

Supplementary Materials

Predicting the composition and mechanical properties of seaweed bioplastics from the scientific literature: A machine learning approach for modeling sparse data

Davor Ibarra-Pérez^{1,*}, Simón Faba⁴, Valentina Hernandez-Muñoz², Charlene Smith³, María José Galotto⁴, Alysia Garmulewicz^{5,*}

¹ Department of Mechanical Engineering, University of Santiago of Chile (USACH), Avenida Libertador Bernardo O'Higgins 3363, Santiago 9170022, Chile

² Department of Industrial Engineering, University of Santiago of Chile (USACH), Avenida Libertador Bernardo O'Higgins 3363, Santiago 9170022, Chile

³ Materiom C.I.C, E8 4QS London, UK

⁴ Packaging Innovation Center (LABEN-CHILE), Department of Food Science and Technology, Faculty of Technology, Center for the Development of Nanoscience and Nanotechnology (CEDENNA), University of Santiago de Chile (USACH), Santiago 9170201, Chile

⁵ Faculty of Economics and Management, Department of Management, University of Santiago of Chile (USACH), Avenida Libertador Bernardo O'Higgins 3363, Santiago 9170022, Chile

* Correspondence: davor.ibarra@usach.cl; alysia.garmulewicz@usach.cl

Supplementary Materials

Once the criteria for each section are validated, the information is extracted in tabular form, documenting in linear form each of the formulations and results obtained, with their respective units and identification fields. These are summarized in the following structure in Table S1 (It should be noted that all the columns of units respective to each property were suppressed to focus only on the properties obtained):

Table S1. Structure of data extraction in table form.

Category	Name Column
Material ID	Paper DOI
	Unique ID
	Material System
	Abbreviations
Ingredients	Ingredient 1
	Abbreviations 1

	Concentration 1
	Units 1
	Portion 1
	Ingredient 2
	Abbreviations 2
	Concentration 2
	Units 2
	Portion 2
	Ingredient 3
	Abbreviations 3
	Concentration 3
	Units 3
	Portion 3
	Ingredient 4
	Abbreviations 4
	Concentration 4
	Units 4
	Portion 4
	Ingredient 5
	Abbreviations 5
	Concentration 5
	Units 5
	Portion 5
Method	Process
	Tools
	Tensile strength
Mechanical	Tensile strength units
Properties	Elongation at break
	Elongation at break units

Table S2 represents Pearson correlation with mechanical properties by raw data, should be possible to observe which components promote the respective property according to the subset of the extracted data, but how sparse data the metrics is not useful for indicating the relationship of the components.

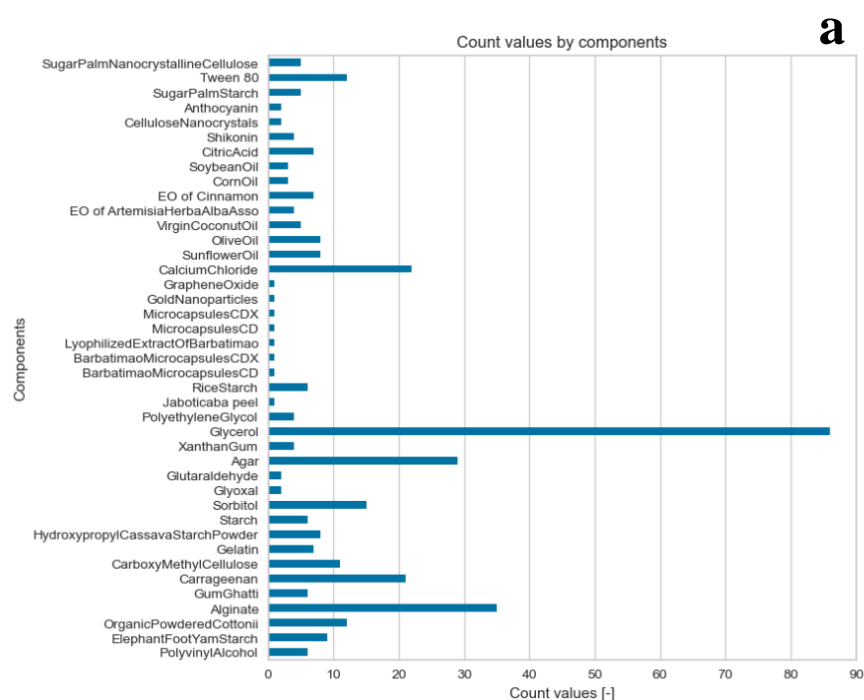
Table S2. Pearson correlation for raw data with respect to mechanical properties.

