Pilot-scale polyhydroxyalkanoate production from food waste assisted by online spectroscopy

Beatriz Vivar de Sousa
Supervisors: Doctor Nídia Dana Lourenço, Doctor Maria Teresa Cesário, MSc Fernando Ramos Silva

Abstract

Given today's plastic disposal global crisis, bio-based and biodegradable polymers may have an important role in leading the way towards a more sustainable plastics industry. Among these "green" polymers, polyhydroxyalkanoates (PHA) are aliphatic polyesters synthesized by bacteria and stored as intracellular granules. In this study, poly(3-hydroxybutyrate-co-3-hydroxyvalerate), P(3HB-co-3HV), was produced by a mixed microbial culture fed with fruit pulp waste fermentate. The accumulation reactor was monitored at-line by Raman and near-infrared (NIR) spectroscopy to develop models for predicting total suspended solids (TSS) concentration, PHA concentration and intracellular PHA content. A principal component analysis (PCA) model performed on the Raman spectral data revealed a gradual change in the shape of the signal over time, which was believed to result from shifts in the bacterial community. NIR spectra were used for the development of quantitative calibration models through partial least squares (PLS) regression. The models were subjected to cross-validation and external test set validation and proved to be suitable for predicting all three parameters under study. TSS concentration was predicted with a root mean square error of prediction (RMSEP) of 1.0g/L and a coefficient of determination (R^2) of 0.86, while PHA concentration showed a RMSEP of 0.69g/L and a R^2 of 0.89 and PHA content had an RMSEP of 14.6% and a R^2 of 0.86. Finally, the results of the TSS and PHA concentration models were combined to give an indirect PHA content prediction. This model performed exceptionally in predicting high cell density samples, with a RMSEP of 4.52% and a R^2 of 0.96.

Introduction

We are currently facing a society that is strongly dependent on plastics throughout many sectors: from packaging to agriculture, electronics, construction industry, transportation or health care. However, at a time in which a global environmental conscience is on the rise, the issues associated with the origin and fate of conventional plastics cannot possibly be ignored. On the one hand, over 99% of all commercially available plastics are produced from fossil fuels, and on the other hand, their non-biodegradability, together with inefficient collection and recycling systems, is leading to an alarming accumulation of mismanaged plastic waste. It is becoming widely accepted that a possible approach to the current plastic waste crisis might be the development and production of biorelated plastic materials, which both reduce nonrenewable energy use and present an innovative solution for the end of life issues [1, 2, 3].

Polyhydroxyalkanoates (PHAs) are a family of aliphatic polyesters produced via bacterial metabolic transformation of sugars or fatty acids, which serve as intracellular reserves for carbon and energy. To date, over 150 hydroxyalkanoate (HA) monomers and over 300 PHA-producing bacterial species have now been identified. The wide range of possible monomeric compositions, microstructure and molecular weight distribution grants PHAs a great versatility in terms of material properties [4, 5]. In fact, properties ranging from hard, brittle plastics to soft elastomers can be obtained, using conventional processing techniques [6]. This means PHAs show potential to be used as alternatives for several traditional polymers and applied to a range of different industries [7].

The most relevant polymers belonging to the PHA family are poly(3-hydroxybutyrate) (P(3HB)), poly(3-hydroxybutyrate-co-3-valerate) (P(3HB-co-3HV)) and poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) (P(3HB-co-3HHx)), all of which are already commercially available [8, 9, 10]. At an industrial scale, PHA production is conducted by bacterial pure culture fermentation using pure substrates (e.g. glucose) and sterile conditions. However, production costs remain too high for this process to replace that of petrochemical commodity polymers. Hence, there has been an increasing effort towards developing low-cost alternative processes. One strategy has focused on using mixed microbial cultures (MMCs), which are microbial populations operating in open (non-sterile) biological systems to which a selective pressure that favors PHA-accumulating organisms is applied. MMCs are particularly interesting when agro-industrial wastes with high organic fraction are used as substrates, such as cheese whey effluents, fruit and vegetable solid wastes, olive oil mill wastewater, crude glycerol and the organic fraction of municipal solid waste leachate. Apart from reducing operational costs by not requiring sterile conditions nor costly substrates, MMC-based processes offer the possibility of reducing the pollutant load of these waste streams and allow for a possible process integration in current wastewater treatment plants (WWTP) or industrial facilities.

Processes for PHA production involving MMCs are commonly operated in three-stages, comprising: (1) an acidogenic fermentation of the organic carbon substrate to produce volatile fatty acids (VFAs), which are used as precursors for PHA biosynthesis; (2) a culture selection stage, where the culture is enriched with PHA-producing organisms by applying a selective pressure method for PHA storage; and (3) a PHA accumulation stage where the selected culture is fed with an excess of VFAs in order to exploit its PHA storage capacity to its maximum [11]. This three-stage process has been demonstrated at both laboratory and pilot scale and PHA contents comparable to those of pure culture processes have been demonstrated [12, 13]. However, the low cell densities and consequent volumetric productivities attained with MMCs are the main bottleneck of the process, along with some challenges in the cost and environmental impact of the downstream stage.

Real-time monitoring of the process bioreactors is considered essential for a reliable and robust process control and, consequently, for improved efficiency, productivity and reproducibility [14]. For instance, measurements of cell growth, substrate consumption, and product concentration are traditionally obtained through off-line analytical methods like gas chromatography (GC) and high-performance liquid chromatography (HPLC). However, these are labor-intensive and time-consuming techniques, meaning the results are usually too late to be of any use for process control [15]. Spectroscopy, on the other hand, has been gaining attention as a low-cost on-line monitoring tool for bioreactors. Spectral data can be treated through chemo-

metrics methods for process modelling and multi-analyte quantification. Among the different spectroscopy methods available, vibrational spectroscopy techniques, such as Raman and near-infrared (NIR), have received the most attention.

NIR spectroscopy is based on absorption phenomena taking place between 13,000 and 4,000 cm^{-1}. This energy is high enough to induce molecular vibrational transitions corresponding to overtones and combinations of the different fundamental bands seen in the mid-infrared (MIR) region [16]. Chemical bonds between hydrogen and a heavier atom such as carbon, nitrogen or oxygen (i.e. X-H bonds), are particularly susceptible to NIR spectroscopy, and thus strong water interference is frequent in aqueous samples. NIR spectroscopy has already been proven useful for detection and quantification of PHA and other relevant elements of the production process. For instance, Cruz et al. developed an at-line monitoring strategy based on NIR technology for the production of P(3HB) by *C. necator* using waste cooking oil as a substrate [17]. Dai et al. carried out a similar experiment using MMCs and crude glycerol as a substrate and were able to follow P(3HB) content and substrate concentration in-situ [18].

In Raman spectroscopy, on the other hand, the incident light is not absorbed but rather interacts with the molecule and is scattered from it. In this case, a single frequency radiation is used and the frequency difference between the incident and the scattered light is detected. Raman spectroscopy is best for the detection of symmetric vibrations of non-polar groups, and thus its spectra are not sensitive to water. However, it provides an inherently weak signal. Most studies using Raman as a monitoring tool for PHA production have focused on intracellular polymer quantification in pure cultures [19, 20]. The technique has further been used to quantify the molar fraction of (3HV) in samples containing commercially available copolymers of P(3HB-co-HV) [21]. No studies have been developed regarding the application of Raman spectroscopy to PHA production with MMCs.

The present study aimed to develop an at-line monitoring strategy for the accumulation assays of a pilot-scale 3-stage PHA production process, using MMCs and fruit pulp waste as a substrate. Both Raman and NIR spectroscopy were investigated as potential tools for the development of quantitative calibration models to determine total suspended solids (TSS) concentration, PHA concentration and PHA content in the biomass. Spectral data was analyzed through multivariate analysis tools, such as principal component analysis (PCA) and partial least squares (PLS) regression.

