Bioflocculation and Activated Sludge Separation: A PLS Case Study

G. Van De Staey* G. Gins* I. Smets**

* KU Leuven Department of Chemical Engineering, Division of Bio- and Chemical Reactor Engineering and Safety, Heverlee, Belgium.
** Corresponding author, (e-mail: Ilse.Smets@cit.kuleuven.be)

Abstract: Sedimentation and filtration are the most common techniques for activated sludge separation in wastewater treatment plants. Using partial least squares (PLS), the influence of bioflocculation related variables on removal efficiency was assessed. Small particles and dissolved polysaccharides are deemed detrimental for filtration, while hydrophobic large flocs improve the filtration performance. Settling worsens when filaments are present and improves with the presence of large flocs. The potential of using PLS is demonstrated, although more measurements and samples of a wide diversity would improve the modeling performance. Such models can then pinpoint crucial measurements for bioflocculation monitoring in relation to separation performance in wastewater treatment plants.

Keywords: Wastewater treatment, Bioflocculation, Activated Sludge, Sedimentation, Filtration, Membrane Bioreactors, Fouling, Partial Least Squares

1. INTRODUCTION

With the ever increasing number of measurements in (bio)chemical process industry, the use of multivariate statistical methods has gained interest as a valid tool to reduce the amount of data and to increase its specific information content (Eriksson et al., 2013).

In wastewater treatment facilities, activated sludge flocs, which predominantly consist of bacteria, reduce dissolved organic pollutants by degrading them in order to replicate. After replication, the sludge flocs are removed from the treated water, either by sedimentation as is common practice in conventional activated sludge systems, or by filtration, as is done in membrane bioreactors (MBRs) (Judd, 2011).

The use of the partial least squares in activated sludge research has already been demonstrated several times. E.g., Mesquita et al. (2013b) applied this method as a multivariate calibration tool, enabling prediction of intracellular storage polymers using fluorescence microscopy and quantitative image analysis rather than time consuming chromatographic methods. A review on other uses of quantitative image analysis and chemometric techniques has been presented by the same research group (Mesquita et al., 2013a). Alternatively, partial least squares has been demonstrated by Van den Broeck et al. (2011) to be an adequate assessment tool to unravel multivariate correlations between activated sludge characteristics and their MBR-filterability.

In this study, the potential of using the multivariate technique of partial least squares to unravel which activated sludge characteristics are most influential for both mentioned separation techniques is assessed. This work aims at identifying practical and relatively easy-to-do measurements that can be used by a wastewater treatment plant manager to monitor the system, and alleviate problems at their causes to alleviate disturbances before they become problematic.

2. MATERIALS AND METHODS

2.1 Sample origins

A total of fourteen samples of different origin was analysed. Samples included activated sludge from four municipal wastewater, from four long-term labscale experiments and from six different industrial processes.

2.2 Sample characterisation

All samples were subjected to an extensive characterisation using typical methods in wastewater treatment research. Mixed liquor (volatile) suspended solids (ML(V)SS) were determined using standard methods using a vacuum filtration setup.

The sludge volume index (SVI) was used as a measure of settleability, with a value surpassing 150 mL/g meaning the sludge did not settle well (APHA, 1998).

The relative hydrophobicity (RH) was determined using the microbial adhesion to hydrocarbons (MATH) test, as defined by Rosenberg et al. (1980). In this test, a sludge sample is mixed with a hydrocarbon (e.g., hexane). A more hydrophobic sludge floc will adhere to this hydrocarbon, leaving a less turbulent watery phase after settling and decanting.

The surface charge of sludge (SC) was determined by colloid titration, using polybrene and polyvinyl sulphate.
potassium salt as the cationic and anionic standards (Morgan et al., 1990).

Soluble microbial products (SMP) were extracted from the sludge using the method proposed by Van Dierdonck et al. (2012). Bound extracellular polymeric substances (eEPS) were released using heat treatment (10 min, 80 °C) and extracted using the same method. Polysaccharides (PS), proteins (PN) and DNA were measured as the major components of SMP and eEPS, using the methods of Dubois et al. (1956), Lowry et al. (1951) and Tataurov et al. (2008), respectively.

