Biocomposites made from starch and natural fibers: study of a new processing method and modeling of the biocomposites mechanical properties

Auteur : Delahaye, Louise
Promoteur(s) : Richel, Aurore
Faculté : Gembloux Agro-Bio Tech (GxABT)
Diplôme : Master en bioingénieur : chimie et bioindustries, à finalité spécialisée
Année académique : 2019-2020
URI/URL : http://hdl.handle.net/2268.2/10523

LIÈGE université
Gembloux
Agro-Bio Tech

BIOCOMPOSITES MADE FROM STARCH AND NATURAL FIBERS: STUDY OF A NEW PROCESSING METHOD AND MODELING OF THE BIOCOMPOSITES MECHANICAL PROPERTIES

LOUISE DELAHAYE

TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE MASTER BIOINGÉNIEUR EN CHIMIE ET BIOINDUSTRIES

ANNÉE ACADÉMIQUE 2019-2020

PROMOTEUR: PROF. AURORE RICHEL

Le présent document n'engage que son auteur.

[1] Dans ce cas, l'autorité académique est représentée par le(s) promoteur(s) membre du personnel(s) enseignant de GxABT

LIÈGE université
Gembloux
Agro-Bio Tech

BIOCOMPOSITES MADE FROM STARCH AND NATURAL FIBERS: STUDY OF A NEW PROCESSING METHOD AND MODELING OF THE BIOCOMPOSITES MECHANICAL PROPERTIES

LOUISE DELAHAYE

TRAVAIL DE FIN D'ÉTUDES PRÉSENTÉ EN VUE DE L'OBTENTION DU DIPLÔME DE MASTER BIOINGÉNIEUR EN CHIMIE ET BIOINDUSTRIES

ANNÉE ACADÉMIQUE 2019-2020

PROMOTEUR: PROF. AURORE RICHEL

Plastic is wonderful because it's durable and *plastic is terrible because it's durable* (Leeson, 2016)

Acknowledgments

First, I would like to thank my promoter, Professor Aurore Richel, for allowing me to do my master thesis in the Laboratory of Biomass and Green Technologies, for her help and her guidance during this work but especially during this crisis, by making sure I could finish my work in the best conditions possible.

I would also like to thank my two supervisors, Lionel Dumoulin and Sophie Morin. This work would not have been possible without you. Thank you for your help, your countless proofreads and advice, your time, and dedication. Thank you for staying by my sides throughout this journey, physically and virtually. You helped me building my confidence as a scientist and growing up as a person, and for this I am deeply grateful.

Many thanks to the entire Laboratory of Biomass and Green Technologies for their help, technical support and kindness. I would also like to thank the people that made it possible to go back safely to the labs.

Thank you to Professor Yves Brostaux for his statistical advice.

I would like to finish by thanking my family and friends. Thank you Mamy for hosting me for more than 2 years when I started to study in Gembloux and for your support during these 5 years. Thank you, Mom and Dad, for giving me the opportunity to study in Gembloux and in Sweden and for your continuous support. Thank you, Laurence, for always having kinds words when I needed it. Thank you to my sister, Mathilde, for everything we have shared during these 5 years, the study periods, the laughers, the cries, the trips and for always believing in me. Thank you to my friends, with whom I've shared my struggles and success and so many good memories. I would not be the person I am today without you all and I am deeply grateful for you.

Lastly, thank you Ashley for proofreading my work and for your support.

Abstract

Biocomposites made from starch and natural fibers were studied in this work. Starch is abundantly produced by fractionation processes of crops and legumes while natural fibers, such as flax and hemp fibers, are largely produced in Europe.

A new processing method, through microwave-assisted plasticization, was studied in this work. Thermoplastic starch (TPS) samples were produced from pea starch, glycerol and water. Low percentages of starch (20% (w/w)) and high temperatures (190°C) gave the most optimal results in terms of homogeneous plasticization and ability to be molded. Flax, hemp and microcrystalline cellulosic natural fibers were processed with the selected TPS matrices to create biocomposites. FTIR analyses and optical microscopy highlighted the presence of matrix around the fibers, indicating a good compatibility between the initial components. No degradation of the TPS matrix or the fibers relatable to the microwave process was identified.

The fibers and starch composition of biocomposites, as well as their processing parameters, were related to their tensile test measurements through multilinear regression modeling. The database built with TPS and biocomposites data gave models with most of the variability explained when studying the Young's modulus and tensile strength ($R^2 > 0.96$). The analysis of the regression coefficients significance indicated that many variables and interactions had an impact on the mechanical properties of the final material. To help the scientific literature in their further research, a list of significant parameters was produced.

Key words: Thermoplastic starch; Biocomposites; Natural fibers; Microwave; Explanatory model; Mechanical properties

Résumé

Des biocomposites formulés à partir d'amidon et de fibres naturelles ont été étudiés dans le cadre de ce travail. L'amidon est abondamment produit par fractionnement de céréales et de légumineuses, tandis que les fibres naturelles, comme les fibres de lin et de chanvre, sont largement produites en Europe.

Une nouvelle méthode de formulation, par plastification assistée par microondes, a été étudiée dans le cadre de ce travail. Des thermoplastiques d'amidon (TPS) ont été produits à partir d'amidon de pois, de glycérol et d'eau. De faibles pourcentages d'amidon (20% de la masse totale) et des températures élevées (190°C) ont donnés les résultats les plus optimaux en ce qui concerne l'homogénéité de la plastification et la capacité à être moulé. Des fibres naturelles de lin, de chanvre et de cellulose microcristalline ont été ajoutées aux matrices TPS sélectionnées pour créer des biocomposites. Les analyses par FTIR et l'étude par microscopie optique ont mis en évidence la présence de matrice TPS autour des fibres, indiquant une bonne compatibilité entre les composants initiaux. Aucune dégradation de la matrice TPS ou des fibres, liées au procédé par microondes, n'a été identifiée.

La modélisation par régression multilinéaire a été utilisée pour étudier la relation entre la composition des fibres et de l'amidon des biocomposites, ainsi que leurs paramètres de formulation, avec leurs mesures d'essai de traction. La base de données construite avec les échantillons de TPS et de biocomposites, pour l'étude du module de Young et de la résistance à la traction, a produit des modèles dont la plus grande partie de la variabilité était expliquée ($R^2 > 0,96$). L'analyse de la significativité des coefficients de régression a indiqué que de nombreuses variables et interactions avaient une incidence sur les propriétés mécaniques du matériau final. Pour aider la littérature scientifique dans ses recherches ultérieures, une liste de paramètres significatifs a été produite.

Mots-clés : Amidon thermoplastique; Biocomposites; Fibres naturelles; Microondes; Model explicatif; Propriétés mécaniques

Table of content

List of figures

List of tables

Abbreviations

ATR: Attenuated total reflectance

BC: Biocomposite

DW: Dry weight

EaB: Elongation at break (%)

FTIR: Fourier-transform infrared

HDPE: High-Density Polyethylene

HMT: heat-moisture treatment

LDPE: Low-Density Polyethylene

Mt: Million tons

NRMSE: Normalized root mean square error

PE: Polyethylene

PET: Polyethylene terephthalate

PP: Polypropylene

PS: Polystyrene

PUR: Polyurethanes

PVC: Polyvinyl chloride

RH: Relative humidity (%)

RMSE: Root mean square error

RPM: Rotation per minute

T_g: Glass transition temperature (°C)

T_m: Melting temperature (°C)

TPS: Thermoplastic starch

TS: Tensile strength (MPa)

WVP: Water vapor permeability ($g/m^2/day$)

YM: Young's modulus (MPa)

I. State of the art

1. Plastic industry and its economy

1.1. History of the plastic industry

Plastic materials occupy an essential place in our daily life. They can be found as packaging (food wrappers and trays, drinks bottles, soap or toothpaste containers), as building and construction materials (window frames, pipes, building insulation) or as automotive, electrical and electronic parts. They equip our household with many objects (clothes, kitchen appliances and utensils, bathroom items) or for our leisure and sports equipment. Many other examples exist in many other sectors including the medical and pharmaceutical industry or agriculture (PlasticsEurope et al., 2019; Shrivastava, 2018a).

Plastic history is marked by significant discoveries and inventions that led to the plastic industry known today. For example, in 1862, Parkesine was presented to the world at the Great International Exhibition in London by Alexander Parkes. This material is considered as the precursor of celluloid, one of the first semi-synthetic plastic materials. It was produced from nitrocellulose, some solvents, and natural plasticizers such as camphor or vegetable oil. In 1907, the first fully synthetic resin was produced from phenol and formaldehyde. It was invented by Leo Baekeland who named this plastic Bakelite. This invention marked the beginning of the plastic industry (Crawford and Quinn, 2017a). However, it was only after World War II, that plastic materials started being mass produced (Shrivastava, 2018a). The abundance and low price of oil, and the improvements in manufacturing processes decreased the production costs and allowed the production of new plastic materials such as polyurethanes, polyesters, or polypropylene. Quickly, numerous new plastic products appeared on the market such as Nylon or Teflon (American Chemisty Council, n.d.).

The success of plastic materials comes from several factors: they are inexpensive, they are known to have constant and reliable quality over the years and the seasons, and by means of the specific functions they offer, they can be used in countless fields of applications (American Chemisty Council, n.d.; Shrivastava, 2018a). They were developed as functional alternatives to other materials, mainly harvested directly as raw material (e.g. horn, tortoiseshell, leather, ivory or wood) or manufactured (e.g. glass, natural fibers, or metals). Today, plastic materials have been developed and improved in such ways that they have enhanced the comfort and standard of living as they bring hygiene and safety to our world (Shrivastava, 2018a).

1.2. Plastic end-of-life

A plastic material becomes a plastic waste at the end of its life and can take different directions depending on if it is collected or not. In Europe, collected plastic is either recycled, burned for energy recovery, or disposed in a landfill (PlasticsEurope et al., 2019; Shrivastava, 2018b). However, some countries such as Belgium, has vastly banned the disposal of plastic waste in landfill.

In 2018, Europe produced 62 million tons (Mt) of plastic materials, and 29 Mt post-consumer plastic waste was collected the same year. From that collected waste, most of it was converted as energy (42.6%) or recycled (32.5%). However, 24.9% was still sent to landfill.

Plastic products have different lifespans, from less than a year to several decades. Thus, the amount of plastic produced does not always correlate with the plastic waste collected that same year. Even if the percentage of collected plastic waste sent to landfill decreases each year, it can be seen in **Figure 1** that many European countries still do not have landfill restriction implemented, and many improvements in terms of waste management (recycling and energy recovery) still have to be achieved (PlasticsEurope et al., 2019).

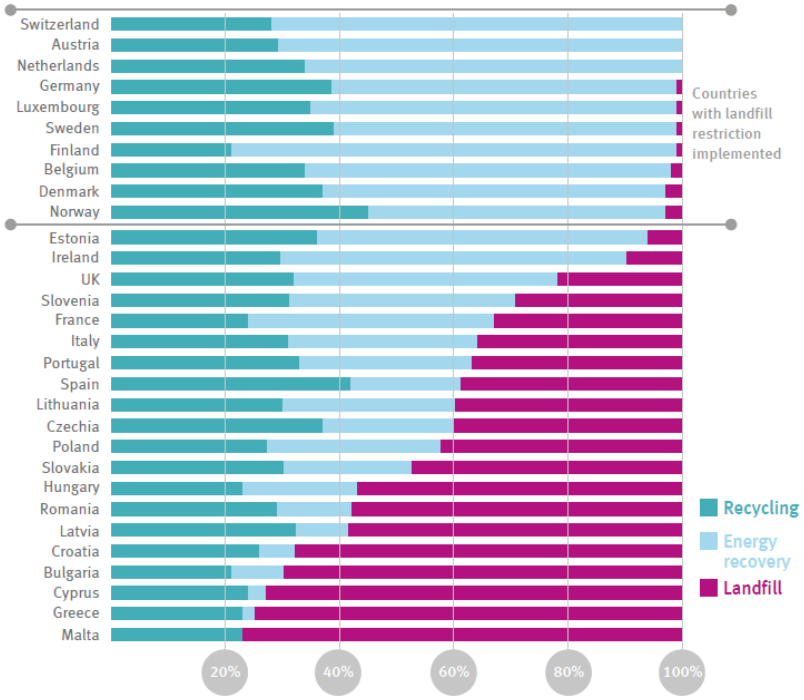


Figure 1: Post-consumer plastic waste rates of recycling, energy recovery and landfill per country in 2018 (PlasticsEurope et al., 2019)

Unfortunately, not every plastic waste is collected: they are then dumped or littered in the environment and mostly end up in the rivers and oceans (Shrivastava, 2018b). But that is not the only way plastic waste enters the aquatic environment. It arrives there by different ways: through dumping, via landfills, or by accidental spillage. Up today, approximately 10% of all the plastic ever produced has been released

in the ocean. Furthermore, a third of the plastic yearly produced is considered as single-use plastic and is dumped before the first year after manufacture (Crawford and Quinn, 2017b).

From this overview, it seems that the field of plastics faces a major challenge: maintaining the plastic materials advantages in terms of properties while reducing their drawbacks occurring at the end of their life. Improvement of the collecting and recycling system and the implement of landfill restrictions to more countries could reduce the percentage of waste sent to landfill. This part is beyond the scope of this study. Plastics with short lifespans are especially concerning; therefore, several options might be considered such as the design of performing materials either biodegradable or recyclable.

1.3. Plastic materials classification and applications

Plastics are not a single material but a group of different materials and blends that each have their own properties and characteristics which are suitable for specific applications. They can be produced from different raw materials such as fossil materials (crude oil, gas and coil) or renewable materials (cellulose, vegetable oils, starch,…) (Shrivastava, 2018b).

Two main types of plastics exist, thermoplastics and thermosets (also called thermosetting plastics). Thermoplastics are found in majority on the market. In 2008, they represented more than 70% of the plastic demand in Europe (PlasticsEurope et al., 2019). Thermoplastics are polymers that can be processed either as soft or liquid materials when heated. The state depends on whether the glass transition temperature (T_g) or the melting temperature (T_m) is respectively reached. When cooled, they solidify into a glassy or semicrystalline solid. This process is reversible, and these materials can be processed repeatedly by applying heat, meaning they can be recycled into new products. However, this reversible process can lead to degradation or affect some properties after a certain number of repetitions (Bîrcă et al., 2019; Verma and Sharma, 2017). On the contrary, thermosets do not have the ability to melt under heating. This comes from a chemical reaction during their processing that forms intermolecular cross-links, creating a complex network. This reaction is called the curing process. As a result, their mechanical properties are not dependent of usage temperature, unlike thermoplastics (Bîrcă et al., 2019). **Table 1** presents a non-exhaustive list of common plastic materials for each group with some application examples.

Table 1: Common thermoplastics and thermosets and their usual applications (PlasticsEurope et al., 2019)

Type of plastic	Examples of applications
Thermoplastics	
High-Density Polyethylene (HDPE)	Milk bottles, shampoo bottles, pipes
Low-Density Polyethylene (LDPE)	Reusable bags, food packaging films, water bottles
Polypropylene (PP)	Food packaging, wrappers, automotive parts
Polyvinyl chloride (PVC)	Window frames, pipes, floor and wall covering
Polyethylene terephthalate (PET)	Bottles of water, juices
Polystyrene (PS)	Food packaging, building insulation, eyeglasses frames
Thermosets	
Polyurethanes (PUR)	Insulation of buildings, pillows and mattresses
Epoxy resins	In special paints for ships or wind turbines, as protective coating on beds, furniture, bicycles
Silicone	Artificial corneas, bakeware, cookware, medical devices, personal care products

1.4. Plastic industry marketplace and economy

The plastic industry holds a significant place in Europe by providing direct employment to over 1.6 million people in around 60,000 companies, placing the industry as 7th in industrial value added contribution (PlasticsEurope et al., 2019).