Materials and Methods

Bioreactor Operation

Pilot-scale PHA production was carried out through a 3-stage process. Two different operations of the pilot plant were followed for the present thesis. The first one took place between April and July 2019 (Operation 1), when bioreactor operation was stable, and the second one in October of the same year (Operation 2), right after restarting production.

An Upflow Anaerobic Sludge Blanket (UASB) reactor with working volume of 60 L was used to carry out the acidogenic fermentation, during which apple pulp waste provided by Sumol+Compal S.A. was converted into a VFA-rich stream and served as substrate for the selection and accumulation reactors.

The culture selection was conducted in a selection batch reactor (SBR) with a working volume of 100 L, operating under a feast and famine regime. The reactor was inoculated with activated sludge from the Mutela WWTP. The carbon source was fed in the beginning of a 12 hour-cycle. Synthetic VFA supplementation was often necessary to adjust the OLR to 75 Cmmol/cycle. Nutrient supply was uncoupled from the carbon source and was performed 2 hours after the beginning of the cycle, with a typical C:N:P ratio of 100:6.5:1 [Cmmol:Nmmol:Pmmol]. Overall, a sludge retention time (SRT) of 4 days and a hydraulic retention time (HRT) of 1 day were applied to the reactor.

Accumulation assays were performed in a 60L fed-batch reactor. A biomass purge collected at the end of the famine phase of two SBR cycles was placed in the reactor and fed with the VFA-rich stream in a pulse-wise manner under nutrient limitation. Each feeding pulse was typically equivalent to that of the SBR. Carbon depletion was indirectly monitored through dissolved oxygen (DO) levels and the accumulation assays were terminated once no significant DO drop was observed after pulse feeding. Samples were collected before each accumulation pulse for analysis through Raman and NIR spectroscopy techniques. GC was used as a reference method for PHA concentration (triplicate measurements) and the corresponding TSS were further determined for each data point (duplicate measurements). A total of 16 accumulations were monitored during Operation 1 and 13 during the Operation 2. The first set was exclusively analysed through Raman, whereas the second set was investigated through NIR spectroscopy.

Analytical Procedures

Polymer Quantification GC was used to quantify the intracellular polymer and to characterize it in terms of monomer composition [22]. Dried biomass (3-7 mg) was weighed into digestion test tubes to which acidified methanol (20% H_2SO_4, 2mL) and chloroform (2 mL, containing 1 g/L of heptadecanoate (HD) as internal standard) were added, and placed in a digestor for 3.5 hours at 100 °C. After cooling, water (2 mL) was added., the mixture was vigorously shaken and then allowed to settle. The organic phase was extracted and placed in GC vials. A gas chromatograph (Varian CP-3800 Gas Chromatograph, Bruker), equipped with a Restek column (60 m, 0,53 mm internal diameter, 1 μm film thickness, Stabilwax) was used for quantification. Standards were made using an Aldrich copolymer of P(3HB-co-3HV) containing 14% (HV) and 86% (HB) (%(w/w)) with concentrations between 0 and 6.3 g/L. Squared correlation coefficients (R^2) of 0.987 and 0.994 were achieved for HB and HV calibration curves, respectively. The PHA concentration of the samples was considered as the sum of the concentrations of both monomers.

Solids Determination The determination of TSS concentration in the samples was carried out according to the standard procedure described in [150]. Samples were filtered through a pre-weighed standard glass-fiber filter, which was then placed in an oven at 105 °C overnight. Once dried, the increase in filter weight represented TSS. The sample volume was generally fixed at 5 mL. This method offers an estimation of the amount of matter present in the accumulation reactor, which roughly relates to its biomass concentration. TSS were chosen as a measure of cell density because they give a close estimation of PHA concentration from the PHA content obtained by GC, which is based on total solids (TS).

Cell staining and Visualization Nile blue staining was performed according to Bengtsson et al., 2008 [23], to visualize PHA granules and confirm that the selected culture was effectively enriched with PHA-accumulating organisms. A drop of Nile blue stain was added to about 1 mL of sample, which was then placed in a thermostatic bath at 55 °C for 15 min. After this incubation period, samples were examined under an Olympus BX51 epifluorescence microscope at 1000X.

Phylogenetic Characterization Fluorescence *in situ* Hybridization (FISH) was carried out as previously described by Amann et al. (1995) [24]. Samples were fixed with 4% paraformaldehyde (PFA), which is specific for Gram-negative bateria: 1 mL of PFA was added to 0.5 mL of sample, which was then incubated at 4 °C for 3 hours, centrifuged, washed twice with 1 mL of phosphate-buffered saline (PBS) and resuspended in 0.5 mL of absolute ethanol. Samples were applied to the wells of specific glass-slides, dried at 46 °C for 10 min, dehydrated in a grading series of ethanol solutions at 50%, 80% and 98% (3 minutes in each solution) and dried under compressed air at the end. 8 μL of an hybridization buffer (2 M NaCl, 10% sodium dodecyl sulphate (SDS), 1 M Tris-HCl, formamide and miliQ water, pH 8) were applied to each well. Then, 0.7 μL of EUBMIX (see below) and 0.7 μL of a specific probe were added to each well. Each slide was incubated at 46 °C for 1.5-3 hours and subsequently washed with a washing buffer (2 M NaCl, 10% SDS, 1 M Tris-HCl, 0.5 M ethylenediaminetetraacetic acid (EDTA) and miliQ water) pre-heated to 48 °C, then with MiliQ water at 4 °C and dried with compressed air. Vectashield mounting media was added to the dried slides. The slides were visualized using ZEISS Axiocam 506 mono fluorescence microscope, at 1000X. The fluorescein isothiocyanate (FITC)-labelled EUBmix was applied to differentiate bacteria from detritus. Specific cyanine 3 (Cy3)-labelled probes for PHA-accumulating genera were applied: *Paracoccus*, *Azoarcus*, *Thauera*, *Lampropedia*, *Amaricoccus*, *Zoogloea* and *Plasticicumulans*.

Spectral Analysis

Raman spectra were obtained with a modular spectrometer (Ocean Optics QE65 Pro), coupled with a 785 nm excitation laser (RGBLase LLc, California, USA). SpectraSuite® was the software platform used for operating the spectrometer. A 15 s integration time was applied to all samples. Each spectrum corresponded to a single scan and quadruplicates were measured every time. The detector covered a Raman shift range between -62.34 and 2677.68 cm^{-1} in a total of 1044 data points. Samples were concentrated by centrifugation and the supernatant was partially discarded before the spectra were acquired.

NIR spectra were obtained with a Matrix-F Fourier Transform-NIR Spectrometer (Bruker OPTIK GmbH, Ettlingen, Germany), coupled with a Matrix-F Fourier Transform-NIR Spectrometer (Bruker OPTIK GmbH, Ettlingen, Germany), coupled with a fiber optic probe (Falcata XP 12 NIR, Hellma® Analytics, Müllheim, Germany) with a mechanical pathlength of 1 mm (total optical path length of 2mm). OPUS (Version 8.2) spectroscopy software was used to operate the equipment. Spectra were acquired in transflectance mode with a resolution of 16 cm^{-1} and a scanner velocity of 40kHz. The detector covered between 12,000 and 4,000 cm^{-1} in a total of 999 data points per spectra. Quadruplicates were measured for every sample and each spectrum further comprised 64 scans. Air was used as background. No sample pretreatment was performed before spectral acquisition. Around 50 mL of sample were collected from the accumulation reactor and placed in a glass beaker under constant agitation, where the fiber optic probe was directly inserted.

