Bioplastic degradation

Abstract

Bioplastic are polymers produced by micro-organisms. An example of a bioplastic is PHBV. PHBV is short for Poly(3-hydroxybutyrate-co -3-hydroxyvalerate). PHBV is a thermoplastic polymer, and most im portant, is renewable.

The goal of the project is to calculate the degradation of these PHBV particles. This is done in the following steps:

- Data acquired from flow cytometry.
- Statistical analysis on this data.
- Creating a machine learning algorithms to distinguish bacteria from PHBV.
- Calculating degradation over time through clustering.

Material and method

The programing language R has been used for analyzing the dataset. To be specific: R version 4.1.2 (2021-11-01) -- "Bird Hippie"

Platform: x86_64-w64-mingw32/x64 (64-bit).

The data was retrieved form a flow cytometer. Which works with lasers to scatter light of a practical and catches those scatter values.

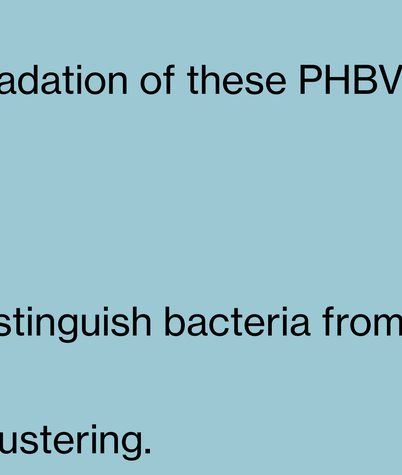

Figure 1: Flow cytometer

The data it self contains of 4 folders with different measurement techniques and different samples. Each sample also had different rates of flow cytometer cycles, which created difficulties.

The most important column variables were the Forward scatter (FSC), Sideward scatter (SSC) and the width. These variables tell something about the area and surface complexity of the particle. Below an example of how the data looks in R.

FSC.A <int>	SSC.A <int>	FL1.A <int>	FL2.A <int>	FL3.A <int>	FL4.A <int>
183714	83123	92	109	94	249
135029	50422	125	219	284	467
1152415	963584	136	51	382	354
471739	125265	4	27	87	238
432272	103531	127	113	259	463
158793	59836	185	99	363	290
107982	23163	136	132	278	232
450685	78495	54	62	117	432
111624	28196	88	54	0	331
126381	58323	187	116	319	103

Figure 2: example of the .fsc data

Conclusion

Many challenges arose form the dataset. As the data was contained in a uncommon format

and formulated strangely by the prior experiments. But through discussion with the original author of the data and data analysis a final dataset was created to be used for the remainder of the project.

Results

Freeze – thaw:

A method used in the flow cytometer data is Freeze thaw. Where bacteria of different samples where killed.