In 2018, almost 360 Mt of plastics were produced worldwide. Europe represented 17% of the world production with 62 Mt of plastics produced. As seen in **Figure 2**, the packaging and the building and construction sectors are the most demanding in plastics (39.9 and 19.8%, respectively). As for the plastic types, PE and PP represent almost half of the plastic demand (29.7 and 19.3%, respectively).

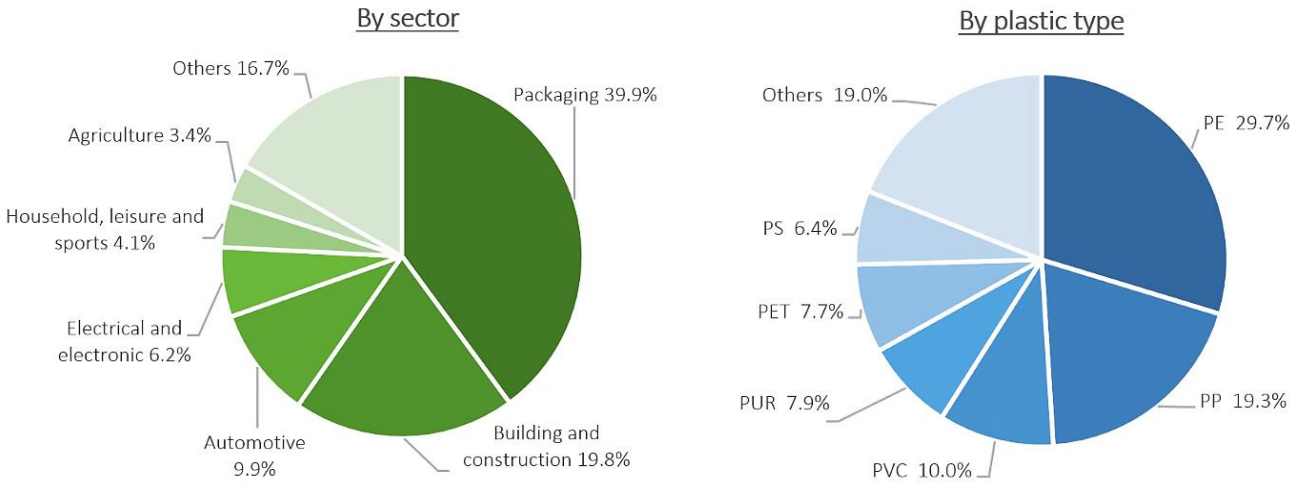


Figure 2: Distribution of the plastic demand by sector and by plastic type in Europe in 2018 (Reproduced from PlasticsEurope et al., 2019). PE: Polyethylene, PP: Polypropylene, PVC: Polyvinyl chloride, PUR: Polyurethanes, PET: Polyethylene terephthalate, PS: Polystyrene.

1.5. Bioplastics

Bioplastics are a subcategory of the previously described materials. Whether they are thermoplastics or thermosets, plastics are considered bioplastics if they are partially or totally biobased and/or biodegradable. Thus, biobased does not imply biodegradability. “Biobased” means “produced (partially) from renewable resources” while biodegradation is a biochemical process that transforms the biomass into water, carbon dioxide and compost, that depends on environmental conditions (European Bioplastics, 2019).

In 2018, bioplastics represented only one percent of the 360 Mt of plastic produced worldwide. Despite that low proportion, the bioplastics demand is increasing, and the market is in constant diversification and growth. From the bioplastic production of 2019, 55.5% were biodegradable with mainly starch blends, polylactic acid (PLA) and polybutylene adipate terephthalate (PBAT) (21.3%, 13.9% and 13.4%, respectively). The 44.5% left were biobased and non-biodegradable bioplastics with, in majority, biobased-polyethylene (PE), -polyamide (PA), -polyethylene terephthalate (PET) and -polytrimethylene terephthalate (PTT) (11.8%, 11.6%, 9.8% and 9.2%, respectively) (European Bioplastics, 2019).

Bioplastics are used in many sectors as diverse applications. Packaging and food services are major sectors for bioplastics. They are found as foamed packaging chips, cosmetics tubes and jars, shopping bags, trays, nets and films for fruits and vegetables, beverage bottles, or even catering products such as cups, plates and cutlery. Another important sector is agriculture and horticulture mainly with mulch films but also plant twine, clips, or pots. Other examples are found as pharmaceutical and medical applications, as consumer electronics, in the automotive industry, as building and construction materials, as textiles, and many more (European Bioplastics, 2019; Thielen, 2014).

Biobased bioplastics are produced from raw materials such as polysaccharides (e.g. cellulose or starch), proteins (e.g. casein), lignin, natural rubbers, sugar, and oils originated from various plants (Thielen, 2014). The main advantages of these bioplastics are the use of renewable resources that regenerate faster than fossil resources and the fact that their life cycle has the potential to be carbon neutral, as seen in **Figure 3**. This means that the carbon dioxide released during the bioplastic’s life (production, utilization, end-of-life) can be reabsorbed by the plants that will be used to create new biobased bioplastics in a human-life time lap (European Bioplastics, 2019). Another advantage of bioplastics is the increased or the creation of unique performances for some applications, such as the use of natural rubbers (extracted from rubber trees) in tires, which are not totally replaceable by synthetic rubbers.

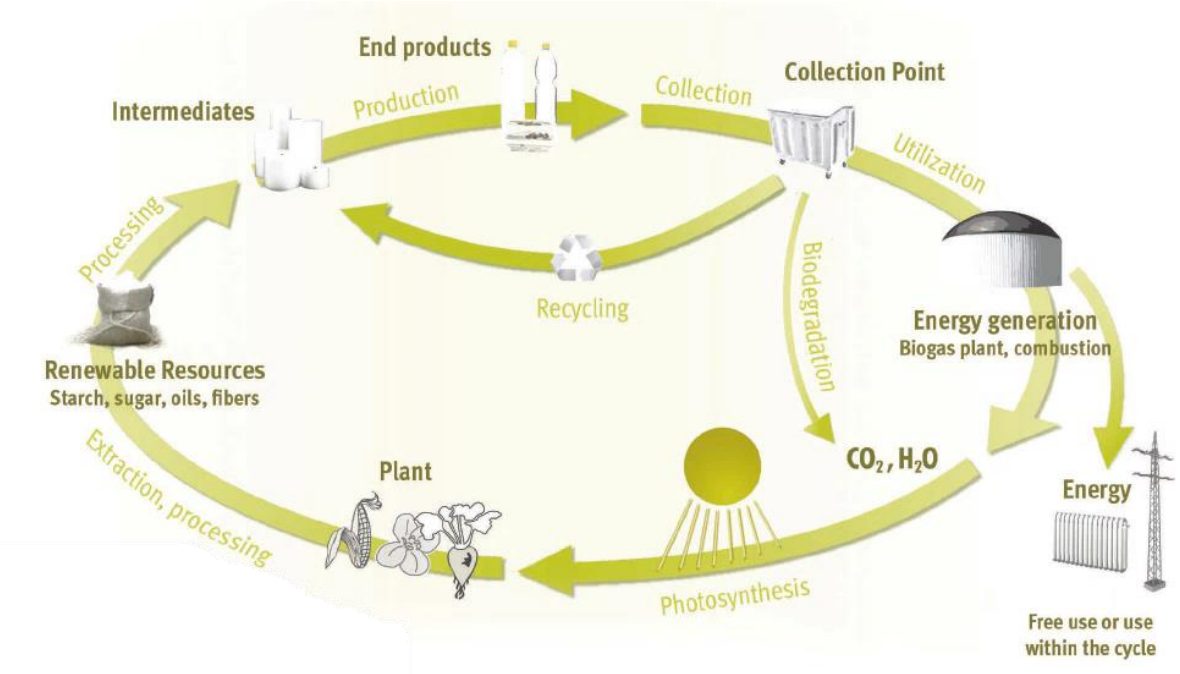


Figure 3: Life cycle of biobased bioplastics (Thielen, 2014)

Despite many advantages, biobased bioplastics remain a controversial topic because they can be made from food or feed crops. This is why many researchers and industries tend to use agriculture residues and waste or lignocellulosic materials to produce 2nd generation bioplastics (Thielen, 2014).

To know what type of plastic is more sustainable for a specific application, Life Cycle Assessment (LCA) methodology can be performed to quantify the potential environmental impacts of the product through its entire life cycle. Starting with the extraction of the raw material, then with the production and the utilization of the product to finally its disposal, this life cycle is illustrated in **Figure 3** for biobased bioplastics. This tool specifically focuses on the environment pressures in terms of climate change, human health effect, or resource depletion to name a few. Thus, it does not include economic or social impacts and needs to be combined with other tools to have a global perspective on the improvements needed (European Commission, 2019).

This work focuses on the formulation of biobased and biodegradable plastics made from starch and natural fibers. The use of starch is motivated by its occurrence as a co-product in the fractionation of legumes for proteins extraction. The hemp and flax fibers were the selected natural fibers, as they are widely produced locally (Belgium and France) and have been largely studied.

2. Thermoplastic starch (TPS) : a biobased and biodegradable plastic

2.1. Origin and structure of starch

Starch can be found in seeds, stems, leaves, roots, bulbs, tubers and fruits (Fraser-Reid et al., 2008). Some plants are known to have a high starch content (around 70% dry weight (DW)) such as wheat, corn, barley, potatoes, cassava, or sweet potatoes (ETIP Bioenergy, 2020).

Plants store starch in semicrystalline granules, which size, shape, surface, morphology, and crystallinity content differ depending on the biomass origin (species and anatomical part) (Fraser-Reid et al., 2008). The **Figure 4** presents the structure of starch granules. Each granule is constructed on different level or organization. As seen in **Figure 4 (a,b)**, granules are constructed as a succession of semicrystalline and amorphous layers of growth rings. Crystalline and amorphous lamellae are stacked within each semicrystalline layer. Crystalline lamellae are mostly composed of amylopectin chains organized in double helices and disposed in parallel. As for the amorphous lamellae, they mostly contain the amylose chains as well as the branching points of the amylopectin chains **(Figure 4 (c))** (Malumba et al., 2011).

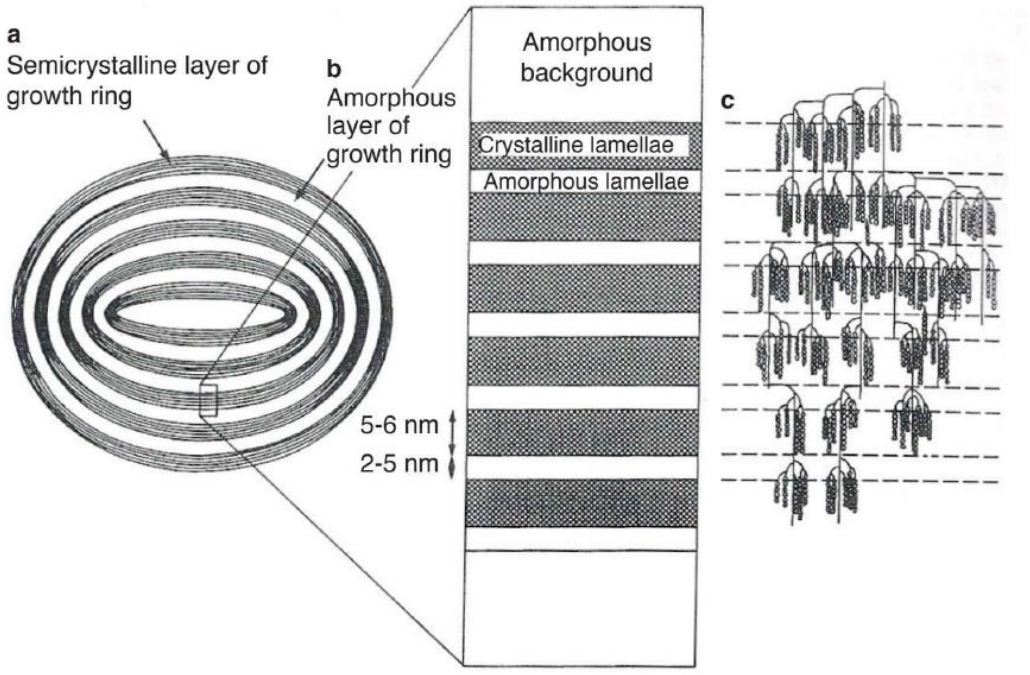


Figure 4: Schematic representation of a starch granule (Malumba et al., 2011)

Amylose and amylopectin are the two major polysaccharides found in starch. Amylose is a linear glucan with α-(1→4)-linkage and amylopectin is a branched glucan with α-(1→4)-linkage (linear backbone), with additional α-(1→6)-linkage at the branch points, as presented in **Figure 5** (Prabhu and Prashantha, 2018).

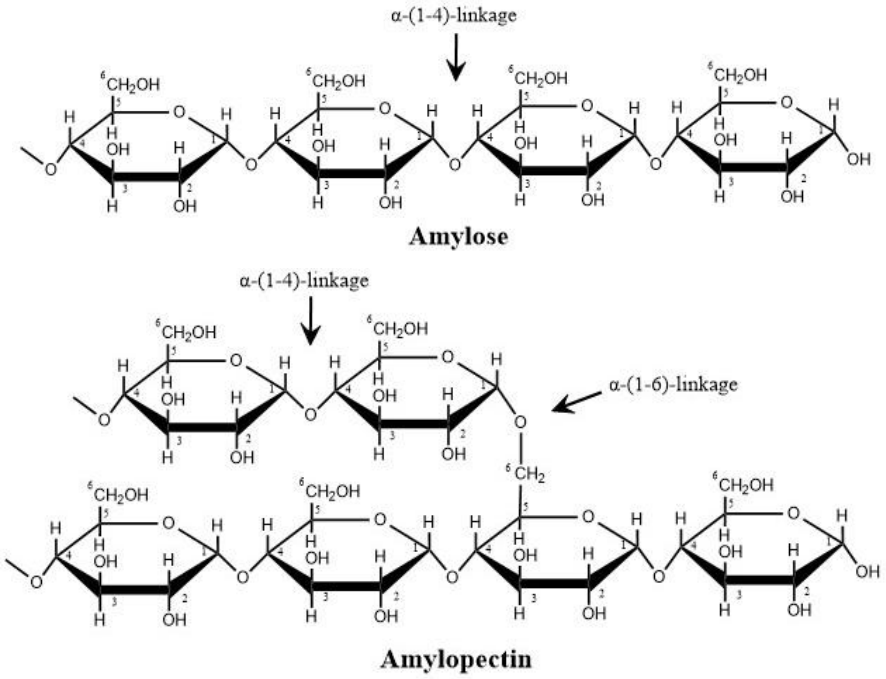


Figure 5: Chemical structure of amylose and amylopectin segments (inspired by Prabhu and Prashantha, 2018)

Starch amylose/amylopectin ratio varies depending on the botanical and anatomical origin of the considered biomass. Most species have starch composed of 20-30% amylose. However, some varieties are composed of only amylopectin (e.g. waxy rice or waxy maize) or have higher amylose content (e.g. amylomaize-V and VII) (Fraser-Reid et al., 2008). **Table 2** presents the amylose and amylopectin content, as well as starch crystallinity, for some common plants.