Data Processing

Multivariate data analysis was performed using Matlab version 9.3, R2017b (MathWorks, Natick, MA, USA) and the PLS Toolbox version 8.2.1 (Eigenvector Research Inc., Wenatchee, WA, USA). The Raman data matrix included 95 spectra and 970 Raman shifts, whereas the NIR data matrix comprised 74 samples and 999 wavenumbers.

For Raman spectra, the preprocessing stage comprised spectral truncation between 200 and 2000 cm^{-1}, Savitsky-Golay smoothing, followed by standard normal variate (SNV) normalization and mean-centering. On the other hand, different preprocessing strategies were explored during the optimization of the NIR models, including first and second derivative, SNV normalization and different spectral truncations. PCA was performed on both Raman and NIR data matrices for data compression and information extraction. Of the 74 NIR samples, 5 were excluded as outliers and 57 were used in PLS regression for quantitative determination of PHA content (%), PHA concentration (g/L) and TSS concentration (g/L). The remaining 12 samples were used for test set validation, after performing a full cross-validation (venetian blinds with 8 splits and 1 sample per split) and pre-selecting the calibration model. The latter was evaluated by its squared coefficient of determination (R^2), root mean square errors (RMSE) of calibration, and cross-validation (RMSEC and RMSECV, respectively). The optimal number of LVs was identified as that leading to the minimal RMSECV and a total of 20 LVs were estimated for each calibration. The final models were chosen based on their root mean square error of prediction (RMSEP).

Results and Discussion

Overview of Process Peformance

The three-stage process was successfully applied to the pilot-plant in two operational periods, allowing for culture enrichment in PHA-accumulating organisms and production of a P(3HB-co-HV) copolymer. Table 1 summarizes the most relevant parameters of every accumulation monitored.

Regarding Operation 1, it is worth highlighting that the SBR had to be started-up again after a wash-out of the culture took place. Only two accumulations were monitored before the wash-out. Since the UASB reactor was stable at the time, all the accumulations used its VFA-rich outlet stream as feed, although supplementation was often carried out in order to increase its organic load or adjust the HB:HV precursors ratio. Overall, the percentage of HV in the copolymer varied between 17 and 26 % (% w/w) and PHA contents up to 78 % (%w/w) were achieved. High cell densities for MMC standards were also obtained, up to 22 g/L as TSS. The results in terms of PHA storage are comparable to the highest storage capacities in the literature and the cell densities are considerably higher than those reported so far [12, 13].

Operation 2 overlapped with the start-up of the entire 3-stage process and, therefore, the UASB reactor was not stable when the accumulation assays began. Hence, the first assays of this operation period used a synthetic VFA mixture as feed,

which was changed to the UASB fermentate outlet as soon as possible. Supplementation was necessary in most of the assays. However, the last two accumulations (12b and 13b in Table 1) used the VFA-rich outlet exclusively as feed. The HV content in the copolymer produced was slightly higher than that of Operation 1, with a range between 22 and 35 %. The two last accumulations were an exception, since the HV content went up to 52 %. This was probably the result of not supplementing the fermentate feed, which at the time presented a higher fraction of HV precursors than expected. Nevertheless, similar PHA contents were obtained during this second operation, with a maximum of 89 % (%w/w) reported - or 71 % if these last couple of accumulation assays were to be excluded. However, lower cell densities, of up to 11 g/L, were observed.

Table 1: Summary of the most important parameters regarding the accumulation assays carried out during Operation 1 and 2. "Day" refers to the number of days since SBR start-up. Regarding Operation 1, the SBR was first started-up on the 21st of December, 2018 (accumulations i and ii) and then again on the 19th of June, 2019 (accumulation 1a to 14a). Regarding Operation 2, the SBR was started-up on 18th of September, 2019 (accumulations 1b to 13b). In some cases, (1) two or (2) three purges of the SBR were used in the accumulation assay.

Accumulation	Day	PHA Content (%)	HB:HV Ratio (%w/w)	TSS (g/L)
i	167	72	76:24	4,43
ii	172	75	79:21	9,34
1a	13	n.a.	76:24	7,77
2a	14	n.a.	76:24	9,80
3a	16	77	78:22	12,3
4a	21	45	77:23	16,8
5a	22	53	76:24	11,7
6a	27	64	68:32	15,1
7a	28	42	74:26	14,1
8a	29	44	78:22	16,1
9a(2)	34	n.a.	79:21	12,0
10a	35	53	83:17	18,4
11a	36	56	80:20	15,8
12a	37	42	83:17	12,0
13a	40	78	79:21	16,9
14a(3)	43	69	80:20	22,4
1b	19	50	67:33	6,24
2b	20	51	65:35	6,39
3b	21	47	76:26	8,31
4b	23	55	77:23	11,1
5b	27	52	78:22	9,69
6b	28	56	74:26	6,35
7b	29	60	73:27	7,19
8b	30	62	74:26	8,79
9b	33	71	77:23	3,58
10b	34	63	77:23	3,34
11b	35	65	77:23	6,29
12b	37	75	48:52	3,63
13b	40	89	49:51	4,57

It is worth highlighting the significant drop in TSS levels observed on day 33, which was certainly due to the changes in C:N:P ratio applied to the SBR (100:6.5:1 to 100:5:1) on day 29 of Operation 2. It appears that there was some serious nitrogen limitation after a few cycles, which eventually resulted in a significant loss of biomass. Although the ratio was increased right after this observation (day 34), the culture was not able to go back to the initial TSS concentrations during the monitored operation period.

In order to ensure the presence of PHA-accumulating microrganisms in the culture and to confirm the production of intracellular PHA granules, Nile Blue staining was performed on samples collected during an accumulation assay. Figure 1 illustrates polymer storage at the end of an accumulation assay.

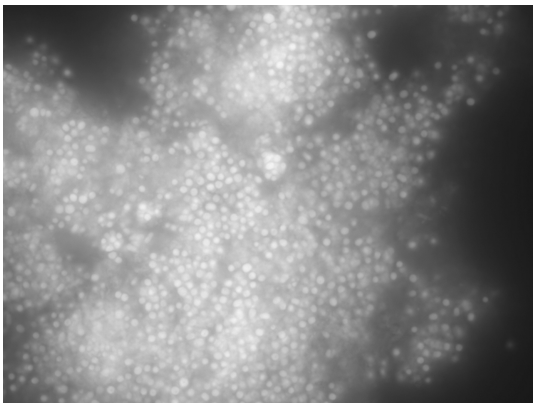

Figure 1: Fluorescence images of the biomass samples taken from after the accumulation assay on day 30 (Accumulation 8b), at 1000X. The image shows the PHA granules stained with Nile Blue, which are enhanced under fluorescent light.

An analysis of the dynamics of the culture during Operation 1 was carried out based on the SBR performance parameters, namely on the feast to famine (F/F) ratio and on the PHA production and VFA consumption rates. These parameters were obtained through weekly monitorings of the reactor and are summarized in Table 2. The significant decrease in the F/F ratio, along with the increase in both PHA production and VFA consumption rates over time indicate that the bacterial community was dynamically changing during most of the first operation period towards a more efficient PHA-producing culture.

Table 2: SBR performance parameters during Operation 1: feast/famine ratio, PHA production rate and VFA consumption rate. The PHA production rate is a global rate, which takes into account the difference in intracellular PHA between the beginning and the end of the feast phase and divides it by the length of this phase, whereas the VFA consumption rate was determined based on the slope of the VFA consumption curve during the same period. COD - chemical oxygen demand.RT

Day	F/F Ratio	PHA Production Rate (gCOD/(L.h))	VFA Consumption Rate (gCOD/(L.h))
8	0.268	1.50	0.70
16	0.064	2.74	5.18
28	0.050	4.11	7.40
35	0.037	7.78	9.30

For the second operation period, on the other hand, FISH was performed on a biomass sample collected from the SBR on day 6. The culture was screened for PHA-producing bacteria

and it was possible to conclude that the selection process was effective in only a few days. In fact, the bacterial community was dominated by the genus *Azoarcus*, with a strong presence of the genera *Thauera* and *Plasticiumulans* and some traces of *Amaricoccus* and *Paracoccus*. No further FISH analyses were carried out after this day and therefore a dynamic study of the culture was not possible.