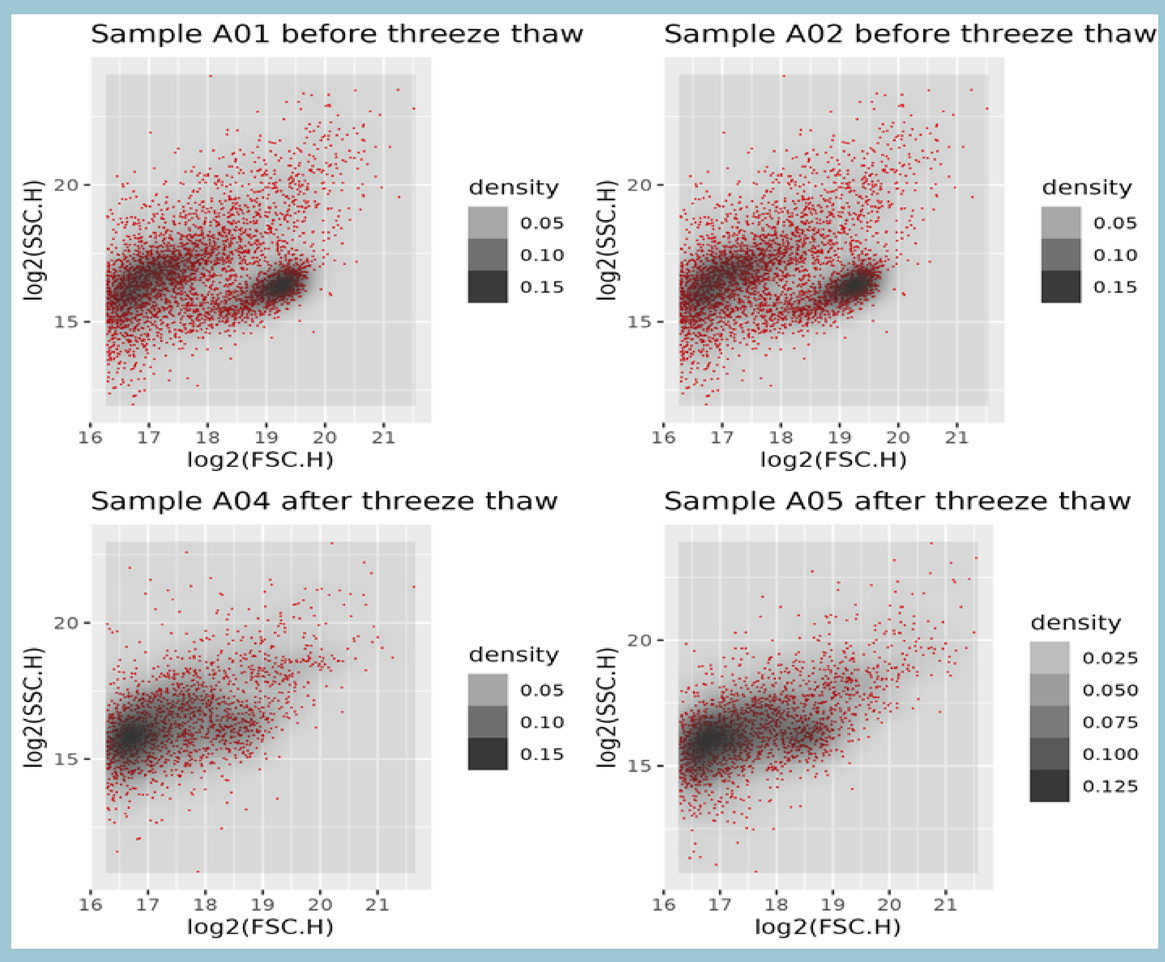

Figure 3: Freeze thaw plots

Here you can clearly see that before Freeze thaw , there were 2 clusters. And after freeze thaw only one cluster. This was very useful to identify the bacteria and plastics.

Principal Component Analysis (PCA):

PCA is a useful statical analysis technique to find new correlation between variables. PCA was done on a part of the data to sneak peek into it, to see if there are any correlations.

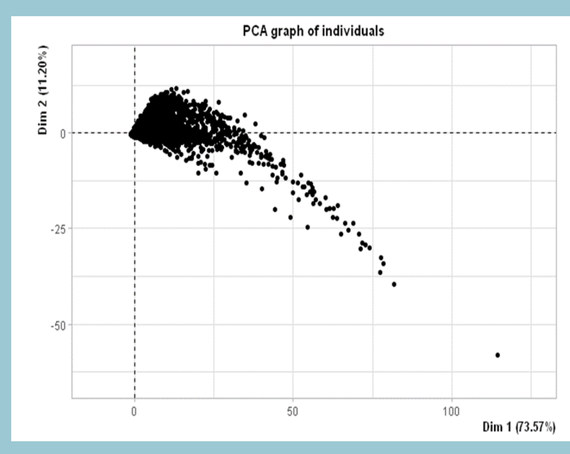

Figure 4: PC1 vs PC2

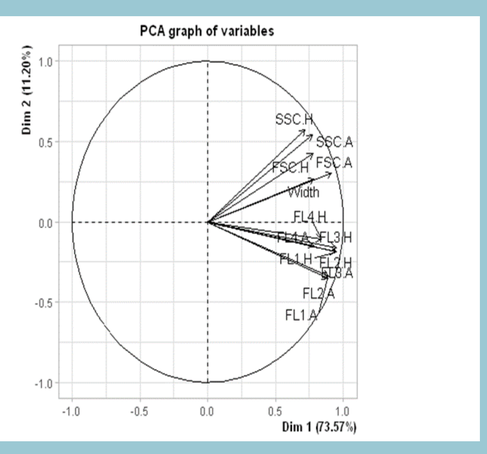

Figure 5: biplot PCA variables

When comparing PC1 against PC2 in this example, there is not a defined cluster to see, but the biplot let's see that certain variables are strong correlated to each other. Which makes this helpful.

Density

The density was also plotted of determined samples. But although there were differences, these differences where not the same across the data.

Figure 6: Density plots of all samples

Hanzehogeschool Groningen
University of Applied Sciences

Pascal Visser
p.visser@st.hanze.nl
BFV3

Rob Meulenkamp
l.r.meulenkamp@st.hanze.nl
BFV3

Rienk Heins
r.d.heins@st.hanze.nl
BFV3