Component	Pearson correlation (Tensile strength)	Pearson correlation (Elongation at break)
Agar	0.519395	-0.169453
CarboxyMethylCellulose	0.383892	-0.238918
Starch	0.324773	-0.171730
BarbatimaoMicrocapsulesCDX	0.267977	-0.085823
Anthocyanin	0.267479	-0.083598
Shikonin	0.266662	-0.080227
LyophilizedExtractOfBarbatimao	0.260380	-0.086128

CelluloseNanocrystals	0.254150	-0.097281
BarbatimaoMicrocapsulesCD	0.244071	-0.085570
CitricAcid	0.167684	-0.169457
GumGhatti	0.154652	-0.152933
PolyethyleneGlycol	0.058711	-0.154382
MicrocapsulesCD	0.045270	-0.083846
MicrocapsulesCDX	0.005680	-0.079106
XanthanGum	-0.001796	-0.093062
ElephantFootYamStarch	-0.015602	-0.148226
Gelatin	-0.050781	0.074423
Jaboticaba peel	-0.060230	-0.019665
SoybeanOil	-0.062953	-0.035249
SunflowerOil	-0.068445	0.043610
CornOil	-0.069902	0.153843
GrapheneOxide	-0.080650	0.154502
GoldNanoparticles	-0.080919	-0.003113
OrganicPowderedCottonii	-0.089731	-0.268127
OliveOil	-0.093707	0.014411
Carrageenan	-0.096762	-0.118346
EO of Cinnamon	-0.105533	-0.048842
Alginate	-0.106741	0.163774
Sorbitol	-0.108914	-0.170174
Glyoxal	-0.116147	0.109537
Glutaraldehyde	-0.117258	0.230607
RiceStarch	-0.121901	0.316595
Glycerol	-0.140844	0.246273
SugarPalmNanocrystallineCellulose	-0.148843	0.065018
SugarPalmStarch	-0.148843	0.065018
Tween 80	-0.152621	0.121144

PolyvinylAlcohol	-0.162440	0.289430
EO of ArtemisiaHerbaAlbaAsso	-0.168590	0.046681
HydroxypropylCassavaStarchPowder	-0.169056	0.441111
VirginCoconutOil	-0.181268	0.086167
DOI_ID	-0.276187	0.030692
CalciumChloride	-0.286061	0.284818

Figure S1 shows the number of values for each of the components (A) with the standard deviation (B). It is observed that there is a high variation between the amount of one component and another, as well as the variation of the component values. This is relevant because it implies that there are characteristics that offer more information than others for the model.