Process Monitoring through Raman Spectroscopy

Preliminary spectral analysis A first inspection of the raw Raman spectra revealed that the shape of the spectra was not constant among the different samples, although it appeared to be coherent within a given accumulation and even within a given week. In order to establish some trend within this variation, the mean spectrum of every accumulation was estimated (see Figure 2). Firstly, there is a clear distinction between accumulations i and ii, hereby Group 1, and those who took place after reinoculating the SBR, as the former exhibit a wide band around 1,400 cm^{-1} which is non-existent in any of the others (I). Regarding accumulations 1a to 14a, there also seems to be an obvious division: accumulations 1a to 5a (Group 2) show a plateau within the first 500 cm^{-1} (II) and then a descending trend along the rest of the spectral range, whereas accumulations 6a to 14a (Group 3) display the same initial elevation but also a clear peak at around 1,200 cm^{-1} (III). It actually appears that the signal is evolving in terms of this latter peak, which is absent in the first accumulations and then starts to increase over time, eventually becoming predominant in the spectra. As was briefly mentioned before, the results gathered in Table 2 clearly indicate that the bacterial community in the SBR was evolving during Operation 1 and thus, there might have been significant changes in the bacterial community to justify the different spectral shapes detected. Unfortunately, no FISH data is available on this operation period to validate this hypothesis.

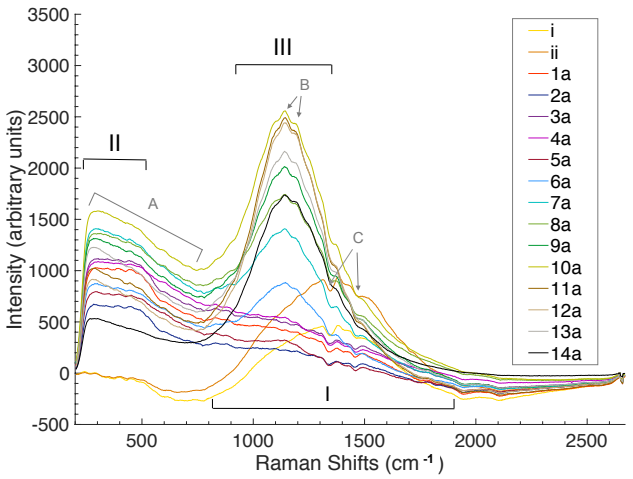

Figure 2: Mean Raman spectra of each of the 16 accumulations monitored during Operation 1. I,II and III indicate the most evident spectral bands, while A, B and C point out some subtle peaks on top of these major bands which seem transversal to the different accumulations. The region below 200 cm^{-1} was discarded.

On the other hand, it is interesting to see that despite these broad bands that dominate the spectra, a few subtle peaks can be identified on top of these bands which are transversal to the different accumulations. By inspecting Figure 2 once more, one can see that the first plateau region presents the same profile throughout the accumulations, up to a Raman Shift of around 800 cm^{-1} (A). Furthermore, two subtle peaks between 1,100 and 1,200 cm^{-1} can be detected both on top of the major band in Group 3 as on the other accumulations which lack this band (B). Similarly, two small depressions can be identified between 1,300 and 1,500 cm^{-1}, which stand out on top of the characteristic broad band of Group 1 but are also subtly present in the others (C). This might indicate that the Raman spectra might indeed be capturing some common structures beyond these most evident global changes.

Preprocessing The spectral range was limited from 200 to 2000 cm^{-1} because the regions outside this interval showed excessive noise. Nevertheless, and despite the averaging of quadruplicate measurements, a significant degree of noise could still be detected. Furthermore, the intensity of the signal was highly variable during acquistion, apparently indicating very low reprodutibility among replicate samples or even consecutive measurements of the sample. Given these considerations, a simple preprocessing strategy based on smoothing followed by SNV normalization was applied to the raw spectra, to reduce random noise and compensate for intensity shifts, respectively. All spectra were further mean-centered before moving on to the subsequent qualitative analysis.

Principal Component Analysis In order to extract further information from the spectral data, a PCA was performed. The 970 Raman Shifts detected by the spectrometer were inserted as variables, along with the 95 spectra as samples. Two principal components (PCs) were enough to capture 98.5 % of the variance in the data. The scores plot (Figure 3) clearly points out the grouping that was predicted above. This was expected, since PCA looks for the greatest amount of variation in the data, which in this case would undoubtedly be the changing shape of the spectra. Furthermore, samples belonging to Group 1 showed the highest Hotelling's T^2 values, some of which were simultaneously outside the Q residuals confidence interval (95%). It is possible to conclude that the model is particularly inadequate for this group of samples, possibly because the number of data points belonging to this group (14 out of 95) is too low to provide the model with the necessary robustness.

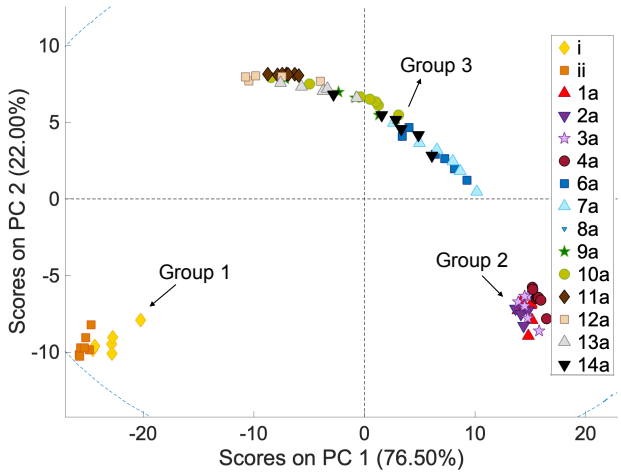

Figure 3: Scores plot of the PCA model performed on Raman spectra obtained from at-line monitoring of the accumulation bioreactor during Operation 1.

A more detailed inspection of the scores on PC1 showed a clear trend among the samples of Group 3, with increasing PC1 scores along each accumulation. This tendency might be

due to a number of factors, from increasing cell density or PHA content to the build-up of feed particles, or even some subtle changes in the overall structure and composition of the biomass. This observation, together with the above-mentioned tranversal subtle peaks, reveals significant potential that ought to be further explored through a quantitative analysis, such as PLS regression. In fact, with the right spectral truncation, the grouping effect might simply be a surmountable complication. Unfortunately, the reference values in terms of PHA content obtained throughout Operation 1 are not accurate and, therefore, a PLS model based on them would be unreliable.

Process Monitoring through NIR Spectroscopy

Preliminary Spectral Analysis By inspecting Figure 4, two saturation regions can be identified in the raw NIR spectra, between 7,000-6,400 cm^{-1} and 5,100-4,600 cm^{-1}, which correspond to the O-H bond first overtone and combinations, respectively. In fact, it is often the case with aqueous matrices that most of the relevant regions are overshadowed by water absorption bands. Furthermore, the increasing baseline shift that can be observed above 8,000 cm^{-1} is probably the result of light scattering by the particles in suspension, namely biomass [25]. In fact, suspended solids are known to induce light diffusion, decreasing the intensity of the incident radiation reaching the detector and thus raising absorption [26]. The baseline shift was very clear across the samples of one accumulation assay, reinforcing the hypothesis that it results from increasing cell density.

