data set is modeled within the first few PCs, depending on the complexity of the sample matrix. This operation reduces the number of vectors needed to describe the variance in the matrix. The resulting PCA matrix can be plotted and studied in order to identify clusters of samples and time dependent variations.

* Corresponding author. Tel + 351 253 604 407. E-mail:ecferreira@deb.uminho.pt

PCA was applied to the auto-scaled set of eleven variables measured during the monitoring period such as: Q_{in}, COD_{in}, COD_{out}, Organic Load, TSS_r and VSS_r in the reactor, TSS_{out}, VSS_{out}, food-to-microorganisms ratio (F/M), $pH_{reactor}$ and $N\text{-}NO_3^-{}_{out}$. With two principal components (PCs) it was possible to explain 71.70% of the variation in the data. Figure 1a shows the biplot obtained presenting in the same graph the samples and the measured variables.

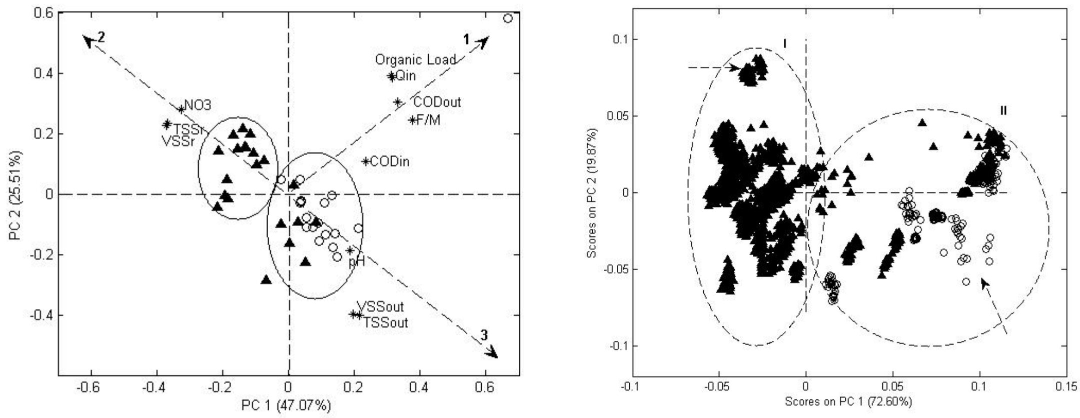

Figure 1. (a) Biplot representing simultaneously the samples and the variables measured during the monitoring period. (b) Score plot representing the UV-Vis spectra variation during the monitoring period. Symbols are explained in the text.

The major advantage of the biplot representation is the possibility of establishing relations between samples and variables. Moreover it is possible to identify how the different variables are related with each other. In this case samples are divided in two main groups – I and II – corresponding to the first and to the second monitoring period respectively. Group I is characterized by a higher nitrification rate, when disturbances i) and ii) were imposed, while group II is associated to a low or even inexistent nitrification process, when disturbance iii) was applied and purges were more frequent. It is possible to identify some samples measured during the first monitoring period in group II. These samples correspond to the analysis performed during the system start-up and after purging biomass from the reactor. In both cases the nitrification rate was low what explains this displacement. Attending the variables, these are also divided in groups according to their influence in the system. It is interesting to notice how PCA can easily identify the close relation between nitrate and biomass concentration in the reactor and how the presence of nitrate is inversely related with the pH in the reactor. The monitored samples are placed along the arrows #2 or #3 depending if they correspond to a high or low nitrification period, respectively. The organic load increase induced during the first period is only slightly noticed by a displacement of the samples in the direction of the arrow #1. However, the influent flow increase during the second monitoring period is clearly identified in the right upper zone of the figure. This result can be explained by the fact that the disturbance imposed to the process in the first period was less significant when compared with the one induced in the second monitoring period. Similar analysis is achieved when spectra are also analyzed with PCA. Spectra were pre-treated by applying a first derivative and mean centering the raw data. With two PCs it was possible to explain 92.5% of the data. The score plot obtained is shown in Figure 1b. Once more the two periods mentioned earlier are clearly distinguished. Spectra acquired during the two different organic load conditions induced to the system – i) and iii) – are indicated by dashed arrows in the score plot. This study points out that spectroscopy associated with chemometrics are potential tools to efficiently, rapidly and intuitively follow-up variations in the system in an equivalent way to the use of different monitoring parameters.