Table 2: Amylose and amylopectin content and crystallinity percentages of starch from different sources

Type of starch	Amylose (%)	Amylopectin (%)	Crystallinity (%)
Corn	17-25[1]	75-83[1]	43-48[2]
Wheat	20-25[1]	75-80[1]	36-39[2]
Potato	17-24[1]	76-83[1]	23-53[2]
Pea	33-88[3]	12-67[3]	17-20[2]
Faba bean	34-40[4]	60-66[4]	20-22[4]
Rice	15-35[1]	65-85[1]	38[2]

[1]: Zakaria et al., 2017; [2]: Zhang et al., 2014a; [3]: Ratnayake et al., 2002; [4]: Punia et al., 2019

Riley (2012) defines the crystallinity in polymers, as "the fraction of a polymer that consists of regions showing three-dimensional order". In starch, the crystallinity represents the percentage of crystalline regions present in the starch granules. That percentage is dependent of the botanical origin as seen in **Table 2**. Furthermore, different crystal patterns exist such as A-, B- and C-types. A-type pattern is found in cereal starches (corn, wheat, rice). B-type pattern is present in fruits, tubers and high-amylose corn. The C-type is found in legumes seed starches and is an intermediate pattern between A- and B-type (Genkina et al., 2007; Zhang et al., 2014a). As presented in **Figure 6 (A)**, the double helix of

amylopectin form a monoclinical lattice and a hexagonal lattice for the A- and B-type pattern, respectively (Genkina et al., 2007). This figure also shows that crystallites with an A-type pattern are denser and not as hydrated as a B-type pattern. X-ray diffraction is commonly used to identify the different crystal patterns as they each possess specific diffraction patterns, as observed in **Figure 6 (B)** (Carvalho, 2013). The crystallinity percentage can also be calculated from the diffractogram (Zhang et al., 2008).

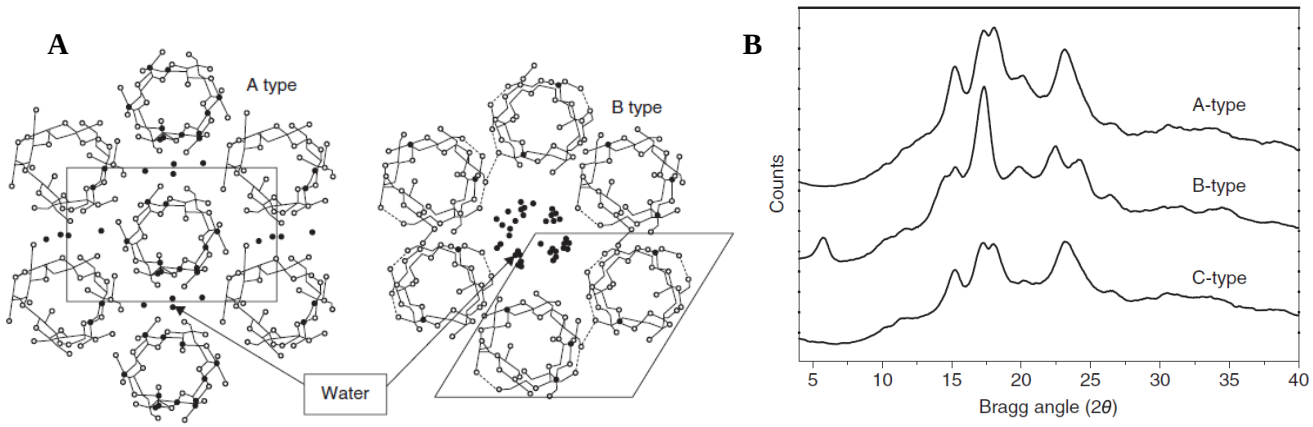


Figure 6 : (A) Plane projection of starch double helices and water molecules arrangements of starch A- and B-type crystalline structures. (Zhang et al., 2008) (B) X-ray diffraction patterns of A-,B- and C-type crystallites (Carvalho, 2013)

2.2. Native and modified starches

Commercial starch is classified in two classes, native and modified starch. Native starch is the original form of starch obtained after its extraction, while modified starch is obtained after treatment of native starch to improve its properties. Modified starch can be obtained through chemical, thermophysical, or enzymatic treatments. Examples of some common treatments and applications are presented below.

Chemical treatments

The three main chemical modifications are esterification, etherification or oxidation. Esterification for example, can result in a reduction of the sensitivity of starch to water by substituting OH groups by hydrophobic groups, which is interesting for dry food applications and also for thermoplastic starch formulation (Nafchi et al., 2013; Shrestha and Halley, 2014). Starch oxidation can reduce retrogradation, bring stabilization and low viscosity of the cooked starch pastes which can be used in the gum confection or in battered meat and fish preparation (Shrestha and Halley, 2014). Other effects are possible such as the cross-linking of starch chains, the addition of positive or negative charges, or the increase of hydrophilic behavior (Fraser-Reid et al., 2008).

Thermophysical treatments

Annealing and heat-moisture treatment (HMT) are common physical treatments. Annealing consists of heating starch below its melting temperature (T_m) but close to its glass-transition temperature (T_g). This allows a reorganization of starch molecules that will realign the starch chains and increase the interactions between them (Ratnayake et al., 2002). In result, the starch crystallization is improved (Shrestha and Halley, 2014). HMT is a physical treatment under specific temperature and moisture conditions that will impact the physico-chemical, rheological, and retrogradation properties of starch (Ratnayake et al., 2002). For example, it can decrease the process time by increasing the heat penetration, and the resulting starch can be used in sterilized soups or sauces (Shrestha and Halley, 2014).

All these modifications can be applied alone or combined and will give the starch product distinct properties (Shrestha and Halley, 2014).

2.3. Thermoplastic starch (TPS)

As the melting temperature of native starch is above its decomposition temperature, it does not have thermoplastic properties. To gelatinize and form a thermoplastic, native starch needs to undergo thermal processes and shear stress in presence of plasticizers. Under these specific conditions, thermoplastic starch (TPS) can be formulated. (Janssen and Moscicki, 2010; Zhang et al., 2014b).

TPS has been widely studied because starch is biodegradable, renewable annually, available worldwide and at relatively low cost (Janssen and Moscicki, 2010; Prabhu and Prashantha, 2018; Zhang et al., 2014a). Starch can be found as a co-product from the fractionation process of some plants, such as pea or faba bean, where the main goal is to extract protein (Ma et al., 2008). The starch fraction can thus be valorized and transformed into TPS materials or sold to TPS manufacturers to increase the process efficiency.

Despite many advantages, TPS materials made from native starch often tend to show little resistance to moisture and have poor mechanical properties (Zhang et al., 2014b). They also tend to recrystallize during storage, which alters the mechanical properties, thus the quality, of these materials (Leroy et al., 2012; Prabhu and Prashantha, 2018). In the scientific literature, these TPS properties are studied with measurements such as the tensile strength (TS), the elongation at break (EaB), the Young's modulus (YM), or the water vapor permeability (WVP) (Prabhu and Prashantha, 2018; Zhang et al., 2014b).

2.3.1. Plasticizer and plasticization

When a plasticizer is added to starch, intra- and inter-molecular hydrogen bonds between starch chains are substituted by starch-plasticizer interactions (Altayan et al., 2017). Because plasticizers molecules are “smaller”, they move more easily in the mixture than starch molecules and can modify the starch crystalline network without breaking it, by incorporating themselves between the starch chains (Nafchi et al., 2013). This phenomenon is illustrated in **Figure 6 (A)** where the water molecules are present inside the starch double helixes. With the addition of plasticizers, the structure is softened, and the macromolecular chains have a higher mobility. As a result, the glass transition temperature (T_g) of starch is lowered, and its melting temperature (T_m) decreases below its decomposition temperature. This whole process is called plasticization (Zhang et al., 2014b). This plasticization process results in the destructuration of the crystalline structure of starch and forms an amorphous TPS (Prabhu and Prashantha, 2018). One of the drawbacks of TPS is the recrystallization that can occur during storage by the formation of hydrogen bonds between the starch chains by expulsing plasticizers. The change in the mechanical properties of TPS caused by this process is called retrogradation or “aging” (Prabhu and Prashantha, 2018).

Common plasticizers are water, glycerol, sorbitol, glucose, sucrose, fructose, glycols, urea, amides, and amino acids. Combinations of these are also used (Zhang et al., 2014b). The nature and content of plasticizer will impact the final properties of the TPS, rendering it more suitable for particular applications. For example, Nafchi et al. (2013) explain that the permeability to gases (O_2, CO_2, water vapor) increases when the plasticizer concentration is increased.

2.3.2. Manufacturing processes

In the context of laboratory research, thermoplastic starch is generally obtained by casting solutions. An aqueous suspension of starch mixed with plasticizers is heated to allow the gelatinization of starch. The solution is then casted on a plate to cool down and dry before peeling the TPS film (Zhang et al., 2014b, 2014a). The extrusion process is commonly used to manufacture TPS on an industrial scale (Prabhu and Prashantha, 2018). A typical extruder is presented in **Figure 7.** The TPS mixture (starch and plasticizers) is fed from the hopper into a heated barrel. The materials are mixed, transported by the screw and heated at the same time. When the mixture arrives at the end of the barrel in a melted state, it passes through a die with a specific shape and size. Usually the first extrusion process will produce pellets. However, TPS extrudates can also be blown into films or molded into shapes before solidification. Pellets can be extruded a second time to add additives and change the properties of TPS (Zhang et al., 2014a). Other plasticization methods exist, such as internal mixing or compression molding (Prabhu and Prashantha, 2018). The choice of the method to use will depend on the TPS composition and the targeted application.

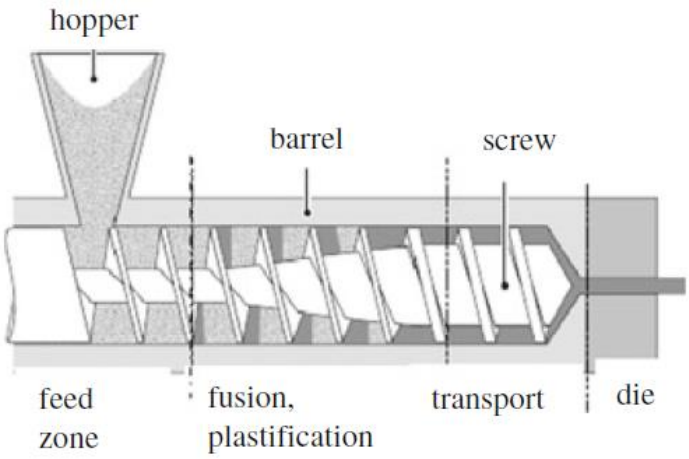


Figure 7: Scheme of a typical extruder (Ponomarev et al., 2012)

Other processing methods can be considered such as heating in a microwave reactor (also called microwave). This device generates and disperses high frequency electromagnetic waves (usually between 0.9 and 2.5 GHz) called microwave frequencies. During a microwave treatment, the water and any other polar molecules that are present in the matter align their dipoles with the alternating field coming from the microwaves. This phenomenon makes the molecules rotate rapidly and the energy created is dissipated as heat (Menendez et al., 2010). Each material possesses a dielectric constant (also called relative permittivity) that expresses how easily this material can align their dipole when it is subjected to an electric field. In result, the higher this value, the faster it will heat in the microwave. As examples, the dielectric permittivity of water and glycerol at 20°C are 80.1 and 46.5, respectively (Engineering ToolBox, 2008). Because the microwave reactor heats the matter directly in the center and performs rapid heating, it is used in many industries (Menendez et al., 2010) and could be a good alternative for the manufacturing of TPS. Different process parameters can be varied such as the temperature, the time of the heating period, the time of treatment, or the heating power.

2.3.3. Improvements of TPS properties

As stated before, TPS materials have limited applications due to low mechanical properties and tendency to absorb water. To address these drawbacks, studies have focused on the composition of the TPS mixture to improve these properties.

Starch source

The first criteria that impacts TPS is the type of starch. Because starches from different plants have different amylose/amylopectin ratio, they will form TPS with different properties (López and García, 2012; Zhang et al., 2014b). For example, López and García (2012) compared TPS films made from starch with different amylose content and found out that the ones with a higher amylose content (corn starch, 23.9% amylose) were stronger, less flexible and more resistant to water while the ones with a

lower amylose content (ahipa starch, 11.6% amylose) were more flexible and showed higher permeability to water.

Plasticizer

As stated before, the properties of TPS materials are affected by the plasticizer(s) used. Its nature and its proportion in the mixture can impact the glass-transition temperature, the water absorption, the physical properties (strength, flexibility, extensibility, permeability) and the homogeneity of the resulting TPS (Zhang et al., 2014b; Zhang and Han, 2006). It is thus important to choose the plasticizer(s) depending on the properties wanted.

Starch modification

Starch can be modified before the production of TPS. As explained before, these modifications will allow the substitution of (hydrophilic)-OH groups for selected groups (hydrophobic, hydrophilic) for specific reactiveness and properties. For example, for TPS materials, it can result in an improvement of the mechanical and barrier properties and a reduction of sensitivity to liquid water (Nafchi et al., 2013). It can also improve the compatibility between starch and polymers in blends thanks to the groups substitution (Zhang et al., 2014b). In this study, only native starch will be used.

Starch blends

Starch can be blended with other polymers, natural or synthetic, biodegradable or not (Janssen and Moscicki, 2010; Zhang et al., 2014b). Some examples include polyethylene (PE), polyvinyl alcohol (PVA) or polylactic acid (PLA). Such blends have been proven to improve the mechanical properties of TPS but still present some issues regarding the compatibility between starch (hydrophilic) and some hydrophobic polymer matrices. This leads to blends with small amount of starch ($\leq$ 40%) in the mix to prevent starch agglomeration. To address this problem, chemical compatibilizers can be added to improve the dispersion of starch in the matrix (Zhang et al., 2014b). For example, Sabetzadeh et al., 2012 showed that the use of polyethylene-grafted maleic anhydride (PE-g-MA) could improve the miscibility, and thus the mechanical properties, between PE and starch.

Other materials, called fillers, can be added to the TPS mixture such as nanoclay**.** Those fillers have demonstrated the enhancement of mechanical properties, thermal stability and water resistance of TPS. When using nanoclay, the results come from the numerous interactions created between nanoclay and starch thanks to the large interface area between them (Zhang et al., 2014b).

Fibers can also be added to TPS and used as a reinforcement to improve the properties of TPS (Zhang et al., 2014b). In a goal of developing biobased and/or biodegradable materials, such option is favored. It will be further discussed in the following section.

2.4. Fibers as a reinforcement in TPS matrices

TPS reinforced with fibers form biocomposites. In fact, composites are multicomponent materials where at least one of the components is a continuous phase (the TPS matrix in this case) (Work et al., 2004). As for biocomposites, they are defined by Jawaid et al., (2017) as "composite materials in which at least one of the constituents is derived from natural resources". If the polymer matrix, as well as the reinforcement material, are from natural resources, they can be called "green biocomposites" (Jawaid et al., 2017).

Fibers, natural or not, are used as reinforcement in composites. Thanks to their high mechanical resistance, they have the ability to transfer that resistance to the matrix they are embedded in through the creation of strong bonds. This results in the improvement of the mechanical properties of the final composite. Thermal, gas properties, and water resistance can also be improved (Zhang et al., 2014b). Fibers reinforced composites are mainly used in applications where the tensile strength has to be high (Pervaiz et al., 2016).

Natural fibers are already used in many composites to substitute glass fibers. They have equivalent mechanical properties but are less dense (Mohanty et al., 2002), which can reduce considerably the weight of the material and in results can help to save fuel when these materials are used in transports (Pervaiz et al., 2016). They also have a better environmental footprint than glass fibers and some biocomposites made from renewable matrices and natural fibers can possess a carbon neutral impact (Mohanty et al., 2002). One of the drawbacks of natural fibers is their variability in quality due to environmental conditions while glass fibers quality is constant (Staiger and Tucker, 2008).