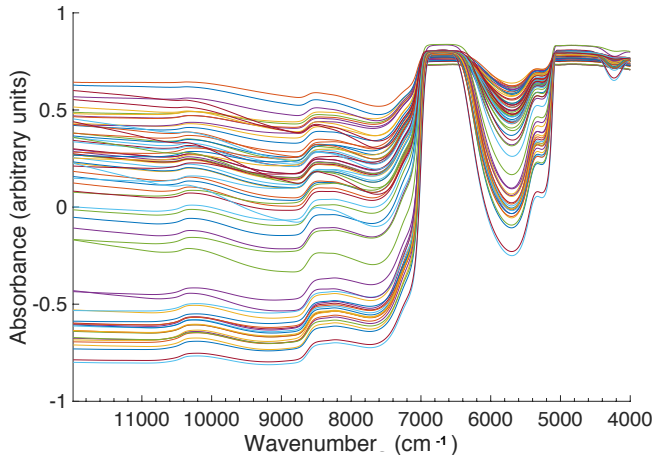

Figure 4: Raw NIR spectra obtained from at-line monitoring of the accumulation bioreactor during Operation 2 (74 spectra in the data set).

Preprocessing As a primary step, only the second derivative spectra were calculated, since this represents one of the most common preprocessing techniques found in the literature for NIR spectra of biomass samples and, in particular, of biomass containing intracellular PHA [17, 18]. Calculating second derivative spectra is especially effective in deconvoluting broad overlapping peaks and removing baseline trends. Furthermore, the differentiated spectra present a negative band with minimum at the same wavelength as the maximum on the original spectrum and were, in this case, used to carry out a detailed analysis of the NIR signal and to identify the various functional groups responsible for each of the absorption bands. Figure 5 portraits the second derivative spectra, on which each region is properly identified with the corresponding overtone or combination.

The combinations and first overtone of water (O-H stretching) are very easily identified, with a strong positive peak followed by a major negative peak around 5,100 and 7,000 cm^{-1}, respectively. This was expected, since these regions were already prominent in the raw spectra. However, it is interesting to see that a pair of very subtle positive and negative peaks seems to be present around 10,500 cm^{-1}, which might correspond to the second O-H overtone. Regarding the functional groups of interest, these mostly come down to aliphatic hydrocarbons, since the effects of any carboxylic or amine groups are completely masked by water absorption bands. Thus, only three more spectral regions were identified: the C-H stretching combinations (4,500-4,100 cm^{-1}), its first overtone (6,150-5,500 cm^{-1}) and second overtone (9,100-7,800 cm^{-1}). All three of these intervals comprehend fine peaks with variable intensity, indicating that they might contain useful chemical information regarding the intracellular PHA polymer.

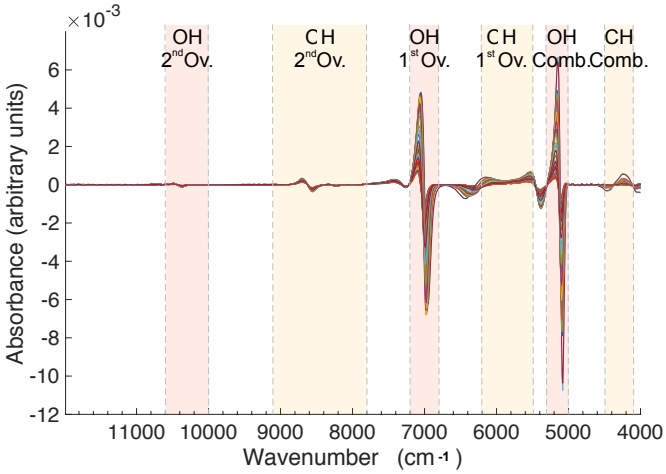

Figure 5: Second derivative NIR spectra from at-line monitoring of the accumulation bioreactor during Operation 2 and absorption bands of relevant functional groups in the NIR region. Red indicates the O-H stretching absorption bands and orange refers to the C-H stretching absorption regions. 'Ov.' stands for overtone and 'Comb.' for combination. Adapted from [27]

Principal Component Analysis A PCA was performed for better data visualization and before moving on to a more quantitative sort of analysis. Primarily, only the second derivative was used as preprocessing and the whole spectral region was investigated, since the goal was to get a broad notion of data trends, sample relationships and variable influence. Again, all samples were further mean-centered.

In this case, two PCs were able to account for 99.0% of the variance in the data. The scores plot 6 reveals that the first accumulations (1b to 8b), which correspond to the period of higher cell densities prior to changing the C:N ratio, group on the first and fourth quadrants, with positive scores on PC1. The remaining accumulations (9b to 13b), comprehending lower cell densities, score negatively on PC1 and seem to spread somewhat randomly across the second and third quadrants, mostly due to the scores on PC2.

A closer look at the scores across PC1 revealed a clear trend among each of the accumulations. In fact, it seemed to be the general case that samples increased their PC1 score along an accumulation, stressing out, once more, the hypothesis that the data is mostly being compared in terms of its biomass concentration. Outlier detection was carried out based on 95% confidence limits on the Q residuals against Hotelling T^2 statistic and 5 samples were removed from the calibration set.

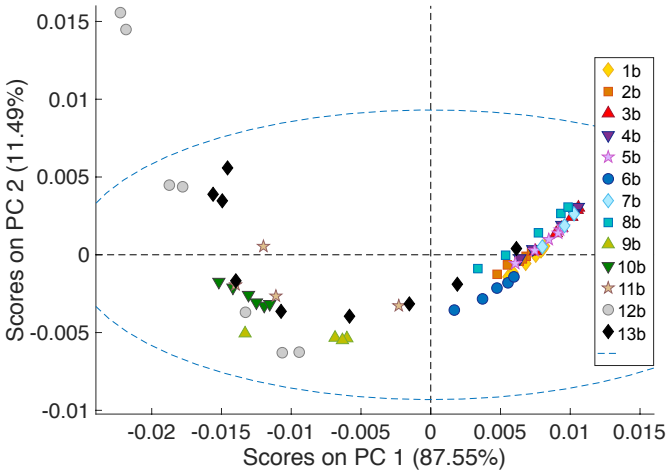

Figure 6: Scores plot of the PCA model performed on NIR spectra obtained from at-line monitoring of the accumulation bioreactor during Operation 2.

Before moving on to the subsequent quantitative model development, the PCA results were further used to decide on which accumulation assays to leave out as an external validation set. It was determined that two assays would be excluded, one corresponding to the first period with high cell densities and the other to the latter period after biomass concentration dropped. Based on the scores plot, accumulations 5b and 10b were chosen for this purpose.

Partial Least Squares Regression and Optimization Given that one of the most useful decision tools for process monitoring is the knowledge of real-time PHA content in the biomass, two different strategies were considered for the development of a quantitative PLS model. The first option was to build a model capable of directly predicting the PHA content (%), whereas the second option was based on building two separate models, one for TSS concentration (g/L) and another for PHA concentration (g/L), and combining these two to obtain PHA content values. Hence, three independent PLS regression models were developed, optimized and validated, one for each predicted variable.

With model optimization in mind, new preprocessing options were introduced and explored. Five different methods were studied: SNV, first derivative, second derivative, first derivative followed by SNV and second derivative followed by SNV. These preprocessing strategies were further combined with spectral truncation. Four different spectral truncations were specifically selected for the TSS concentration model and for the PHA content and PHA concentration models. Hence, by combining both the preprocessing options and the spectral ranges, a total of 20 models were developed and evaluated for each predicted variable.

Regarding the models for PHA prediction, both content and concentration, the four spectral ranges selected include the whole spectrum, the C-H stretching first overtone (6150-5500 cm^{-1}), the first and second overtones simultaneously (6150-5500 cm^{-1}+9100-7800 cm^{-1}) and, lastly, both overtones together with the combinations region (6150-5500 cm^{-1}+9100-7800 cm^{-1}+4500-4100 cm^{-1}).