Phase contrast microscopy and image analysis were applied on diluted samples (1 g/L, using supernatant) to assess morphology and particle size distribution. The image analysis parameters selected in this study are described in Table 1. Images were acquired using an Olympus IX83 fully automated inverted microscope and were analysed by a tailor made program developed with the MATLAB Image Processing Toolbox 4.2 (The Mathworks Inc.) based on the work of Jenné et al. (2007).

Table 1. Image analysis parameters.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>RelFlocNum</td>
<td>Number of sludge flocs per megapixel of image</td>
<td>Megapixel⁻¹</td>
</tr>
<tr>
<td>RelFilNum</td>
<td>Number of sludge filaments per megapixel of image</td>
<td>Megapixel⁻¹</td>
</tr>
<tr>
<td>RelFragNum</td>
<td>Number of sludge fragments (diameter &lt; 5 μm) per megapixel of image</td>
<td>Megapixel⁻¹</td>
</tr>
<tr>
<td>RelFragSurf</td>
<td>Pixel area of sludge fragments per megapixel of image</td>
<td>(-)</td>
</tr>
<tr>
<td>Deq</td>
<td>Mean equivalent floc diameter, μm</td>
<td>(-)</td>
</tr>
<tr>
<td>DeqHw</td>
<td>Mean equivalent floc diameter, weighted with floc size, μm</td>
<td>(-)</td>
</tr>
<tr>
<td>FilLength</td>
<td>Mean filament length, μm</td>
<td>(-)</td>
</tr>
<tr>
<td>FDa</td>
<td>Mean fractal dimension of flocs</td>
<td>(-)</td>
</tr>
<tr>
<td>FDaHw</td>
<td>Mean fractal dimension of flocs, weighted with floc size</td>
<td>(-)</td>
</tr>
<tr>
<td>Rndns</td>
<td>Mean roundness of flocs</td>
<td>(-)</td>
</tr>
<tr>
<td>RndnsHw</td>
<td>Mean roundness of flocs, weighted with floc size</td>
<td>(-)</td>
</tr>
</tbody>
</table>

2.3 Fouling potential assessment

The MBR fouling potential of the activated sludge samples was assessed using a 2L labscale filtration setup. Filtration was flux-driven using a standard A4 Kubota flat sheet membrane with an effective filtration surface of 0.11 m². The transmembrane pressure (TMP) was measured during 9 steps of 15 minutes filtration at constant flux and 5 minutes relaxation. This procedure was repeated for fluxes of 10 and 20 L/m²h. The average fouling rate (AFR) for a flux is measured as the average of the 9 slopes of the TMP-versus-time fitted curve during the filtration steps. The global fouling rate (GFR) for a given flux is measured as the slope of the fitted curve throughout the whole series of filtrations/relaxations. While AFR represents a measure for the reversible fouling rate, GFR accounts more for the irreversible fouling rate. Figure 1 illustrates these concepts.

Partial Least Squares is a multivariate regression modeling technique used for relating two data matrices X and Y, in this case the activated sludge characteristics and the separation performance indicators (AFR, GFR and SVI). Before modeling, all variables are transformed into Z-scores, such that each variable has an average value of zero, and a standard deviation of one.

Both X and Y matrices are transformed into latent variables that have maximal covariance. The positions of the original variables in the new latent space are called the weights. These weights determine the coefficients for the data transformation from the original to the new space. The values of the observations in the latent space are called the scores. Figure 4 illustrates the weights and scores of the first 2 PLS components of such a PLS model. This graph will be further explained in Section 3.1.

The number of PLS components used in each model is chosen by the first number of components for which the predicted residual sum of squares (PRESS) is locally minimal, which is calculated using crossvalidation using the leave-10%-out method with 1000 repeats.