Fibers are the macrofibrils coming from the primary and secondary walls of the plants and are composed of cellulose microfibrils embedded with hemicellulose, lignin and pectin. Fiber resistance is related to the cellulose amount (Staiger and Tucker, 2008), thanks to its complex structure giving it its mechanical strength (Wang et al., 2018). Flax and hemp fibers are thus a promising choice to improve TPS mechanical properties due to their large amount of cellulose (85-87% DW for flax fibers (Kozasowski et al., 2012) and 57-77% DW for hemp fibers (Ravi et al., 2018)).

The final properties of biocomposites are dependent of the initial compounds' properties (Balakrishnan et al., 2016) but also of the interface between each one of them (Gassan et al., 2000). Incompatibility can lead to a decrease in the mechanical properties of the biocomposite (Ravi et al., 2018). When studying the mechanical properties, the orientation and dispersion of fibers in the biocomposite, as well as the fiber dimensions (length and diameter), are parameters that will have an impact on the final mechanical properties of the biocomposite (Castellani et al., 2016).

In this study, native starch reinforced with flax and hemp fibers will be used. Microcrystalline cellulose fibers will also be used as it is widely studied as a standard for fiber integration in plastic matrices.

3. Characterization of TPS and biocomposites

Many analyses exist to evaluate the properties and the composition of biomaterials. Some common tests are used to study the mechanical properties of TPS and fibers reinforced TPS as well as most fossil-based plastics. Thus, these tests allow a comparison of TPS and biocomposites with bioplastics and fossil-based plastics already existing on the market and with many studies in that field of research. Other analyses can be performed to study the chemical bonds and functions present in the material. It can help understand how these materials evolve through the plasticization of starch and the formulation of biocomposites.

3.1. Mechanical analysis: Tensile test

Tensile tests are performed with a specific equipment illustrated in **Figure 8 (A).** As seen on this figure, the sample is attached to jaws by its opposite ends and are pulled apart by a force, increasing the distance at a constant rate. This results in an elongation and then usually the fracture of the sample. The instrument generates a stress-strain curve where the stress (σ) represents the load or force (F) divided by the surface area (A) of the sample ($\sigma = F/A$) and the strain (ε) is the change in length (ΔL) divided by the original length (L_0) of the sample ($\varepsilon = \Delta L/L_0$) (Wiederhorn et al., 2006). **Figure 8 (B)** presents a typical tensile curve. Several points/areas on the curve are used to characterize and compare samples in terms of mechanical properties. The Young's Modulus (YM), also called the elastic modulus, expresses the stiffness of the sample or its ability to be deformed reversibly. The total strain, also called the elongation at break (EaB), represents the maximum deformation a material can withstand before breaking. The (ultimate) tensile strength (TS) is the maximum stress the sample can bear and expresses the strength of the material (Zhang et al., 2014a).

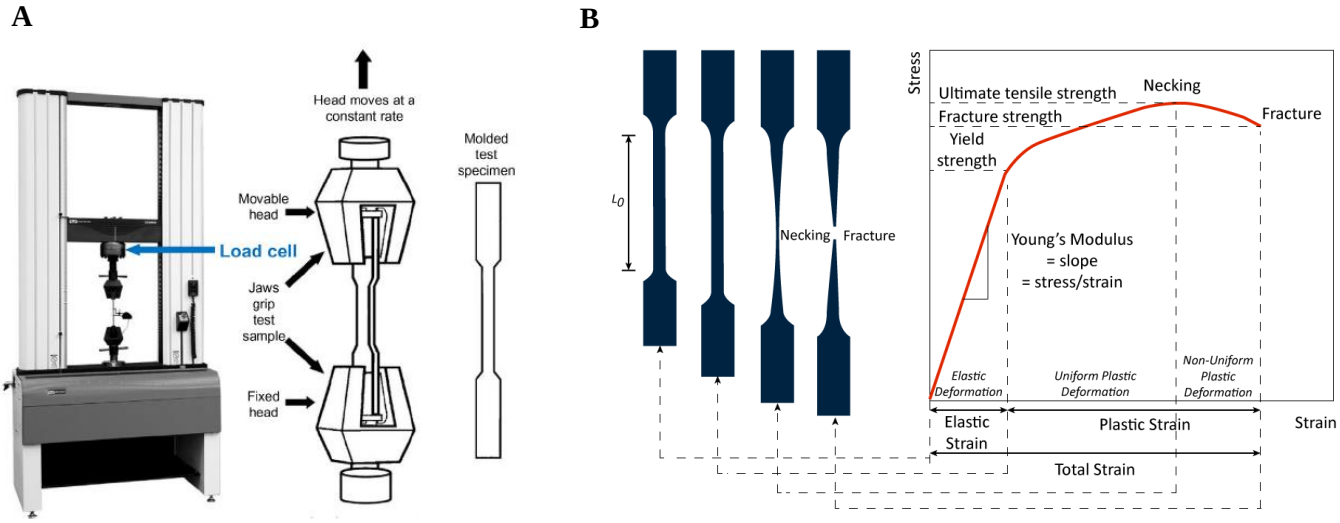


Figure 8: (A) Typical tensile testing instrument (Shrivastava, 2018c), (B) Deformation of a sample during a tensile test and the corresponding stress-strain curve (Yalcin, 2016)

3.2. Surface composition: Fourier-transform infrared (FTIR) spectroscopy

The sample is irradiated by infrared (IR) light, which is usually mid-IR (wavenumber between 4000 and 400 cm^{-1}). When the vibrational frequency of a bond corresponds to the IR frequency, IR light is absorbed. IR spectroscopy consists of recording the transmitted IR light that passed through the sample which brings information about the molecular structure (bonds and functions) of the sample (Wang and Chu, 2013).

The attenuated total reflectance (ATR) method, presented in **Figure 9**, is often used as it implies easy sample preparation and good reproducibility. The infrared beam enters the crystal, that is in direct contact with the sample and creates an evanescent wave that goes through that sample and carries its chemical information. The wave only enters the sample from 0.5 to 5µm, which makes this method a surface analysis (PerkinElmer, 2005). The detector detects the signal as an interferogram (Sharma et al., 2018). The latter is then converted into an IR spectrum through a mathematical transformation called Fourier-transform as presented in **Figure 10**. To obtain the percentage of transmittance in function of the wavenumber, the single beam spectrum must be normalized with the background spectrum, which consists of dividing the single beam spectrum signals by the background signals (Sharma et al., 2018).

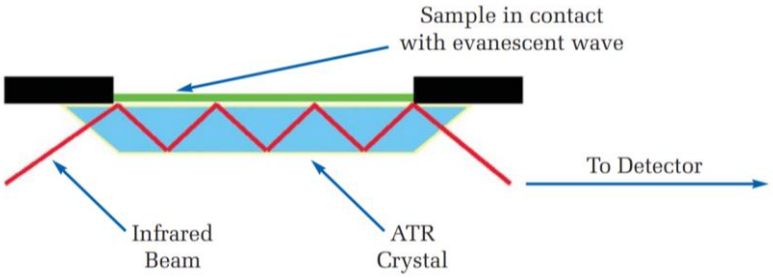


Figure 9: Illustration of the ATR method (PerkinElmer, 2005)

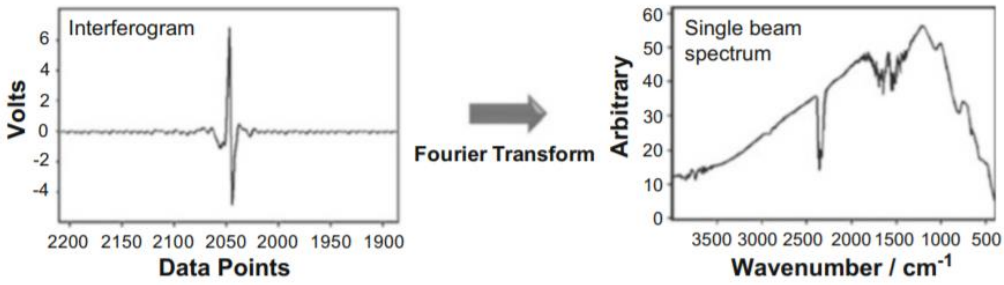


Figure 10: Fourier transform of an interferogram (Sharma et al., 2018)

Based on this literature review, this work will focus on how to facilitate the production of biocomposites on a small scale and how to identify the parameters impacting the final mechanical properties of biocomposites and the degree of their influence. The objectives presented in the next section will explain how to achieve these challenges.

II. Objectives

This work focuses on thermoplastic starch (TPS) and biocomposites (fibers reinforced TPS). These bio-based and biodegradable plastics can be an alternative to some non-biodegradable fossil-based plastics. This is a way to valorize starch found as a coproduct in the protein extraction process of some plants (e.g. pea, faba bean).

This master thesis will be divided in two parts.

The first part will study the microwave-assisted plasticization of TPS and biocomposites as a new processing method, under varied process parameters. A classification based on scores determined by several selection criteria will be performed. These samples will be analyzed by FTIR to understand the evolution of functional groups and linkages during the plasticization of TPS and after addition of fibers. Microscope observations will be performed to determine the impact of the process on the fibers' dimensions (lengths and diameters) and to understand the compatibility between the fibers and the TPS matrix. The orientation of the fibers in the matrix will also be studied.

These few analyses will allow to understand if this new processing method is suitable for the formulation of TPS and biocomposites, and to highlight the limits of the microwave for this kind of study.

In the second part, models will be designed to highlight the importance of the processing and chemical composition on the mechanical properties of biocomposites made from TPS matrices and fibers. Data from literature will be collected to create a database. Multilinear regressions will be performed to study the significance of parameters such as starch and fibers composition as well as process parameters on the tensile test measurements (Young's modulus, tensile strength and elongation at break).

III. Part 1: Microwave-assisted formulation of biocomposites and characterizations

1. Materials and methods

1.1. Materials

Table 3: Brand or origin of the raw materials with technical information

Name	Brand/origin
Glycerol	Alfa Aesar (99%)
Pea starch	Native NASTAR pea starch Provided by Cosucra Belgium (Dry matter 90±2%, ±35% Amylose, ±65% Amylopectin, granulometry <250µm, protein <0.5%, fat <0.4%, ashes <0.2% (Cosucra Socode, n.d.))
Microcrystalline cellulose fiber (C200)	Mikro-Technik GmbH (±200 µm in length and 30 µm diameter (Morin et al., 2019))
Flax fibers	Retted tow flax fibers harvested in France in 2014 and stored in dry and dark conditions before use. Chopped at a targeted length of 5mm (75.8% cellulose, 11.4% hemicellulose, 4.5% lignin, 8.2% extractible (Morin et al., 2019))
Hemp fibers	Retted technical hemp fibers (Fedora 17 or Santhica 27 variety), harvested in France in 2014 and stored in dark and dry conditions before use. Chopped at a targeted length of 5mm (81.0% cellulose, 11.0% hemicellulose, 2.5% lignin, 5.5% extractible (Morin et al., 2019))

Chemicals were used as received without further purification.

1.2. Microwave-assisted thermoplastic starch formulation

TPS were formulated in a microwave reactor (StartSYNTH, Milestone Srl) with distilled water and glycerol. An optimized experimental plan was generated using the JMP 15 statistical software (SAS Institute Inc.), presented in **Table 5**. The aim was to determine which area(s) of the tested matrix gave promising TPS, according to selected parameters.

1.2.1. Samples preparation

Pea starch and plasticizers were weighted directly in the microwave Teflon tubes. The chip was placed first, then glycerol, starch, and water were added in that order. The mixture was stirred by vortex for

1.5min, and then for 5min by magnetic stirring at 250RPM, to ensure proper homogenization. The tubes were placed into closed vessels and set up at equal distance in the microwave reactor as shown in **Figure 11**. Each sample was triplicated within the same batch, one repetition with the temperature probe (C) and two repetitions without (A & B).

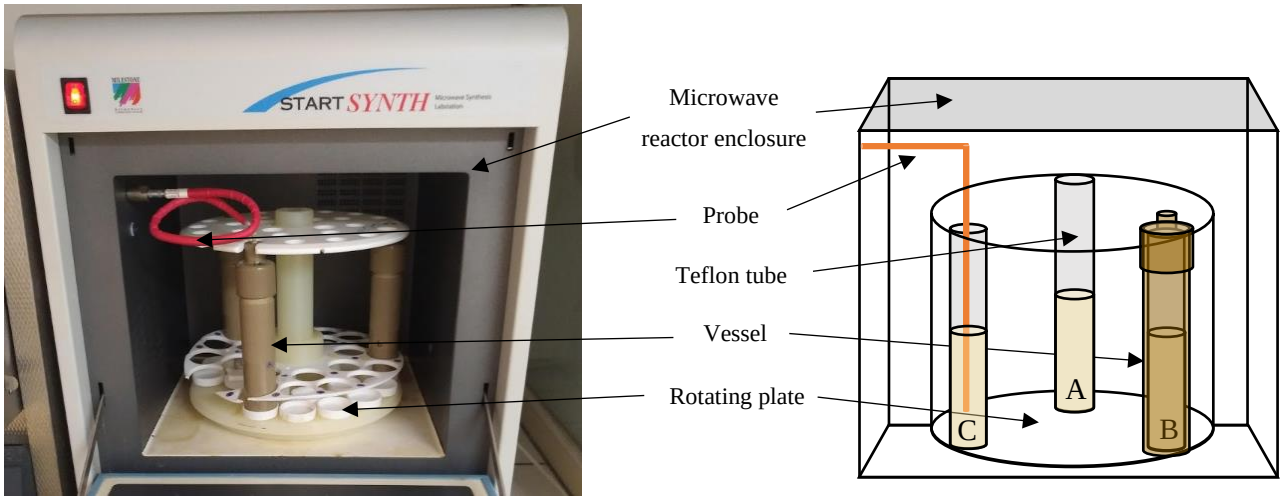


Figure 11: Set up of the samples in the microwave. Left: photo of the system. Right: representation of the microwave enclosure with the samples on the rotating plate and the probe in the sample C.

After plasticizing, the samples were cooled down in a room temperature water bath for 10min before collecting them from the tubes. Collected samples were kept at 80% relative humidity at room temperature, in a desiccator with aluminosilicate gel (from Merck), for a week before being sliced and stored into closed plastic cups and further analyzed.

1.2.2. Process parameters selection and experimental plan

Four process parameters were studied:

- the targeted temperature in the microwave reactor during plasticization
- the heating time to reach the targeted temperature (see **Figure 12**)
- the treatment time at the targeted temperature (see **Figure 12**)
- the percentage of starch (wt%) in the mixture, where water and glycerol completed the mix in equal proportions.

The heating power was set to the device maximum (1200W) during the entire treatment to allow the microwave to adjust the probed mixture to the required temperature. The stirring was set to 20% of nominal power, which should have been enough to ensure good homogenization during the treatment but not too intense, knowing that the formulated fluids were rheo-thickening.

Figure 12 presents the microwave interface and shows how the process parameters were set up and how the power and temperature are monitored during a typical treatment. The first line on the left picture represents the process parameters for the heating period and the second for the treatment period.