For the TSS concentration model, the selected spectral regions were significantly different from the previous ones. Apart from the whole spectral range, the region between 9000 and 7200 cm^{-1} was one of the studied options, to evaluate how the baseline shift in that region was indicative of TSS concentration changes. Furthermore, the water absorption peaks (7200-6800 cm^{-1}+5300-5000 cm^{-1}) were also chosen for the same purpose, since they might serve as an indirect measure of biomass concentration. Lastly, the C-H stretching regions (6150-5500 cm^{-1}+9100-7800 cm^{-1}+4500-4100 cm^{-1}) were also one of the selected truncation solutions, since biomass comprises a significant fraction of organic matter.

Table 3 summarizes the most relevant details of the best five models for PHA concentration, PHA content and TSS concentration. It should be highlighted that before moving on to an external test set validation, the calibration models were validated with the PLS toolbox default cross-validation method (venetian blinds, 8 splits and 1 sample per split). In fact, the most promising models were selected based on both calibration

Table 3: Models selected for PHA concentration (g/L), PHA content (%) and TSS concentration (g/L) with corresponding parameters and results

Y	Model	Preprocessing	Spectral Region (cm^{-1})	%VarX	%VarY	#LVs	RMSEC	RMSECV	R^2 Cal	R^2 CV
PHA (g/L)	8	1st derivative	6150-5500+9100-7800+4500-4100	100	97.63	13	0.2112	0.6626	0.9763	0.7746
	11	2nd derivative	6150-5500+9100-7800	99.86	97.63	10	0.2114	0.6410	0.9763	0.7851
	12	2nd derivative	6150-5500+9100-7800+4500-4100	99.88	99.97	19	0.0227	0.4753	0.9997	0.8832
	19	2nd der+SNV	6150-5500+9100-7800	99.1	97.26	10	0.2271	0.6221	0.9726	0.7969
	20	2nd der+SNV	6150-5500+9100-7800+4500-4100	98.94	98.96	13	0.1406	0.5558	0.9896	0.8394
PHA (%)	25	1st derivative	Global	100	99.24	16	0.0186	0.0924	0.9924	0.8133
	28	1st derivative	6150-5500+9100-7800+4500-4100	99.99	96.74	12	0.0397	0.0914	0.9674	0.8296
	31	2nd derivative	6150-5500+9100-7800	99.83	94.17	8	0.0530	0.1134	0.9419	0.7419
	32	2nd derivative	6150-5500+9100-7800+4500-4100	99.86	99.78	17	0.0103	0.0915	0.9978	0.8270
	36	1st der+SNV	6150-5500+9100-7800+4500-4100	99.93	93.5	11	0.0560	0.1159	0.9350	0.7398
TSS (g/L)	48	1st derivative	7200-6800+5300-5000	100	98.52	19	0.2377	0.6688	0.9852	0.8919
	52	2nd derivative	7200-6800+5300-5000	100	98.48	19	0.2408	0.6631	0.9848	0.8979
	53	1st der+SNV	Global	99.94	96.71	9	0.3548	0.6266	0.9668	0.8974
	57	2nd der+SNV	Global	99.68	96.25	8	0.3789	0.5947	0.9625	0.9080
	60	2nd der+SNV	7200-6800+5300-5000	99.99	96.73	11	0.3539	0.6012	0.9673	0.9065

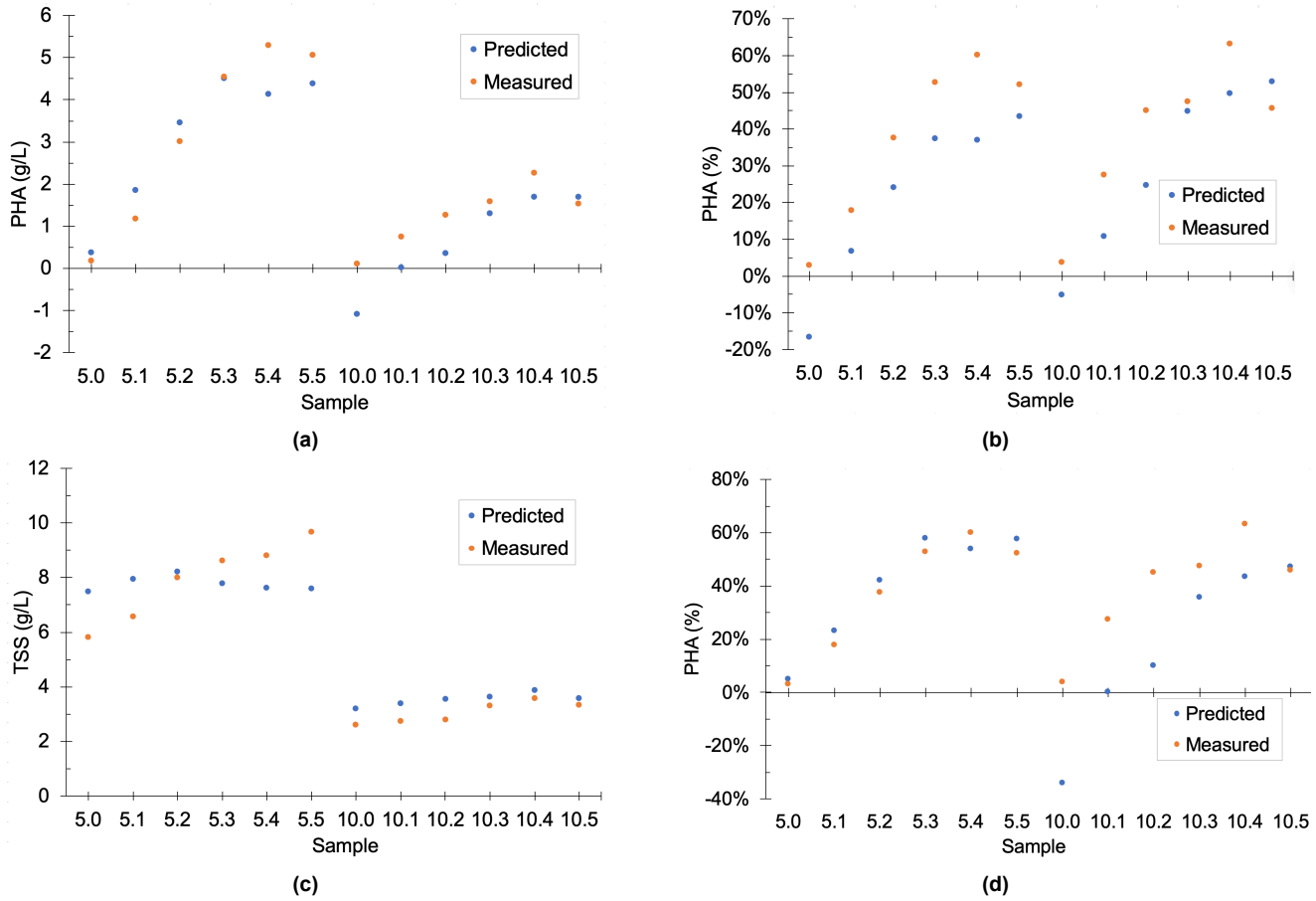

Figure 7: Prediction performance of the models selected for each predicted variable: model 8 was chosen for PHA concentration (a), model 28 for PHA content (b) and model 53 for TSS concentration (c); (d) refers to PHA content values predicted through the combination of model 8 and model 53. Sample 5.0 indicates a sample collected before the first pulse of accumulation 5b, 5.1 a sample collected before the second pulse, of the same accumulation, and so on. Sample 5.5 corresponds to a sample collected at the end of the assay. The rationale is analogous for accumulation 10b.

and cross-validation performace, namely regarding their RMSEC, RMSECV, calibration R^2 and cross-validation R^2 values.