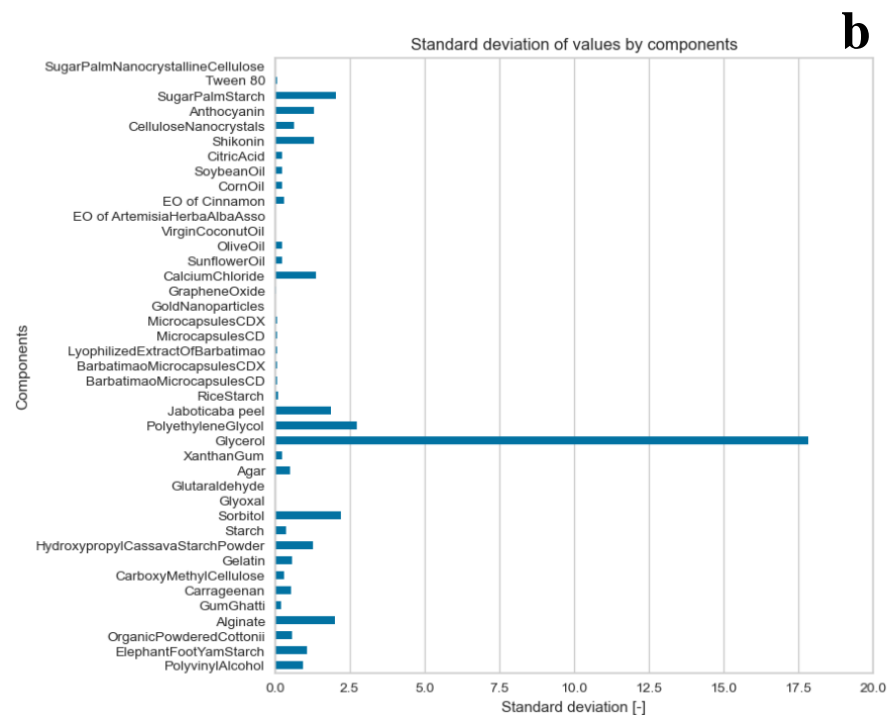


Figure S1. Features count values (A) and standard deviation (B) of each component.

Figure S2 show the jambu elbow plot, this indicates the variation of the inertia of the clustered sets as the number of clusters is varied.

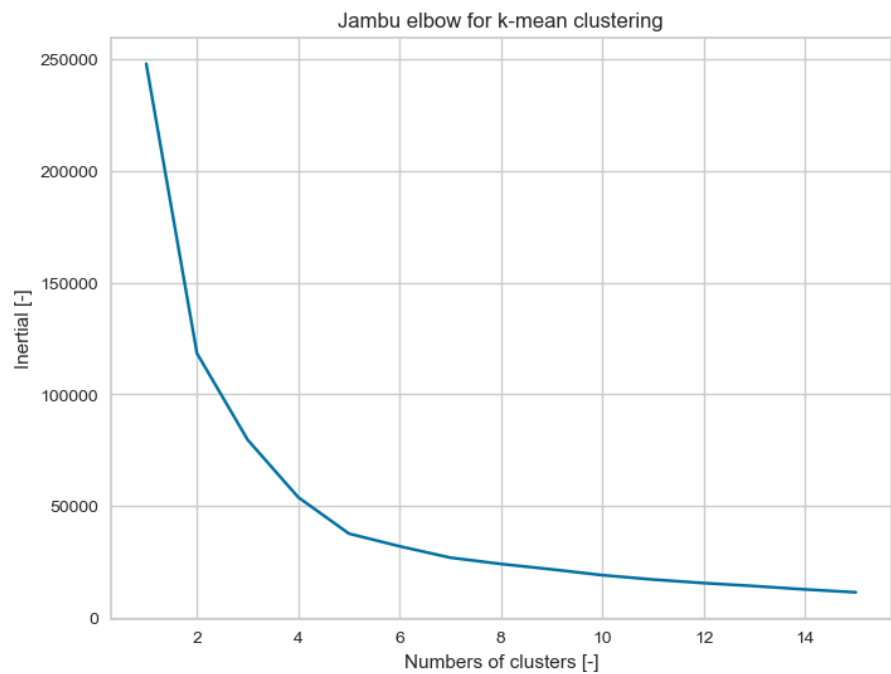


Figure S2. Jambu elbow for k-mean clustering

Table S3, S4, S5, S6 represent Pearson correlation by each cluster and mechanical properties considerate, it is possible to observe which components promote the respective property according to the subset of the extracted data.

Table S3. Pearson correlation for cluster a containing formulations with high elongation at break capacity and low tensile strength.

