Image analysis is, nowadays, an important complement to analytical survey procedures in biotechnological processes. Furthermore, as the quantity of the resulting data can become overwhelming, it is usual for these data to be fed into a multivariate statistical analysis procedure. In the case being considered, an activated sludge reactor was operated for wastewater carbon and nitrate removal, during 35 surveyed days.

COD contents, as well as nitrogen contents, in terms of NH$_4^+$, NO$_3^-$ and NO$_2^-$, were surveyed in the feeding effluent, reactor bulk and settler. COD and NH$_4^+$ removal percentages were determined, as well as NO$_3^-$ and NO$_2^-$ increase inside the reactor, resulting in the determination of 16 analytical parameters. Furthermore, regarding the biomass characterization, a total of 40 image analysis parameters were initially determined, from which 15 were discarded, as they presented cross-correlations over 0.9 with other image analysis parameters. The remainings were set in 4 groups, covering free filamentous bacteria contents, aggregates contents, aggregates size and aggregates morphology. Finally, and with respect to the aggregates characterization, these were divided in 3 classes (large, intermediate and small aggregates) according to their size.

A Partial Least Squares analysis was then performed to the dataset, composed of 35 observations (sampling days), in which 26 days (around 75% of the dataset) were included on the training dataset and the remaining 9 days on the validation dataset. The obtained results allowed to establish an overall reasonable prediction ability (R values above 0.75), for the NH$_4^+$ removal percentage and NO$_2^-$ increase inside the reactor. Furthermore, the parameters found to be more correlated with the NH$_4^+$ removal were identified as belonging to the aggregates size and contents, whereas for the NO$_2^-$ increase, was clearly the NO$_2^-$ contents in the feed effluent, followed by the aggregates morphology.