External Test Set Validation In the final phase of model development, accumulations 5b and 10b, which had been previously excluded from the calibration block, were used as an independent set for validation. Prediction was carried out for all 5 models selected for each variable and the best model was chosen based on the minimal RMSEP. The values predicted by each model were plotted against sample number in the test set, together with the true measured values, and are displayed in Figure 7 (a) to (c). This representation allows for the evaluation of the trend in predicted values along both test accumulation assays, in comparison with the expected trend.

The final PHA concentration model (model 8) refers to a preprocessing comprising the first derivative alone, and uses all three C-H stretching absorption regions. A RMSEP of 0.693 g/L and a prediction R^2 of 0.888 were obtained. This means that the RMSE of cross-validation and prediction are comparable and the prediction R^2 is actually higher than that of cross-validation. This is probably because the desired trend can be found across the predicted values, but these values are slightly underestimated, especially for accumulation 10b. Overall, the model seems to perform better for accumulation 5b and is particularly inappropriate for very low PHA concentrations, predicting a negative concentration for the first sample in accumulation 10b. In fact, the model is likely to be poor at predicting low PHA concentration, namely those obtained after the first few accumulation pulses, given that the RMSEP is considerably high. However, it should provide a more or less significant prediction of the final accumulation pulses. It should be noted that the PHA concentration range in the calibration data set was between 0.16 g/L and 6.16 g/L, which means that sample 10b.0 (0.10 g/L) was extrapolated.

Regarding the PHA content prediction, the final model (model 28) also refers to a preprocessing strategy based on the first derivative alone, and to a spectral truncation including all three C-H stretching absorption regions. A RMSEP of 14.5% and a prediction R^2 of 0.864 were obtained. Again, the prediction R^2 is higher than that of cross-validation and the underlying reason is probably the same. In this case, the underestimation is more accentuated and negative PHA concentrations are predicted for the first samples of both accumulations. However, it should be noted that sample 5b.0 (2.95%) is slighly outside the PHA content range of the calibration set, which covers from 3.16% to 89.0%.

Likewise, the final TSS concentration model (model 53) refers to a preprocessing strategy based on the first derivative followed by SNV normalization, and made use of the whole spectral range. A RMSEP of 1.03 g/L and a prediction R^2 of 0.863 were obtained. Unlike with the previous two models, there does not seem to be any underestimation (nor overestimation) of the concentration values and, in this case, accumulation 10b seems slightly better adjusted. However, since there are no major variations in the TSS concentration values of the test set samples, it is not clear how well the model would react

to significant changes and trends in the data. It should be noted that the TSS concentration values in the calibration set ranged between 2.50 g/L and 11.1 g/L and that the reference values of the test set are all within this interval.

Lastly, a final scenario was evaluated, regarding the combination of the final models for PHA and TSS concentration. Hence, the PHA concentration values predicted by model 8 were divided by the TSS concentration values predicted by model 53 and the resulting PHA content predictions were also plotted against the test set samples, along with the true measured values (Figure 7 (d)). The final model was surprisingly accurate at predicting the PHA content in accumulation 5b, although a considerable underestimation was again noticed in accumulation 10b. It is probably the case that the model performs better at predicting the samples from high cell density accumulations because a higher number of samples belonging to this period was included in the calibration set. A RMSEP of 18.5% was determined for this combined model. However, this value is almost exclusively the result of predicting accumulation 10b. In fact, a RMSEP of 4.52% and R^2 of 0.960 are achieved if considering accumulation 5b alone.

Conclusions

Regarding the application of Raman spectroscopy, the inaccuracy of the reference analytical techniques prevented the development of a quantitative model and thus the suitability of the technique for measuring PHA could not be inferred. Nevertheless, a qualitative analysis revealed some interesting shifts in the shape of the spectral signal over the weeks. This was believed to be caused by changes in the bacterial community, which was not yet stable by the end of the operation period. Since no FISH analyses were carried out during this time, it was also not possible to corroborate this hypothesis. Future work should fall upon gathering accurate analytical information on PHA content and developing a PLS model. It might be necessary to perform a rigorous spectral truncation, given the apparent susceptibility of the equipment to changes in the culture.

On the contrary, NIR spectroscopy showed great promise for the monitoring of the PHA production process. A PCA model using the second derivative spectra revealed some sample grouping based on biomass concentration which essentially separated the data in two sets of accumulations: before and after decreasing the nitrogen levels in the SBR feed. As for the quantitative analysis, two different approaches were explored: a model for predicting PHA content directly, and the combination of two models for predicting PHA and TSS concentration, separately. Model optimization was carried out by investigating different preprocessing strategies and several spectral truncations. Apart from cross-validation, the most promising models for PHA concentration, PHA content and TSS concentration were successfully validated with an external test set. Overall, the models performed better at predicting the accumulation belonging to the high cell density group, probably because more calibration points were available regarding this period. In fact, the model combining the PHA and TSS concentration models revealed extraordinary results in the prediction of high cell density samples, with a RMSEP of 4.52% and a R^2 of 0.96. It might also be the case that the linear relationship between the NIR signal and the analyte concentration depends on the TSS concentration range and then two different models would have to be developed for higher and lower cell densities. Either way, more samples need to be collected in order to develop a more robust and reliable calibration model that can be actively used for monitoring the accumulation stage.

As far the author is aware, this is the first time that a NIR-based quantitative model has been reported for the monitoring of a pilot-scale production of P(3HB-co-3HV) polymer using MMC and a complex feedstock such as fruit pulp waste as substrate.

Ackowledgements

This document was written and made publically available as an institutional academic requirement and as a part of the evaluation of the MSc thesis in Biological Engineering of the author at Instituto Superior Técnico. The work described herein was performed at the Biochemical Engineering (BIOENG) group, in the Faculty of Sciences and Technology (FCT) of Universidade NOVA de Lisboa (UNL), between March and November, 2019. The thesis was supervised by Dr. Nídia Dana Mariano Lourenço and MSc. Fernando Ramos Silva and co-supervised at Instituto Superior Técnico by Dr. Maria Teresa Ferreira Cesário Smolders. The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 773375 and 730349.

References

[1] Ellen MacArthur Foundation. *The New Plastics Economy*. Tech. rep. Ellen MacArthur Foundation, 2016, pp. 1–120. DOI: 10.1080/13563460500344468.

[2] European Commission. *A European Strategy for Plastics*. Tech. rep. 2018, p. 24. DOI: 10.1021/acs.est.7b02368. URL: http://ec.europa.eu/environment/circular-economy/pdf/plastics-strategy-brochure.pdf%7B%5C%%7D0Ahttp://ec.europa.eu/environment/circular-economy/index%7B%5C_%7Den.htm.

[3] *New Plastics Economy Global Commitment*. Tech. rep. June. Ellen MacArthur Foundation, 2019.

[4] Tajalli Keshavarz and Ipsita Roy. "Polyhydroxyalkanoates : bioplastics with a green agenda". In: *Current Opinion in Microbiology* 13.3 (2010), pp. 321–326. ISSN: 1369-5274. DOI: 10.1016/j.mib.2010.02.006. URL: http://dx.doi.org/10.1016/j.mib.2010.02.006.

[5] M. Reis, M. Albuquerque, M. Villano, and M. Majone. "Mixed Culture Processes for Polyhydroxyalkanoate Production from Agro-Industrial Surplus/Wastes as Feedstocks". In: *Comprehensive Biotechnology, Second Edition*. Vol. 6. Elsevier Inc., Sept. 2011, pp. 669–683. ISBN: 9780080885049. DOI: 10.1016/B978-0-08-088504-9.00464-5.

[6] T A Cooper and Argo Group International. *Developments in bioplastic materials for packaging food , beverages and other fast- moving consumer goods*. Woodhead Publishing Limited, pp. 108–152. ISBN: 9780857098979. DOI: 10.1533/9780857098979.108. URL: http://dx.doi.org/10.1533/9780857098979.108.