Next, the variable influence of projection (VIP) is calculated for each original sludge characteristic (X-variables). The VIP of an X-variable is a weighted sum of squares of the PLS weights for that variable over all PLS components, taking thus into account the importance of each PLS component. Hence, for a given model, there will always be one VIP vector, summarizing the importance of all X-variables. Variables with a VIP >1 are deemed important.

Finally the regression coefficients β, that link the sludge characteristics to the separation performance indicators are calculated using the created PLS model. More information on PLS can be found in, e.g., Eriksson et al. (2013).
3. RESULTS

3.1 Reversible fouling

Reversible fouling, which is measured by the AFR, is characterised as membrane fouling that can be removed by relaxation or backwashing. In the next paragraph, a PLS model is explained that attempts to link the activated sludge characteristics to reversible fouling.

First Run: All data. Figure 2 shows the measured versus the predicted normalised values of the AFR at 20 L/m²h. It is clear that, while most of the data have a moderate filterability, two samples have a far worse filterability, depicted by a much larger average fouling rate. As the PLS technique is strongly influenced by such outliers, the predictability of the other observations is undermined. Therefore, it is chosen to remove the two outliers from the dataset for a second run.

Second Run: No outliers. Figure 3 depicts the measured versus predicted normalised values of the AFR at 20 L/m²h after outlier-removal, with a much better data spread. The underlying PLS model consists of 6 PLS components.

Figure 4 shows the position of the weights and scores in the first two dimensions of the latent PLS space. The positions of the average fouling rates are indicated with big (green) dots. As can be seen on the graph, the fouling rates for both fluxes are closely located to each other. Moreover, the SMP-PS and SMP-DNA measurements, as well as the relative fragment surface, are in close proximity to the AFR variables, which relates to direct correlation, while the relative hydrophobicity (RH) can be seen on opposite sides of the origin, thus probably inversely related. Since, e.g., MLSS and MLVSS are closely situated to the origin, these variables are not very influential for the PLS model.
Fig. 5. Regression coefficients $\beta$ between normalised sludge characteristics and normalised average fouling rates @10 L/m$^2$h.

The regression coefficients $\beta$ of the PLS model that link the normalised sludge characteristics to the normalised average fouling rate are shown in Figure 5 for the flux of 10 L/m$^2$h. The same trends are observed for the higher flux of 20 L/m$^2$h (data not shown). Variables are vertically listed in order of ascending VIP, with the red line denoting the critical value of VIP > 1. Variables under the red line are thus, statistically, deemed relevant. As can be seen on the graph, the SMP-polysaccharide content is, together with the surface of fragments the most influential for the average fouling rate among the samples tested.

Although the variables eEPS-PS, fragment, filament and floc numbers, and RH are significantly relevant, their regression coefficients are lower than the previously mentioned ones. Nevertheless, having hydrophobic flocs appears to be beneficial for filtration, while a large number of fragments and flocs is not, as both are indicators of deflocculation.

3.2 Irreversible fouling

The irreversible fouling is denoted by the global fouling rate (GFR) as explained in Figure 1. Using the same dataset as the second AFR run, a PLS with 4 components model was made. The normalised GFR predictions versus measurements from the crossvalidation is shown in Figure 6.

Figures 7 and 8 depict the regression coefficients $\beta$ between the normalised global fouling rates at 10 and 20 L/m$^2$h and the activated sludge characteristics. As can be seen in the figures, both the SMP polysaccharide content and the relative fragment surface remain important for the global fouling rate at both fluxes.

As the suction force decreases when lowering the flux from 20 to 10 L/m$^2$h, the surface characteristics of the flocs become more important than their sizes, probably explaining why the influence of the relative hydrophobicity is more pronounced at the lower flux.

With a higher flux, larger particles are attracted towards the membrane, which explains the increased positive influence of the weighted mean equivalent diameter ($D_{eqW}$) on filterability (although this parameter has a VIP slightly lower than 1). The seemingly contradictory behaviour of the roundness variable can be explained in a similar manner. The roundness exhibits a detrimental effect in the case of a low flux, but becomes beneficial for higher fluxes, which might be because small flocs generally are round of shape, due to their limited number of pixels, but when larger flocs are attracted, having flocs of a regularly round shape may attribute to an easier removal by relaxation or backwash.