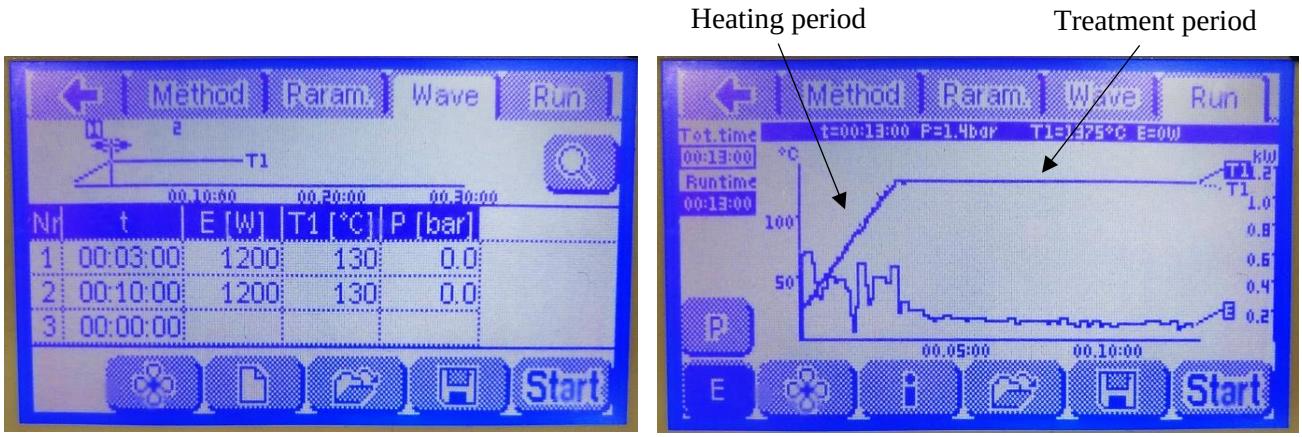


Figure 12: Examples of the microwave process parameters set up (left) and the temperature and energy monitoring curves (right)

Table 4 presents the range of values for each process parameter studied, where the extreme values were selected based on the literature (see appendix 1). Then, an optimized experimental plan, presented in **Table 5**, was generated with the JMP 15 statistical software including first and second interactions between factors. As presented in **Table 4**, the software added an in-between value for each process parameter, which allowed to have enough combinations of process parameters to study the first and second interactions.

Table 4: Range of values for process parameters used for the optimized experimental plan of TPS formulation

	Minimal value	**In-between value**	**Maximum value**
Temperature (°C)	130	160	190
Time of heating (min)	1	2	3
Time of treatment (min)	0.5	5.25	10
Starch (% w/w)	20	40	60

Table 5: Optimized experimental plan of the TPS formulation

Sample	Temperature (°C)	Time of heating (min)	Time of treatment (min)	Glycerol (%)	Starch (%)	Water (%)
TPS 1	130	1	0.5	40	20	40
TPS 2	130	1	0.5	20	60	20
TPS 3	130	1	10	30	40	30
TPS 4	130	2	5.25	30	40	30
TPS 5	130	3	0.5	30	40	30
TPS 6	130	3	10	40	20	40
TPS 7	130	3	10	20	60	20
TPS 8	160	1	5.25	30	40	30
TPS 9	160	2	0,5	30	40	30
TPS 10	160	2	5.25	40	20	40
TPS 11	160	2	5.25	20	60	20
TPS 12	160	2	10	30	40	30
TPS 13	160	3	5.25	30	40	30
TPS 14	190	1	0.5	30	40	30
TPS 15	190	1	10	40	20	40
TPS 16	190	1	10	20	60	20
TPS 17	190	2	5.25	40	20	40
TPS 18	190	2	5.25	28.7	42.6	28.7
TPS 19	190	3	0.5	40	20	40
TPS 20	190	3	0.5	20	60	20
TPS 21	190	3	10	30	40	30

1.2.3. Samples selection

The formulated TPS were selected using a scoring grid based on selection criteria presented in **Table 6** (observations without equipment). For each criterion, a score was attributed, giving a final score for each TPS.

Table 6: Formulated TPS selection criteria with associated score

Criterion	Score
Consistency	0 = Hard block, powder or liquid 1 = Block between soft and hard 2 = Soft block 3 = Paste that can be molded
Color	0 = Dark orange, brown 1 = Light yellow to light orange 2 = White or transparent
Color homogeneity	0 = Non-homogeneous 1 = Almost homogeneous 2 = Homogeneous
Air bubbles	0 = Big air bubbles (cavities) or air bubbles of different sizes 1 = Small homogeneous air bubbles 2 = No air bubbles or very few

Depending on the score obtained, the TPS were classified as optimal (complete plasticization), non-optimal (uncomplete plasticization) or promising (optimal if the homogeneity is improved). This classification was defined following **Table 7**.

Table 7: Classification of the formulated TPS based on the score obtained based on the selection criteria

TPS Classification	Score
Optimal	≥ 7
Promising (homogeneity to improve)	≥ 4 and ≤ 6
Non-optimal	< 4

1.3. Biocomposites formulation in a microwave reactor

Biocomposites were formulated in the microwave reactor from the TPS samples that were optimal and from non-modified natural fibers: microcrystalline cellulose fiber (C200), flax fibers, and hemp fibers.

1.3.1. Sample preparation

The samples preparation was similar to TPS. First, the TPS mixture was prepared as before (starch with plasticizers and homogenization). Then, the fibers were added to the mixture and blended by hand. After the process in the microwave, the samples were also cooled down in a room temperature water bath for 10min before removal from the tubes.

The process parameters, as well as the compositions of the TPS matrices, were the same as the TPS samples selected (all composed of 20% starch, 40% water and 40% glycerol (w/w)). The biocomposite total mass was composed of 90 or 95% (w/w) of TPS matrix and 10 or 5% of fibers. The corresponding experimental plan is presented in **Table 8**.

The biocomposites samples are denominated BC followed by the temperature of process, the first letter of the type of fiber, and the percentage of fiber. As example, BC130C10 is a biocomposite processed at 130°C with 10% of C200 fibers. For more clarity, the TPS n°6, 10 and 17 will be denominated TPS130, TPS160, and TPS190 in reference to their process temperature. It will be explained later why these three TPS were selected.

Table 8: Experimental plan of biocomposites formulation in the microwave

Sample	TPS matrix	Temperature (°C)	Time of heating (min)	Time of treatment (min)	Type of fiber	Fiber percentage (%)
BC130C10	N°6 (TPS130)	130	3	10	C200	10
BC130F10		130	3	10	Flax	10
BC130H10		130	3	10	Hemp	10
BC160C10	N°10 (TPS160)	160	2	5.25	C200	10
BC160F10		160	2	5.25	Flax	10
BC160H10		160	2	5.25	Hemp	10
BC190C10	N°17 (TPS190)	190	2	5.25	C200	10
BC190F10		190	2	5.25	Flax	10
BC190H10		190	2	5.25	Hemp	10
BC190C5	N°17 (TPS190)	190	2	5.25	C200	5
BC190F5		190	2	5.25	Flax	5
BC190H5		190	2	5.25	Hemp	5

1.3.2. Sample selection

The formulated biocomposites were selected using a scoring grid, similar than for TPS, based on selection criteria presented in **Table 9** (observations without equipment). For each criterion, a score was attributed, giving a final score for each biocomposites.

Table 9: Formulated biocomposites selection criteria with associated score

Criterion	Score
Consistency	0 = Hard block, powder or liquid 1 = Block between soft and hard 2 = Soft block 3 = Paste that can be molded
Color homogeneity	0 = Non-homogeneous 1 = Almost homogeneous 2 = Homogeneous
Air bubbles	0 = Big air bubbles (cavities) or air bubbles of different sizes 1 = Small homogeneous air bubbles 2 = No air bubbles or very few

Depending on the score obtained, the biocomposites were classified as optimal, non-optimal, or promising (optimal if the homogeneity is improved). This classification is presented in **Table 10** and was defined following **Table 9**.

Table 10: Classification of the formulated biocomposites based on the score obtained based on the selection criteria

Biocomposites Classification	Score
Optimal	= 7
Promising (homogeneity to improve)	≥ 4 and ≤ 7
Non-optimal	< 4

1.4. Characterization of samples and their compounds

1.4.1. Fibers morphology characterization with an optical microscope

Fibers isolated from the biocomposites as well as unprocessed fibers were observed with an optical microscope (Leica DM2700 P). Pictures were taken and the diameter (in µm) and length (in mm) of fibers were measured with the computer program ImageJ. The Weibull distributions (only for positive variables) of these measurements were compared. Pictures of fibers in the TPS matrix were also taken to observe the orientation of the fibers in that matrix.

1.4.2. FTIR analyses

The TPS samples selected, their corresponding biocomposites as well as their initial compounds such as glycerol, starch, water and fibers (before and after integration), were analyzed by an ATR-FTIR spectrometer (Bruker Vertex 70) under an inert nitrogen atmosphere. The transmittance was recorded in function of the wavenumber from 4000 to 400 cm^{-1} with a resolution of 4 cm^{-1}.

The FTIR spectra were analyzed statistically with the R software (The R foundation, Version 4.0.0) using a code presented in appendix 5. The goal was to determine whether the process parameters were related to the presence of some chemical bonds (covalent and hydrogen bonds) between the different samples. The code performed different types of spectra transformations (normalization, first and second derivations and combinations of these). Several correlations were estimated between each spectrum transformation and each studied parameter (i.e. formulation conditions, chemical composition). Peaks were selected according to their ability to improve fitting the correlation of the established model.

2. Results and discussion

2.1. TPS formulation

The TPS samples formulated in the microwave were classified based on their final score determined by the selection criteria presented before. These data are presented in appendix 2. As a reminder, four process parameters were tested in the microwave to understand which conditions gave optimal TPS samples. With the resulting TPS scores, we can see if any process parameter has an impact on the final score of formulated TPS. It is important to note that the following results are general assumptions based on visual observations and are specific to the experimental conditions.

Figure 13 represents the total scores of TPS formulated in the microwave in relation to the percentage of starch in the mixture and the process temperature. As seen on that figure, there is a trend between the scores and the starch percentage; the scores are higher when the percentage of starch is low. In fact, all the TPS samples that were classified as optimal (score ≥ 7) had a starch percentage equal to 20.

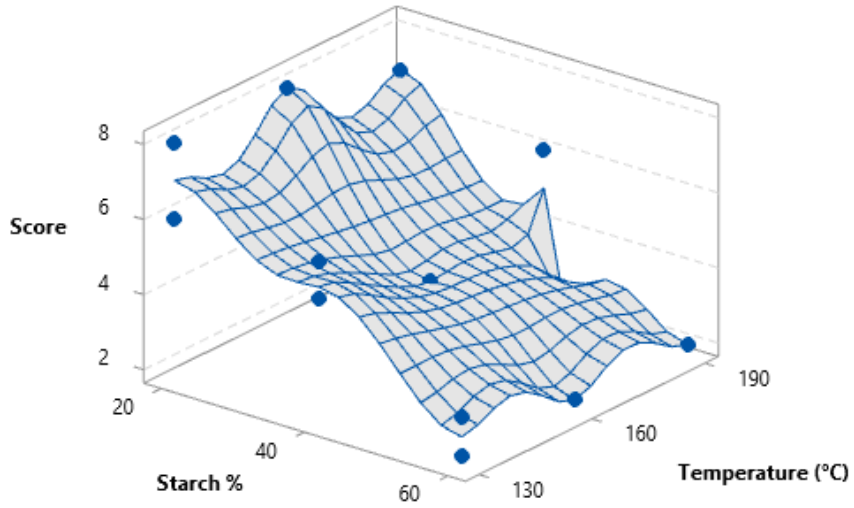


Figure 13: Surface diagram of TPS formulation in the microwave: TPS scores in function of starch percentage and temperature of the microwave

To have a better idea of the relationship between the starch percentage and the TPS scores, these two variables are presented in **Figure 14**. The estimated linear trend has an $R^2 = 0.84$, which is not particularly high to build a precise model but is high enough to show the evolution of the score depending on the starch percentage. As stated before, these data are from visual observations and provide general insights.

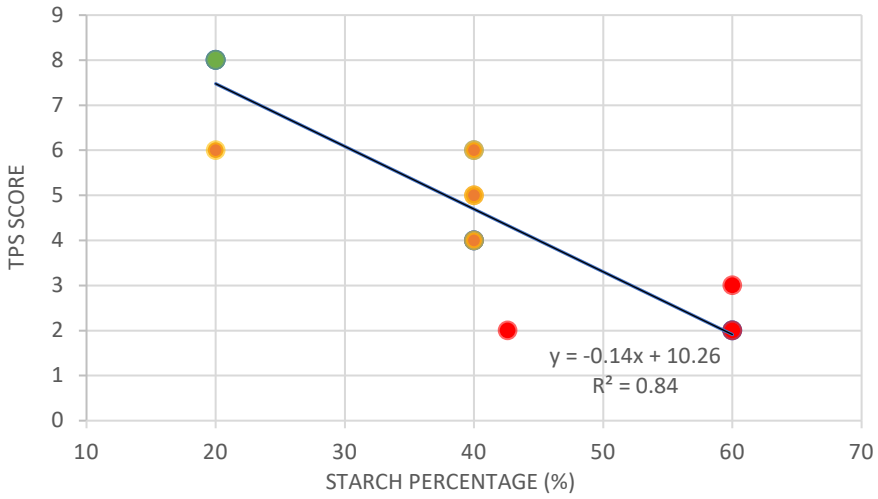


Figure 14: Final scores of microwave formulated TPS in function of the starch percentage in the TPS mix. Green = optimal, orange = promising, red = not optimal.

As for the temperature of treatment, optimal TPS samples were found for each tested temperature (130, 160 and 190°C) and **Figure 13** shows that there is no clear relationship between the score and the temperature. In fact, the estimated linear trend between them has an R^2=0.0056 (see appendix 4). However, among the optimal samples (score ≥ 7), some differences in terms of consistency can be noticed and are illustrated in **Figure 15**. The samples formulated at 130°C and 160°C (**Figure 15 A and B**) had a soft consistency but took the shape of the microwave tube (consistency score = 2). In contrast, the ones formulated at 190°C (**Figure 15 C**) were more liquid and could be molded in the shape of the plastic cup, after being removed from the tubes (consistency score = 3). When considering processing and applications where the samples need to be molded, it seems that 190°C is the most suitable temperature tested for these conditions.

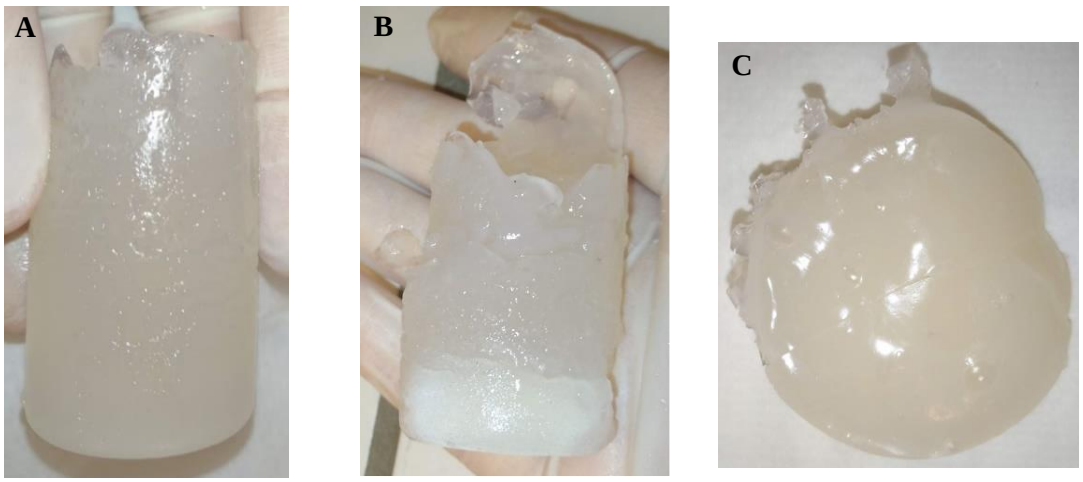


Figure 15: Pictures of the microwave formulated TPS. A: TPS n°6 (130°C), B: TPS n°10 (160°C), C: TPS n°17 (190°C)

For the times of heating and treatment, the data from this experimental plan, presented in **Figure 16**, do not show a significant linear trend between those two factors and the TPS scores. The estimated linear trend between the time of heating and the TPS scores has an $R^2=0$ and the one between the time of treatment and the TPS scores has an $R^2=0.0034$ (see appendix 4).