Component	Pearson correlation (Tensile strength)	Pearson correlation (Elongation at break)
RiceStarch	0.555631	-0.096436
Agar	0.547528	-0.134639
HydroxypropylCassavaStarchPowder	0.547528	-0.134639
Glycerol	0.507423	-0.018233
Gelatin	0.361734	-0.021574
Tween 80	0.289891	0.058294
CornOil	0.289891	0.058294
Carrageenan	0.289891	0.058294
SugarPalmStarch	0.075039	0.744508
EO of Cinnamon	0.075039	0.744508
Sorbitol	0.075039	0.744508
SugarPalmNanocrystallineCellulose	0.075039	0.744508
GrapheneOxide	-0.187558	-0.069068
Glyoxal	-0.222571	-0.057000
Glutaraldehyde	-0.321208	-0.051827
Alginate	-0.330436	-0.319252
SunflowerOil	-0.346307	0.078626
OliveOil	-0.347420	-0.388179
PolyvinylAlcohol	-0.507423	0.018233
CalciumChloride	-0.537359	-0.239779

Table S4. Pearson correlation cluster b containing formulations with low elongation at break capacity and medium tensile strength.

Component	Pearson correlation (Tensile strength)	Pearson correlation (Elongation at break)
CitricAcid	0.721592	0.013091
CalciumChloride	0.721592	0.013091
Alginate	0.719959	0.015169
GumGhatti	0.660508	0.040407
MicrocapsulesCD	0.195797	-0.178259

Glycerol	0.176437	0.112316
XanthanGum	0.060150	0.268787
MicrocapsulesCDX	0.053681	-0.137951
ElephantFootYamStarch	0.050361	0.426807
Gelatin	0.040073	-0.141593
Carrageenan	0.022489	0.055953
Agar	-0.010992	-0.143049
CarboxyMethylCellulose	-0.044388	-0.427716
SoybeanOil	-0.123781	0.256677
Tween 80	-0.123781	0.256677
EO of Cinnamon	-0.222601	0.008947
PolyethyleneGlycol	-0.278631	-0.250930
Sorbitol	-0.340857	-0.226521
SugarPalmNanocrystallineCellulose	-0.415389	0.212925
SugarPalmStarch	-0.415389	0.212925
OrganicPowderedCottonii	-0.515192	-0.391069

Table S5. Pearson correlation for cluster c containing formulations with low elongation at break capacity and high tensile strength.

Component	Pearson correlation (Tensile strength)	Pearson correlation (Elongation at break)
Alginate	0.701726	-0.628384
BarbatimaoMicrocapsulesCDX	0.436752	-0.341703
LyophilizedExtractOfBarbatimao	0.397124	-0.347939
BarbatimaoMicrocapsulesCD	0.312038	-0.336505
PolyethyleneGlycol	0.187790	-0.251271
OrganicPowderedCottonii	0.187790	-0.251271
Anthocyanin	0.027371	0.439068
CelluloseNanocrystals	-0.040511	0.020062
Shikonin	-0.138876	0.414499
Glycerol	-0.187790	0.251271
CarboxyMethylCellulose	-0.302092	0.747987

Starch	-0.308184	-0.174259
Agar	-0.730500	0.699671

Table S6. Pearson correlation for cluster d containing formulations with medium elongation at break capacity and low tensile strength.

Component	Pearson correlation (Tensile strength)	Pearson correlation (Elongation at break)
PolyvinylAlcohol	0.537084	-0.039951
Tween 80	0.293103	-0.152541
Carrageenan	0.283033	-0.280829
CornOil	0.182947	0.125143
HydroxypropylCassavaStarchPowder	0.090140	0.398835
Agar	0.090140	0.398835
RiceStarch	0.090140	0.398835
SoybeanOil	0.086147	-0.033031
SunflowerOil	0.039553	-0.185899
Jaboticaba peel	-0.007902	-0.337326
OliveOil	-0.070261	-0.062286
Glyoxal	-0.110603	-0.219330
GoldNanoparticles	-0.110603	-0.219330
VirginCoconutOil	-0.253321	0.176719
EO of ArtemisiaHerbaAlbaAsso	-0.259046	-0.070086
Alginate	-0.267281	0.204694
CalciumChloride	-0.391001	0.149942
Glycerol	-0.537084	0.039951