In the same sense, having filaments present might create a secondary network of flocs and filaments which is more easily removed during flux relief, explaining why the relative filament number has a positive impact on filtration.

3.3 Settleability

Aside from filterability, the potential use of PLS models was also tested on the settleability indicator, depicted by the sludge volume index (SVI).

Figure 9 illustrates the biplot of the first two PLS components, showing variable weights and score values. As can be seen in the figure, the relative filament surface and the mean filament length are the most important variables (largest distance from the origin). Both parameters, together with the surface weighted mean equivalent floc diameter ($D_{eqW}$) seem to be closest related to the sludge
Fig. 7. Regression coefficients $\beta$ between normalised sludge characteristics and normalised global fouling rates @10 L/m$^2$h.

Fig. 8. Regression coefficients $\beta$ between normalised sludge characteristics and normalised global fouling rates @20 L/m$^2$h.

Fig. 9. Biplot of SVI PLS regression. Weights are denoted by the blue lines, scores are denoted by the small (red) dots. The position of the SVI is indicated by the (big) green dot. Volume index (depicted with the big (green) dot), while the relative hydrophobicity is situated at the other side of the origin.

Figure 10 validates these findings, although the relative hydrophobicity does have a VIP slightly smaller than one. The importance of a good balance between filaments and flocs for adequate settling is confirmed. The negative correlations between ML(V)SS and SVI is most likely caused by the occurrence of MLSS in the denominator of the SVI formula.

3.4 Evaluation of PLS

Following the three examples above, the potential use of partial least squares as a predictive modeling tool for activated sludge separation techniques can be assessed.

It is clear that although a large number of variables are measured during measurement, only a few of these variables appear to have a significant impact on the PLS model. In the bi-plot, those are the variables with high absolute weight values, situated close to, or on opposite sides of the origin, of the performance indicator of interest. These variables should be monitored more closely in the process, and by tuning process conditions, like aeration or feeding regime, steered in the direction that increases performance, as depicted in the $\beta$-plots.

However, it should be noted that many of these PLS models are very sensitive towards outliers. This has been demonstrated in Section 3.1, where the two high outliers had a strong influence on the model, reducing its effectiveness and had to be removed from the dataset. Therefore, having a good data spread, both in training as validation set, together with adequate data preparation is of high importance. To this end, a comprehensive tutorial that
Fig. 10. Regression coefficients $\beta$ between normalised sludge characteristics and normalised sludge volume index (SVI).

This handles data spread and model validation has been published by Westad and Marini (2015). Moreover, since PLS is a data driven technique, adequate physicochemical background knowledge on the process, remains of key importance e.g., in selecting the variables included in the model and removal of outliers.

4. CONCLUSION

The use of PLS as a multivariate regression tool has been demonstrated in the case of two filtration related parameters (average fouling rate and global fouling rate), expressing the fouling potential of activated sludge during constant flux filtration experiments, as well as in the case of assessing settleability.

From the results, one can conclude that the occurrence of small sludge fragments, colloidal matter and polysaccharide content in the soluble microbial products were the most detrimental for filtration among the samples tested. Filtration improved when flocs were large, hydrophobic and roundly shaped.

Abundance of filaments was confirmed to be the most prevalent cause for sedimentation problems. A good balance between flocs and filaments is proven to be vital for proper settlement.

Due to the limited number of measurements, no real validation of the proposed models could be performed yet. More data will be gathered, a proper model validation is expected.

In general, image analysis related information regarding morphology and size distribution of flocs, together with hydrophobicity were shown to be most relevant for detecting changes in bioflocculation that can affect both filtration and settling performance in wastewater treatment systems.

ACKNOWLEDGEMENTS

Glenn Van De Steay is funded by a PhD Fellowship of the Research Foundation - Flanders. This work is supported by Project OT/10/035 of the KU Leuven Research Council. The scientific responsibility is assumed by its authors.

REFERENCES