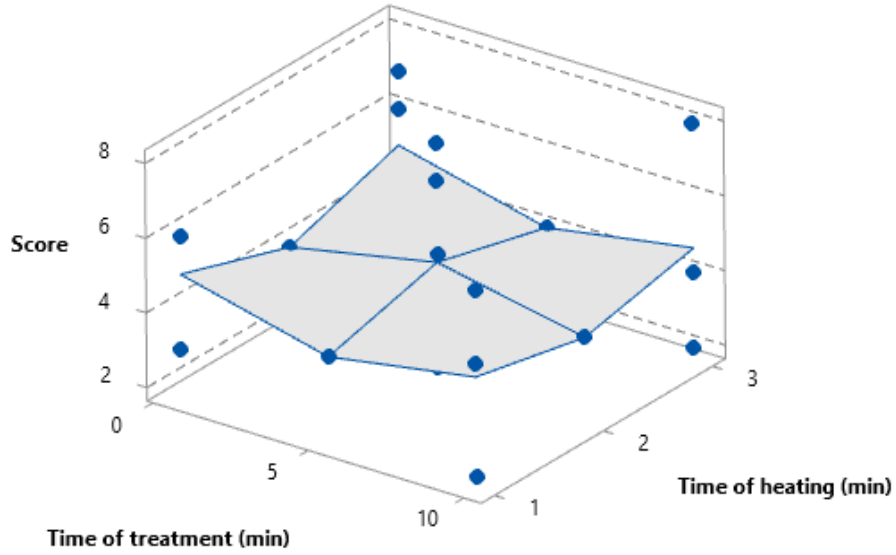


Figure 16: Surface diagram of TPS formulation in the microwave: TPS scores in function of time of heating and time of treatment

All the samples that were processed with the same temperature and the same amount of starch were classified in the same category except for one case. The samples that were processed at 130°C with 20% of starch and with different times of heating and treatments were not classified the same. This difference comes from the large air bubbles present in the sample processed with shorter times of heating (1min) and treatment (0.5min), as seen in **Figure 17 A and B**, which was classified as promising. The other sample (3min of heating and 10min of treatment) was completely homogeneous and was classified as optimal (**Figure 17 C and D**).

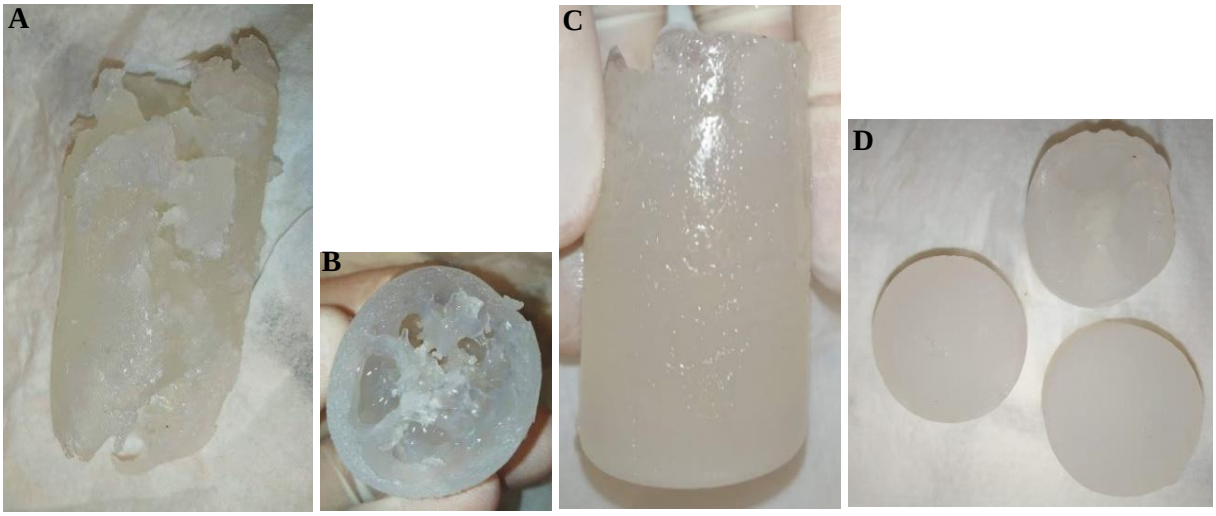


Figure 17: Pictures of the microwave formulated TPS. A,B: 130°C, 20% starch, 1min heating, 0.5min treatment, C,D: 130°C, 20% starch, 3min heating, 10min treatment. C and D are sliced samples.

To be able to understand what non-optimal samples look like, some examples are presented in **Figure 18**. All the presented TPS show heterogeneous plasticization. Picture B,C and D also show overheating zones and probably sample degradation.

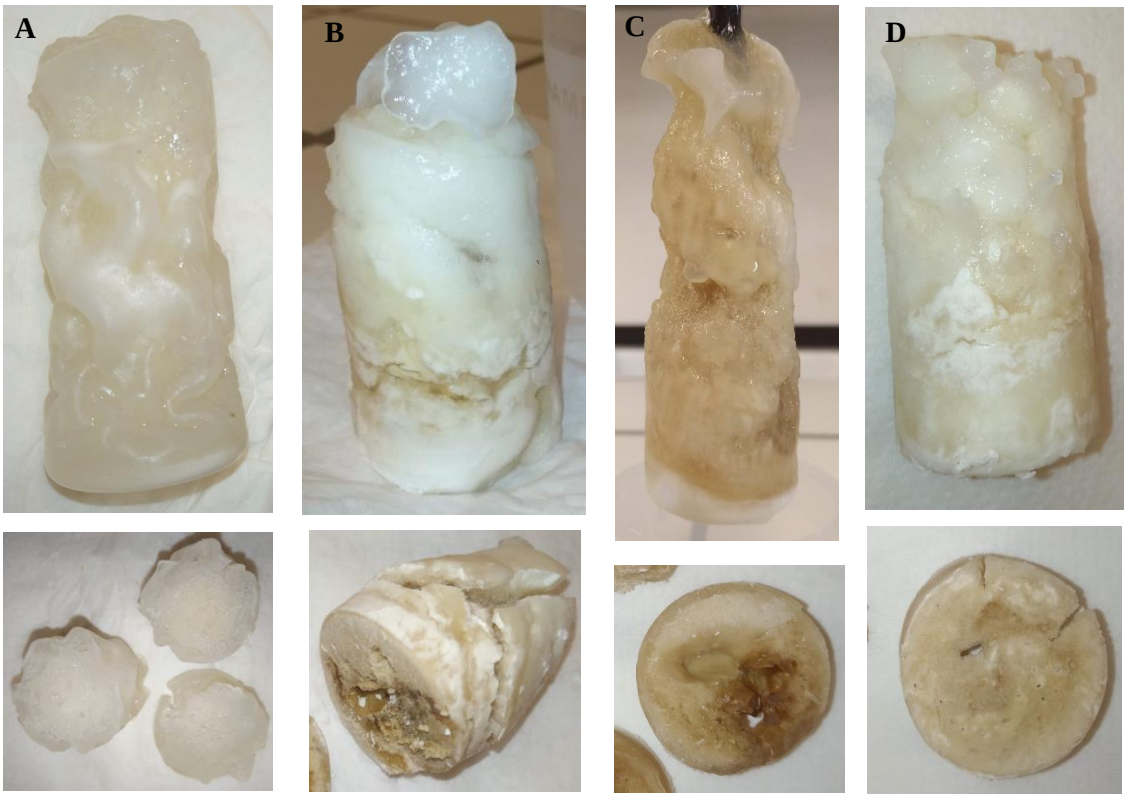


Figure 18: Pictures of the microwave formulated TPS. A: TPS n°8, B: TPS n°11, C: TPS n°16, D: TPS n°20. Bottom pictures are the sliced samples.

From these results, it can be concluded that:

- The optimal samples are found in each range of temperature but only the optimal ones formulated at 190°C have the ability to be molded.
- The percentage of starch impacts the final score of the TPS. In this experiment, 20% (w/w) of starch gives the highest scores.
- These data do not show a clear relationship between the times of heating and treatment and the final TPS scores, in the tested range.
- For the lower temperature tested (130°C), it seems that a longer time is required to obtain a homogeneous sample.

As shown in **Figure 15**, the temperature of treatment had an impact on the final consistency of the TPS sample. Only the optimal samples processed at 190°C were liquid enough and able to be molded. This could be explained by the temperature of demolding. All samples were cooled down for 10min in a room temperature water bath, meaning that the samples that reached 190°C had a higher temperature of demolding than the ones at 130 and 160°C. It is possible then that the samples processed at 190°C were still warm enough to be molded when the others were already too cold. As a consequence, it is possible that the samples processed at higher temperatures continued the plasticization process while cooling down and thus had a longer time of plasticization.

In the case of this study, it seems that the microwave was only suitable when the TPS mixture was in a liquid form with lower viscosity, i.e. containing 20% (w/w) of starch. Previous studies have shown promising results for TPS with starch percentages up to around 80% (Averous and Boquillon, 2004; Ma et al., 2008; Thunwall et al., 2008). The main difference lies in the processing device. In fact, TPS is often formulated with an extruder, mainly different in three aspects compared to the microwave: the way samples are prepared, the heating and the stirring method.

In an extruder, all the compounds are fed through the hopper. Then, the TPS mixture is constantly stirred with an endless screw and the mixture is in contact with a heating barrel (as presented in **Figure 7**). This device allows optimal homogenization and homogeneous heating. Concerning the heating in the microwave, it is known to heat samples more evenly than other conventional heating methods as the energy targets each molecule in the sample. The monitoring curves show that all the temperatures were reached in the center of the tubes, even when the period of heating was short (< 3min).

The main problem seems to come from the stirring during the sample preparation and in the microwave. The magnetic chips chosen were small to fit the tubes. As the samples were rheo-thickening or had a powdery consistency, it was not enough to offer proper homogenization (for starch percentages above 20%) before and during the treatment. This means that the mixtures with more than 20% of starch were heterogenous before the treatment, with some areas containing more plasticizers than others. This

resulted in samples with heterogeneous plasticization or no plasticization at all. The speed, the size, and shape of the magnetic chip are thus important to consider, depending on the composition of the mixture.

The homogenization is the limiting parameter for the microwave formulation of TPS. The samples must be liquid enough, meaning containing low starch amounts. From this observation, two potential solutions are here presented in order to produce TPS in a microwave with higher percentages of starch.

The first is the sample preparation. Many studies that have produced TPS with high starch percentages (Averous and Boquillon, 2004; Ma et al., 2008) prepare their samples by blending the starch and the plasticizers with a mixer. The samples are then stored for hours up to weeks to allow the plasticizers to penetrate between the starch particles before the plasticization process. This could be a good solution, especially when the starch percentage is high (from 60%) to avoid heterogenous plasticization where only parts of the sample are mixed with the plasticizer(s).

Another improvement would be to set up the microwave reactor differently. By adding a mixer from the top, as presented in **Figure 19**, optimal homogenization would be applied during the treatment. It would imply the process of one sample at a time and would need a larger container. The probe and the mixer would need to pass through the lid in a hermetic way, if no loss of water is wanted.

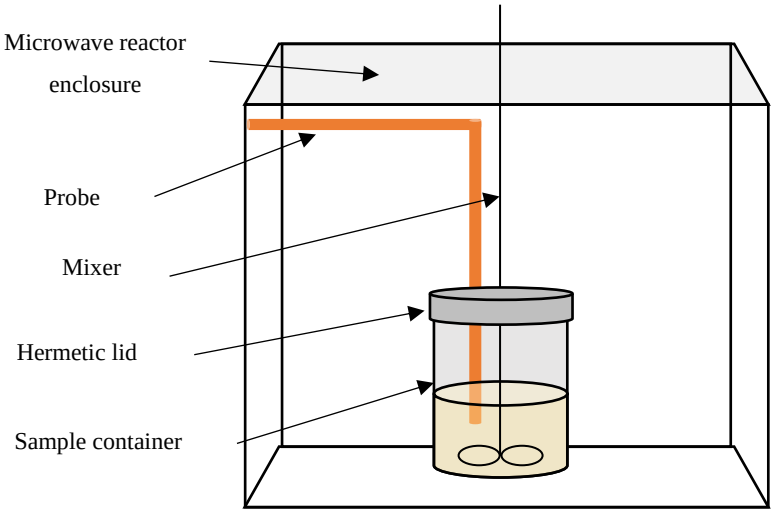


Figure 19: Representation of a new microwave set up

2.2. Biocomposites formulation

After the classification of TPS samples, three of the five optimal TPS were selected (one for each temperature; TPS130, TPS160 and TPS190) and fibers were added to formulate biocomposites (**Table 8**). As a reminder, all the selected TPS samples were composed of 20% starch, 40% water and 40% glycerol (w/w).

As well as for the TPS, the biocomposites were classified based on their final scores determined from selection criteria presented before (**Table 9**). These final scores as well as the detailed scores for each criterion are represented in the appendix 3.

The color criterion was not used to classify the biocomposites. This criterion was used to classify the TPS and to select the ones that did not show strong coloration, which could have been an indicator of starch thermo-degradation. As a result, all the TPS samples selected (and further used to formulate the biocomposites) did not show strong coloration (white or lightly yellow). The biocomposites' color mainly came from the fibers and could not be used to classify the samples.

In terms of fibers homogeneity, all samples seemed to have a good fiber dispersion, after visual comparison. However, it was not possible to differentiate the samples visually, that is why this criterion was not used to determine the biocomposites score. Samples analysis, such as FTIR analysis will determine if some samples are more homogeneous than others in terms of composition.

The final scores of biocomposites are represented in **Figure 20**. First, 10% ((w/w) compared to TPS mass) of each fiber were added, for each temperature. Many studies have added fibers from 5 to around 20% (Avérous et al., 2001; Averous and Boquillon, 2004; Curvelo et al., 2001; Gironès et al., 2012). It seemed a good compromise to start with 10% of fibers, knowing that the mixture would thicken from pretests. None of the samples were considered optimal (total scores between 1 and 4). However, as seen on the figure, as the temperature increased, the scores increased also. With an addition of 10% (w/w) of fibers, the best consistency was obtained at 190°C but the samples were still too viscous to be molded and not homogeneous enough.

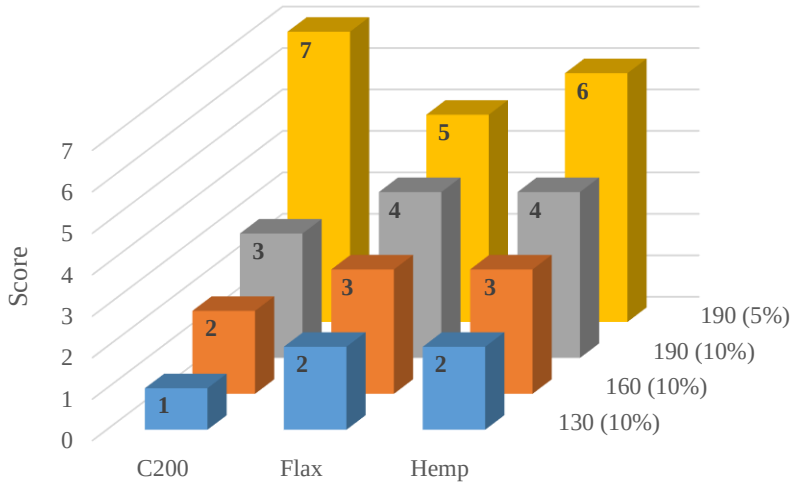


Figure 20: 3D histogram of biocomposites scores formulated in the microwave in function of the temperature of treatment and fiber type. The fiber percentages are in brackets. The numbers represent the scores.