[7] Nafisa Jabeen, Ishrat Majid, and Gulzar Ahmad Nayik. "Bioplastics and food packaging: A review". In: *Cogent Food & Agriculture* 1.1 (2015), pp. 1–6. DOI: 10.1080/23311932.2015.1117749.

[8] *Biomer*. URL: http://www.biomer.de/IndexE.html (visited on 08/29/2019).

[9] *TianAn Biopolymer - Nature's Ecofriendly Solution*. URL: http://www.tianan-enmat.com/%7B%5C#%7D (visited on 08/29/2019).

[10] *Danimer Scientific - a Biotechnology Company*. URL: https : / / danimerscientific . com (visited on 08/29/2019).

[11] Anouk F Duque, Catarina S S Oliveira, Inês T D Carmo, Ana R Gouveia, Filipa Pardelha, Ana M Ramos, and Maria A M Reis. "Response of a three-stage process for PHA production by mixed microbial cultures to feedstock shift : impact on polymer composition". In: *New Biotechnology* 31.4 (2014), pp. 276–288. DOI: 10.1016/j.nbt.2013.10.010.

[12] Jelmer Tamis, Kätlin Luˇ, Yang Jiang, Mark C M Van Loosdrecht, and Robbert Kleerebezem. "Enrichment of Plasticicumulans acidivorans at pilot-scale for PHA production on industrial wastewater". In: *Journal of Biotechnology* 192 (2014), pp. 161–169. DOI: 10.1016/j.jbiotec.2014.10.022.

[13] Simon Anterrieu, Luca Quadri, Bert Geurkink, Inez Dinkla, Simon Bengtsson, Monica Arcos-hernandez, Tomas Alexandersson, Fernando Morgan-sagastume, Anton Karlsson, Markus Hjort, Lamija Karabegovic, Per Magnusson, Peter Johansson, Magnus Christensson, and Alan Werker. "Integration of biopolymer production with process water treatment at a sugar factory". In: *New Biotechnology* 31.4 (2014), pp. 308–323. ISSN: 1871-6784. DOI: 10.1016/j.nbt.2013.11.008. URL: http://dx.doi.org/10.1016/j.nbt.2013.11.008.

[14] N D Lourenço, J A Lopes, C F Almeida, M C Sarraguça, and H M Pinheiro. "Bioreactor monitoring with spectroscopy and chemometrics : a review". In: *Analytical and Bioanalytical Chemistry* 404 (2012), pp. 1211–1237. DOI: 10.1007/s00216-012-6073-9.

[15] Martin Koller and Alejandra Rodríguez-Contreras. "Techniques for tracing PHA-producing organisms and for qualitative and quantitative analysis of intra- and extracellular PHA". In: *Engineering in Life Sciences* 15 (2015), pp. 558–581. DOI: 10.1002/elsc.201400228.

[16] Matthew Scarff, S Alison Arnold, Linda M Harvey, and Brian McNeil. "Near Infrared Spectroscopy for Bioprocess Monitoring and Control : Current Status and Future Trends". In: *Critical Reviews in Biotechnology* 26 (2006), pp. 17–39. DOI: 10.1080/07388550500513677.

[17] Madalena V Cruz, Mafalda Cruz Sarraguça, Filomena Freitas, João Almeida Lopes, and Maria A M Reis. "Online monitoring of P (3HB) produced from used cooking oil with near-infrared spectroscopy". In: *Journal of Biotechnology* 194 (2015), pp. 1–9. DOI: 10.1016/j.jbiotec.2014.11.022.

[18] Jing Dai, Erik R Coats, and Armando G Mcdonald. "Multivariate near infrared spectroscopy for predicting polyhydroxybutyrate biosynthesis by mixed microbial consortia cultured on crude glycerol". In: *Biomass and Bioenergy* 81 (2015), pp. 490–495. ISSN: 0961-9534. DOI: 10.1016/j.biombioe.2015.08.009. URL: http://dx.doi.org/10.1016/j.biombioe.2015.08.009.

[19] Joke De Gelder, Diana Willemse-erix, Maarten J Scholtes, Jorge I Sanchez, Kees Maquelin, Peter Vandenabeele, Patrick De Boever, Gerwin J Puppels, Luc Moens, and Paul De Vos. "Monitoring Poly (3-hydroxybutyrate) Production in Cupriavidus necator DSM 428 (H16) with Raman Spectroscopy". In: *Analytical Chemistry* 80.6 (2008), pp. 2155–2160. DOI: 10.1021/ac702185d.

[20] Ota Samek, Stanislav Obruca, Martin Siler, Petr Sedlacek, Pavla Benesova, Dan Kucera, Ivana Márova, Jan Jezek, Silva Bernatová, and Pavel Zemánek. "Quantitative Raman Spectroscopy Analysis of Polyhydroxyalkanoates Produced by Cupriavidus necator H16". In: *Sensors (Basel)* 16.11 (2016). DOI: 10.3390/s16111808.

[21] Celly M S Izumi and Marcia L A Temperini. "Vibrational Spectroscopy FT-Raman investigation of biodegradable polymers : Poly (3-hydroxybutyrate) and poly (3-hydroxybutyrate- co -3-hydroxyvalerate)". In: *Vibrational Spectroscopy* 54.2 (2010), pp. 127–132. ISSN: 0924-2031. DOI: 10.1016/j.vibspec.2010.07.011. URL: http://dx.doi.org/10.1016/j.vibspec.2010.07.011.

[22] Luísa S Serafim, Paulo C Lemos, Rui Oliveira, and Maria A M Reis. "Optimization of Polyhydroxybutyrate Production by Mixed Cultures Submitted to Aerobic Dynamic Feeding Conditions". In: *Biotechnology and Bioengineering* 87.2 (2004), pp. 145–160. DOI: 10.1002/bit.20085.

[23] Simon Bengtsson, Jakob Hallquist, Alan Werker, and Thomas Welander. "Acidogenic fermentation of industrial wastewaters : Effects of chemostat retention time and pH on volatile fatty acids production". In: *Biochemical Engineering Journal* 40 (2008), pp. 492–499. DOI: 10.1016/j.bej.2008.02.004.

[24] Amann RI. "In situ identification of microorganisms by whole cell hybridization with rRNA-targeted nucleic acid probes." In: *Molecular Microbial Ecology Manual*. Ed. by ADL Akkermans, JD van Elsas, and FJ de Bruijn. Dordrecht: Springer, 1995, pp. 1–15. DOI: https://doi.org/10.1007/978-94-011-0351-0_23.

[25] R. J. A. do Nascimento, G R de Macedo, E. S. dos Santos, and J A de Oliveira. "Real Time and In Sity Near-Infrared Spectroscopy (NIRS) for Quantitative Monitoring of Biomass, Glucose, Ethanol and Glycerine Concencentration in an Alcoholic Fermentation". In: *Brazilian Journal of Chemical Engineering* 34.02 (2017), pp. 459–468. DOI: dx.doi.org/10.1590/0104-6632.20170342s20150347ISSN.

[26] J P Reed, D Devlin, S R R Esteves, R Dinsdale, and A J Guwy. "Performance parameter prediction for sewage sludge digesters using reflectance FT-NIR spectroscopy". In: *Water Research* 45.8 (2011), pp. 2463–2472. ISSN: 0043-1354. DOI: 10.1016/j.watres.2011.01.027. URL: http://dx.doi.org/10.1016/j.watres.2011.01.027.

[27] Jorg-Peter Conzen. *Multivariate Calibration - a Practical Guide for Developing Methods in the Quantitative Analytical Chemistry*. 3rd Englis. Bruker Optik GmbH, 2014, pp. 40, 41.