After seeing that the results with 10% of fibers were not optimal, other tests were performed with an addition of 5% (w/w) fibers, at the treatment temperature of 190°C. The scores increased when the percentage of fiber was decreased. **Figure 21** shows that the samples had an improved ability to be molded, and the one with C200 had a comparable appearance with the corresponding optimal TPS. However, the ones with flax and hemp fibers were not optimal in terms of homogenization. That is why only the biocomposite formulated at 190°C with 5% C200 was classified as optimal (score = 7). Hemp fibers samples were very similar than flax fibers samples.

As for the TPS samples, the raise of process temperature improved the consistency of the biocomposites samples and their ability to be molded. It could come from the same explanation, namely the demolding temperature being higher when the process temperature is higher, which gives a better ability to be molded.

Once again, the homogenization was a limiting step in the preparation and the process of biocomposites. After the TPS mixture was prepared as before, the fibers were added and stirred by hand until having a homogeneous paste. The more fibers, the thicker the paste and therefore the more difficult it was to blend in the microwave, just as for the percentage of starch in the TPS matrix. It is clear that the magnetic chip was not adequate to have proper homogenization. This explains that the decrease of the fiber percentage in the mixture increased the biocomposites final scores as well as their ability to be molded. This phenomenon could limit the improvement of the mechanical properties as it was demonstrated that the Young's modulus and the tensile strength increase with the amount of reinforcing fibers (Avérous et al., 2001; Gironès et al., 2012; Prachayawarakorn et al., 2010).

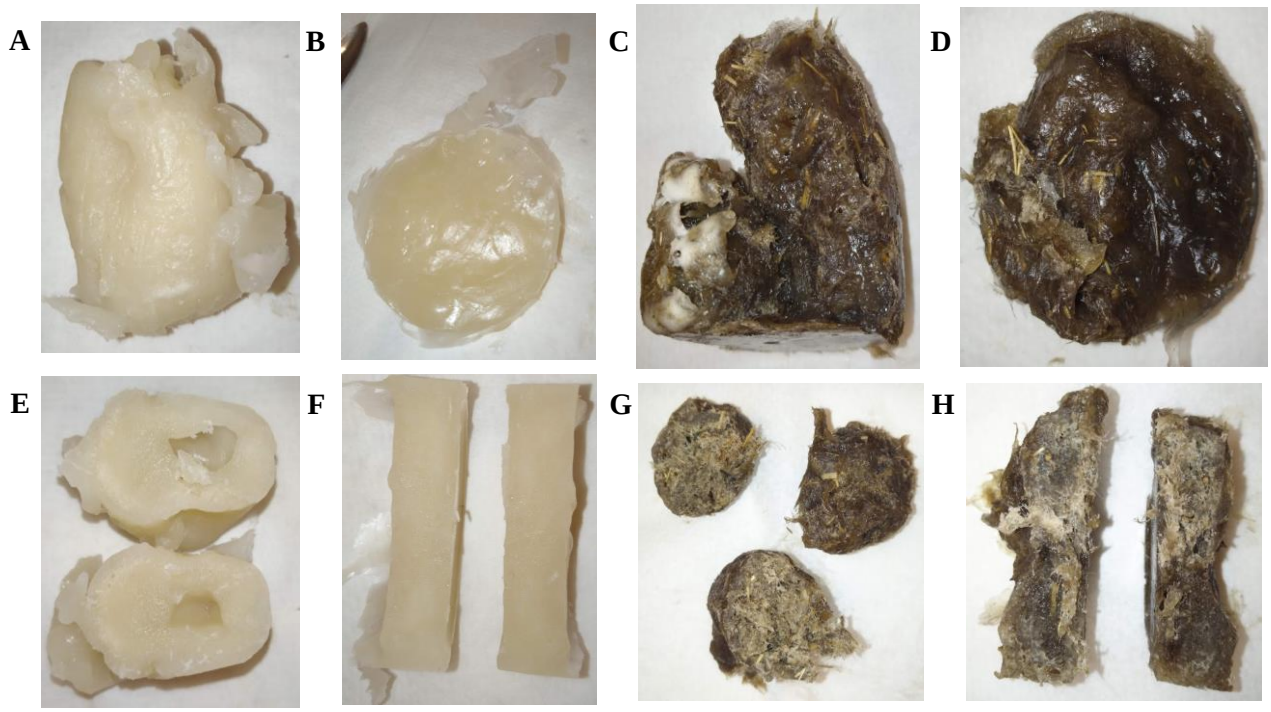


Figure 21: Pictures of biocomposites. A, B: 10% and 5% C200 fibers, C,D: 10% and 5% flax fibers. Pictures from E to H are the sliced samples of the ones above.

Results also show that biocomposites processed at 190°C made with C200 fibers seemed more homogeneous in terms of consistency than with other fibers. This could be explained by the size of the different fibers. The C200 fibers have a length of ±0.2 mm when the flax and hemp fibers have a length of ±5mm. It is then easier for the C200 fibers to disperse in the mixture and it gives more liquid samples before process. In results, the stirring in the microwave is more efficient, the samples have a better consistency and are more homogeneous in appearance.

Figure 22 and 23 present images of biocomposites with flax and hemp fibers observed with an optical microscope. Both of these pictures show that the fibers are positioned in many directions and cross each other, as pointed by the arrows. In the biocomposite, the forces are absorbed by the fibers. When the fibers are all in the same direction, the mechanical strength of the material is dependent of that direction (Mathes, 2018). In the case of this study, the mechanical strength is not direction dependent and the biocomposites seems to be able to handle forces from many directions.

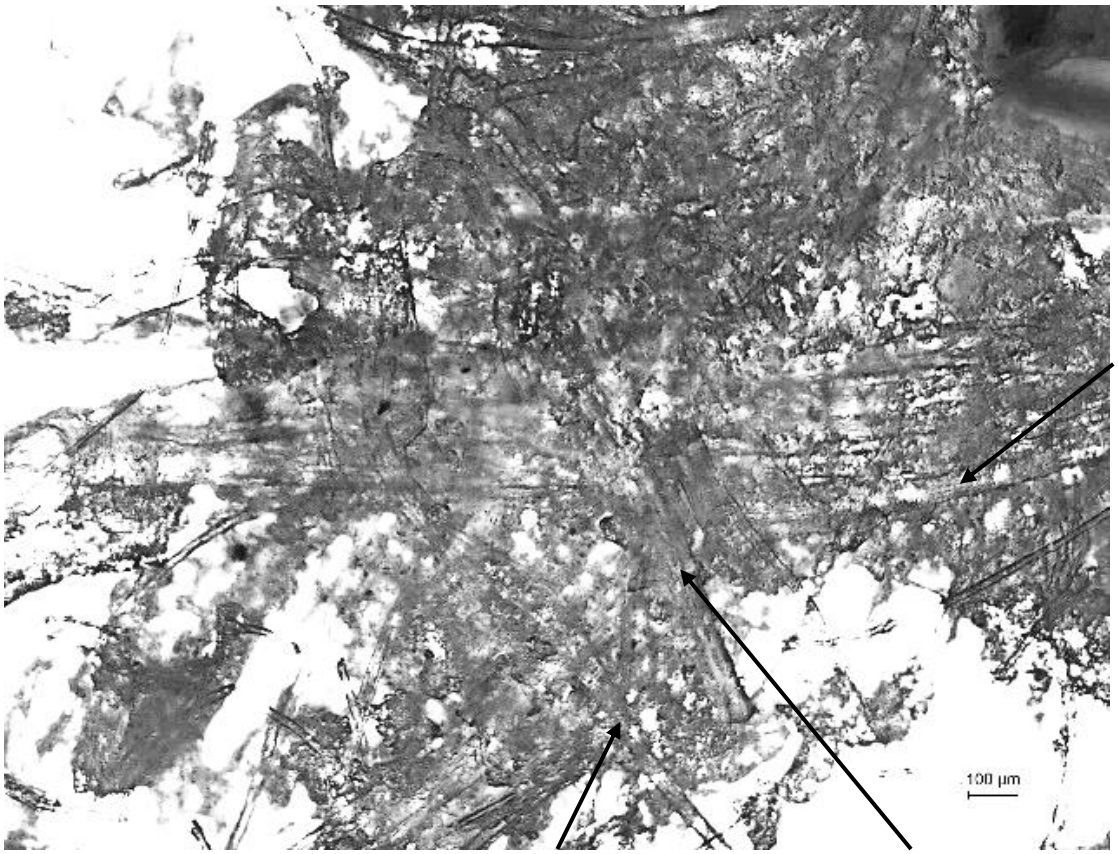


Figure 22: Optical microscope images of biocomposite with flax fibers

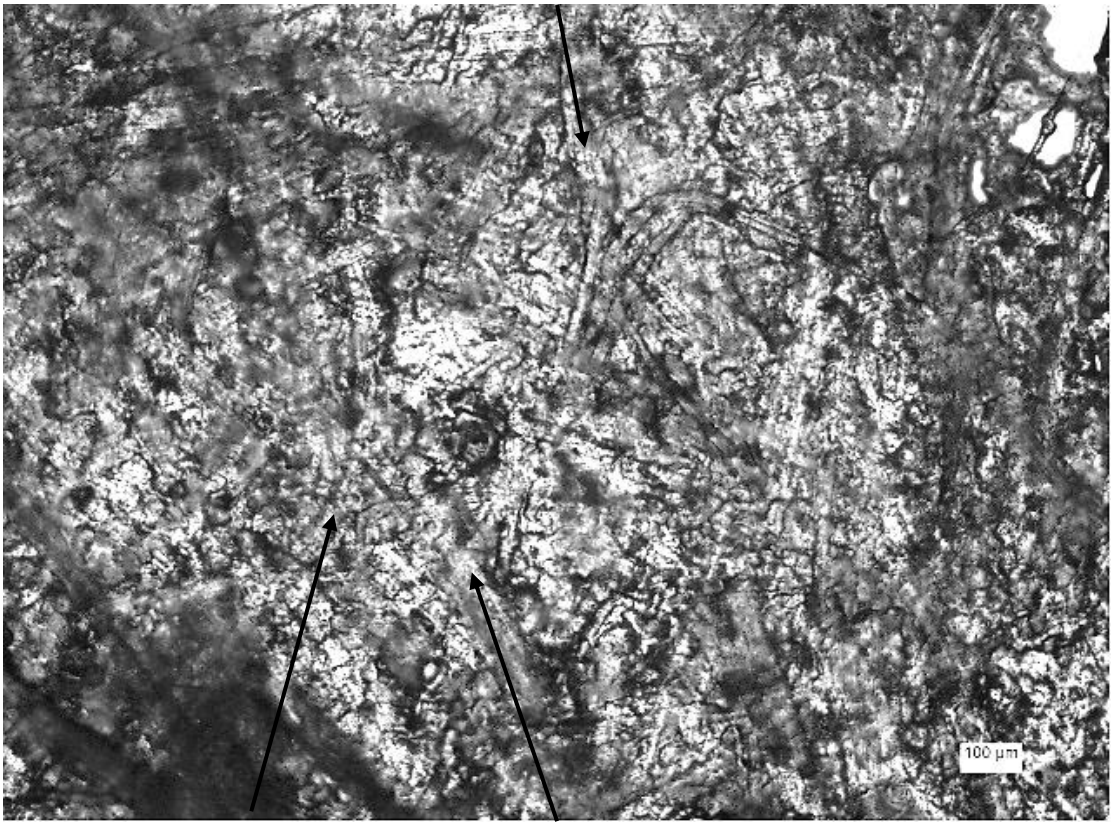


Figure 23: Optical microscope images of biocomposite with hemp fibers

2.3. Fibers morphology characterization with an optical microscope

The diameters and length of flax and hemp fibers have been measured before and after plasticization into biocomposites. Their distributions are represented in **Figure 24**. C200 fibers were not observed as they were too small to be isolated from the biocomposites. The objective was to study the impact of the microwave-assisted plasticization on the fibers' dimensions (diameter and length) as it is known that some processes, like extrusion, reduce the fibers length during composite formulation (Castellani et al., 2016). Another objective was to observe the way fibers interact physically with the TPS matrix and so evaluate their compatibility.

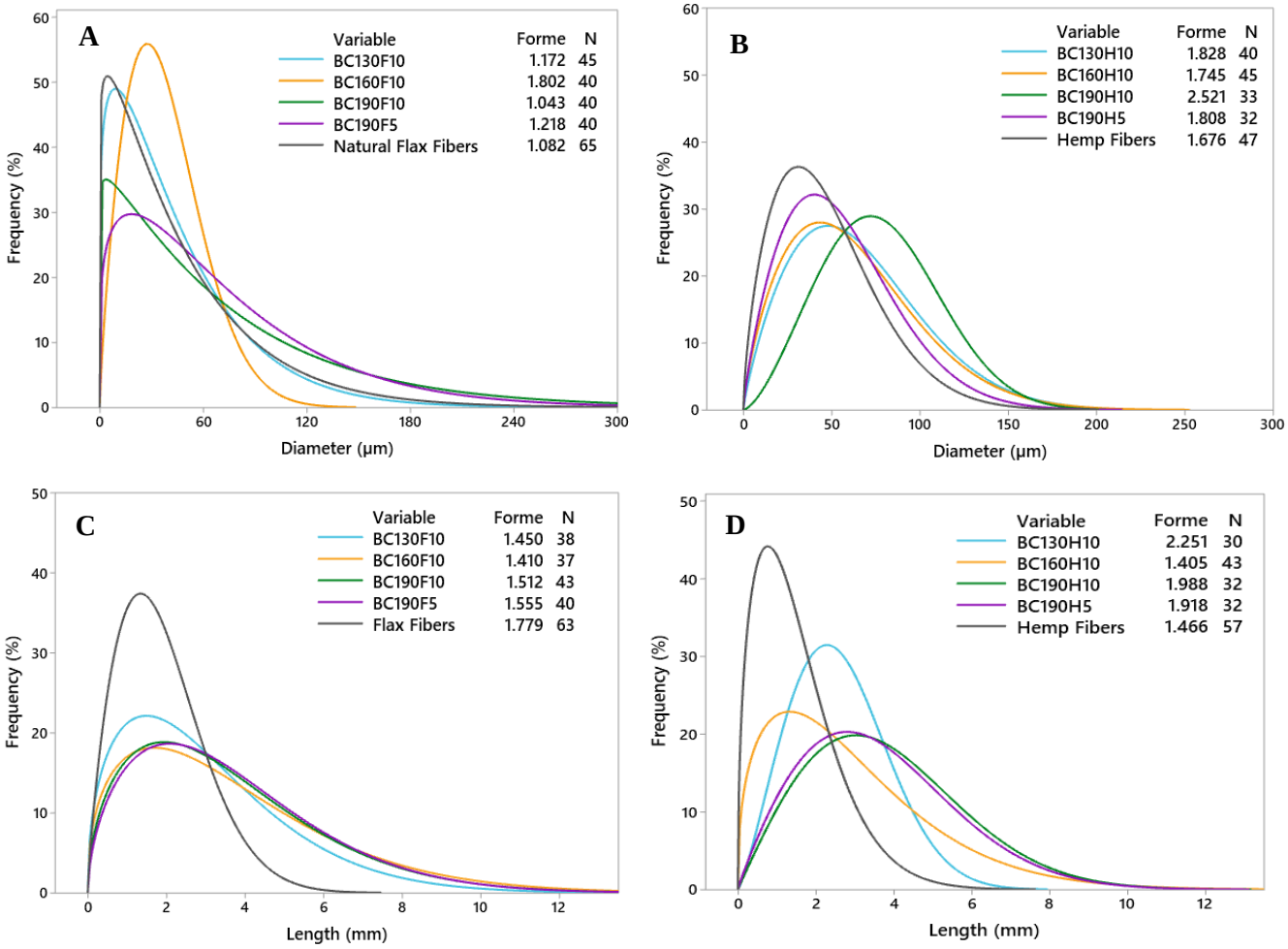


Figure 24: Distribution of natural hemp and flax fibers before and after integration into biocomposites. A: Flax fibers diameters, B: Hemp fibers diameters, C: Flax fibers lengths, D: Hemp fibers lengths. «Forme» = shape and N is the number of fibers measured.

To study the symmetry of the studied populations, the shape factor (referred as "Forme" in **Figure 24**) was studied. A shape of 3 represents a "normal curve"-shape, a shape close to 1 represents a right-skewed curve and a shape close to 10 represents a left-skewed curve.

As seen in **Figure 24A**, there is not a clear change of the diameter distribution of the biocomposites (BC) flax fibers, before and after plasticization. The distribution of BC130F10 is close to the unprocessed fibers and the great majority of the diameters is between 0 and 180µm. BC160F10 has, on average, smaller diameters with a range between 0 and 120µm. BC190F10 and BC190F5 show the largest diameters distribution with a range between 0 and 240µm. BC160F10 is the closest to a normal distribution diameter with a shape of 1.802, centered around 30µm. The other ones have shapes closer to 1 and the diameters are centered around 0 to 20µm.

The BC hemp fibers diameters increase slightly after plasticization as illustrated in **Figure 24B**. The majority of the unprocessed hemp fibers have a diameter range between 0 and 150µm and BC190H5 between 0 and 175µm. All the biocomposites with 10%(w/w) fibers have a diameter range between 0 and 200µm. BC190H10 has a distribution closer to a normal curve (shape=2.521), than all the other distributions, where the shapes have a value around 1.75. BC190H10 has diameters centered around 75µm and the other ones around 25 to 50µm.

Figure 24C shows that the BC flax fibers lengths have increased in a similar way after each treatment. The length range increases from 0 to 6 mm for the majority of the unprocessed fibers to 0 to 12mm for the other ones. A flattening of the distributions is also observed as the maximum goes from almost 40% in frequency to less than 20% for all the treatments except for BC130F10 which have a maximum slightly above 20%. This shows that after the integrations, the fibers distribution is more spread on the entire range than before. All the distributions have lengths centered around 2mm.

Figure 24D presents the BC hemp fibers lengths, which increase also after each treatment, but with different lengths and shapes. The fibers lengths seem to increase with the process temperature; the fibers have a length range of 0 to 6mm for the unprocessed fibers, of 0 to 8mm for BC130H10 and of 0 to 12mm for BC160H10, BC190H10 and BC190H5. The distributions of BC130H10, BC190H10 and BC190H5 are close to a normal curve (shape=2.251, 1.988, 1.918, respectively). The distribution of BC160H10 is close to a right-skewed curve (shape=1.405) and with a similar repartition than the unprocessed fibers (shape=1.466). Just as the unprocessed fibers, BC160H10 have lengths centered around 1mm and BC190H10 and BC190H5 around 3mm, even though they have the same length range (0-12mm). BC130H10 has lengths centered around 2mm.

From these observations, the fibers length ranges are clearly increased during plasticization into biocomposites. For the flax fibers, the temperature does not seem to impact the length range, in contrary with the hemp fibers where the longest ones are found when processed at 160 and 190°C. Even though the length ranges are increased, all the flax fibers lengths are centered around 2mm and the hemp fibers vary from 1 to 3mm when looking at centered lengths values. This shows that only a low number of fibers had a length increase in both cases.

As for the diameters, they slightly increase after processing for both flax and hemp fibers, with one exception for the flax fibers with a decrease of diameters when processed at 160°C. The temperature and the time of heating and of treatment do not influence the diameters. As a reminder, BC130 samples were processed during 3min of heating and 10min of treatment and all the other ones during 2min of heating and 5.25min of treatment (**Table 8**). This difference is not translated into the graphs. As example, BC130H10 and BC160H10 have similar diameter distributions but distinct temperature and times of heating and treatment.

As other types of treatments, it was expected that the microwave would degrade the fibers with a decrease in the fiber length and diameter as a result. From this analysis, it can be concluded that the microwave-assisted plasticization does not degrade the fibers (decrease of fibers length and diameter), in the tested conditions. As the fibers lengths and diameters are mostly increased along the plasticization, it could be interesting to measure the fibers after process systematically. As stated before, these parameters are related with the final properties of the biocomposites (Castellani et al., 2016), and will be further demonstrated in **Section IV**. In the majority of the articles describing these parameters in the literature (Avérous et al., 2001; Curvelo et al., 2001; Gironès et al., 2012), only the fibers morphology before treatment is referenced (please confer to the **Section IV**).

Different hypotheses were formulated for the physical interactions between the TPS matrix and the fibers during the plasticization into biocomposite:

a) Formation of a visible TPS matrix layer around the fiber
b) Presence of TPS matrix along the fiber, in a heterogeneous way
c) Absence of TPS matrix or formation of a TPS matrix layer that is not visible

These hypotheses are illustrated in **Figure 25**.

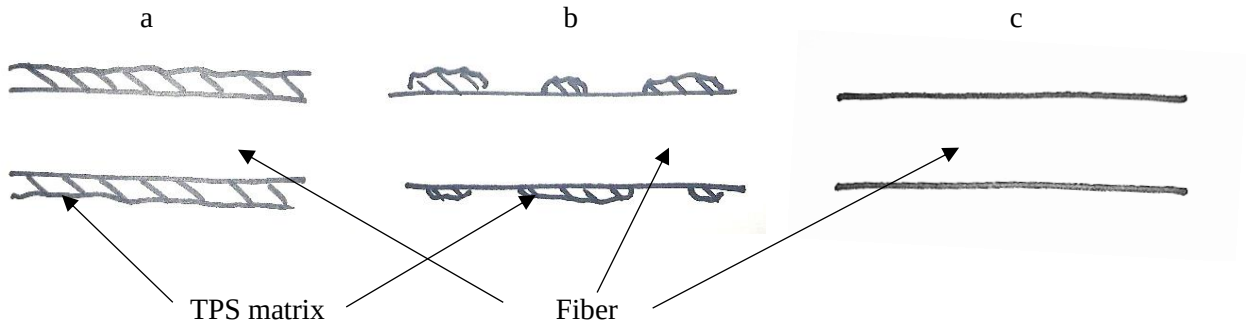


Figure 25: Hypotheses of interactions between the TPS matrix and the fiber during plasticization

Figure 26 depicts examples of pictures were TPS matrix was observed around the isolated fibers with the optical microscope. Different cases are observed, that can be related with the different hypotheses.

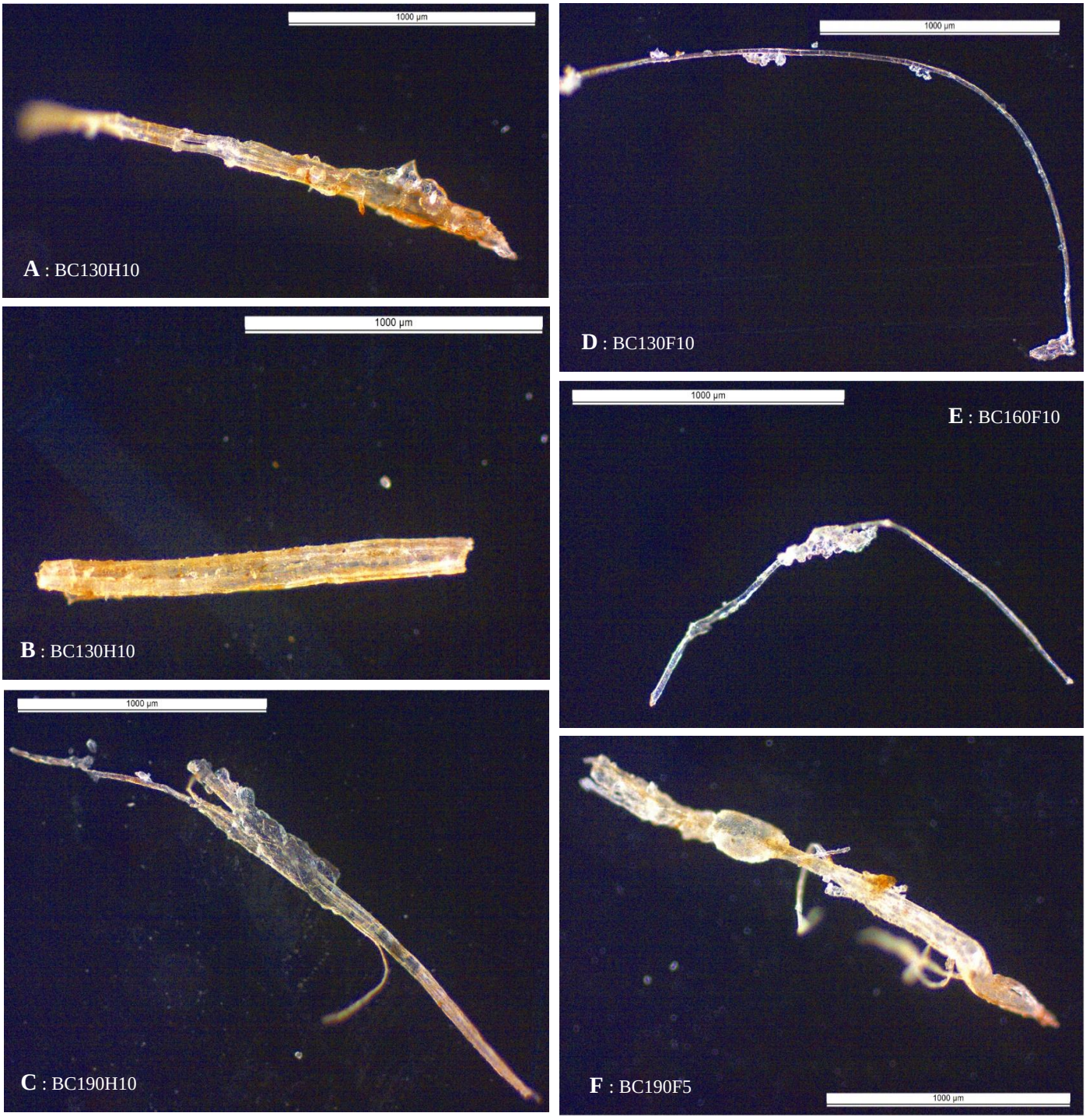


Figure 26: Flax and hemp fibers observed with an optical microscope after plasticization into biocomposites

Two hypotheses could explain **Figure 26B**. The first would be the absence of TPS matrix and the other one would be the presence of a thin layer of TPS matrix that could not be distinguished from the fiber. In the case of **Figure 26C,D and E**, TPS matrix seems to be present in specific areas around the fiber. The last case scenario is for **Figure 26A and F**, where TPS matrix seems to be present almost all around the fiber and the limit between the fiber and the matrix is difficult to determine.

The presence of TPS matrix around the flax and hemp fibers indicates that, when the fibers were pulled out of the BC manually, TPS matrix stayed around them, which could evidence a certain compatibility between the fibers and the matrix.

The increase in diameter observed in **Figure 24** could be explained by the presence of TPS matrix all around the fiber. As the distinction between the fiber and the matrix is not always clear, it is possible that the actual fibers diameters were overestimated. TPS matrix is also present at the surface of fibers heterogeneously. It shows that the fibers could be heterogenous on their surface, in terms of chemical composition, or at least properties. This could be explained by the presence of different polysaccharides profiles at the fiber surface (Morin et al., 2020).

From these observations, it seems that some fibers have more TPS matrix around them than others. It can be reminded that the BC were not completely homogenous, as presented before (**Section III 2.2.**). It may be hypothesized that some fibers were present in a richer TPS matrix environment than others. If this was the case, it could have increased the chance of these fibers having TPS matrix around them.

For the different hypotheses, no trend could be identified in terms of process parameters.

2.4. FTIR analyses

2.4.1. Comparison of the TPS and its initial compounds

The FTIR spectra of TPS130, glycerol and starch were compared and presented in **Figure 27**. For each spectrum, the peaks assignments, based on the literature, were compiled in **Table 11** with the corresponding chemical bonds. Each chemical bond has different vibration types: symmetric or asymmetric stretching, deformation, or scissoring. The vibrations mode depends on the chemical bond position in the molecule and its environment. This can be translated by a shift in the maximum wavenumber of the corresponding peak. The FTIR peaks comparison between TPS, glycerol and starch will help to understand and identify the type of interactions and reactions occurring between the initial compounds during the plasticization.

Table 11: Wavenumbers and bonds associated for TPS sample and initial compounds

Wavenumber (cm^{-1})	Chemical bonds	Starch	Glycerol	TPS130	References
3200-3600	O-H stretching	3200-3600	3200-3600	3200-3600	a, b, c
2930	C-H stretching	2928	2933	2941	a, c
2885	C-H stretching (in CH_2)	2887	2879	2883	c
1650	O-H deformation (of absorbed water)	1637	1651	1643	a, c
1520	C-H or CH_2 deformation	1518	-	-	c
1458	CH_2 symmetric deformation	1458	1454	1456	a, b, c
1415	CH_2 symmetric scissoring	1414	1412	1416	a, b
1330	C-H deformation	1335	1327	1333	b, c
1230	C-O stretching or -O-CH_2-C	1242	1230	1236	b, c
1200	C-O stretching	1205	1209	1211	b
1150	C-O-C asymmetric stretching	1148	-	1153	a
1100	C-C and C-O-C asymmetric stretching	1105	1109	1111	b, c
1080	C-O stretching	1078	-	1074	a, c
1025	C-H deformation + C-C stretching	-	1030	1040	b
1010	C-O stretching	1013	-	-	a
985	C-C stretching	993	993	995	b
925	-OH	928	922	924	b, c
850	C-O-C symmetric stretching	852	852	849	b
765	C-C stretching	762	-	-	b

a : Abdullah et al., 2018, b : Kachel-Jakubowska et al., 2017, c : Oniszczuk et al., 2019

New peaks were not identified in the TPS130 spectrum compared with the initial compounds. This shows that there was no creation of new covalent chemical bonds between the initial compounds during the plasticization. This observation confirms previous research stating that the plasticization is a physical rearrangement where the plasticizers (glycerol and water) penetrate the starch structure and form hydrogen bonds (Prabhu and Prashantha, 2018). These interactions can be identified in the 3200-3600cm^{-1} band. There is no major difference for this band before and after plasticization. The peaks around 2930 and 2885cm^{-1} are more defined after plasticization and confirm the inter- and intra-molecular bonding between glycerol and starch (Stagner et al., 2011).

Almost all the characteristic peaks present on the starch and glycerol spectra are found in the TPS130 spectrum. The fact that some peaks were not found (1518, 1013 and 762cm^{-1}) or shifted (2941,1153 and 1040cm^{-1}) in the TPS130 spectrum could be explained by a rearrangement of the different compounds. As glycerol and water are inserted inside the starch structure, new non-covalent hydrogen bonds are created, and the overall structure is rigidified. Some chemical bonds are thus not able to move as before. All the other bonds are less than 4cm^{-1} different than the initial compounds and are considered equivalent.

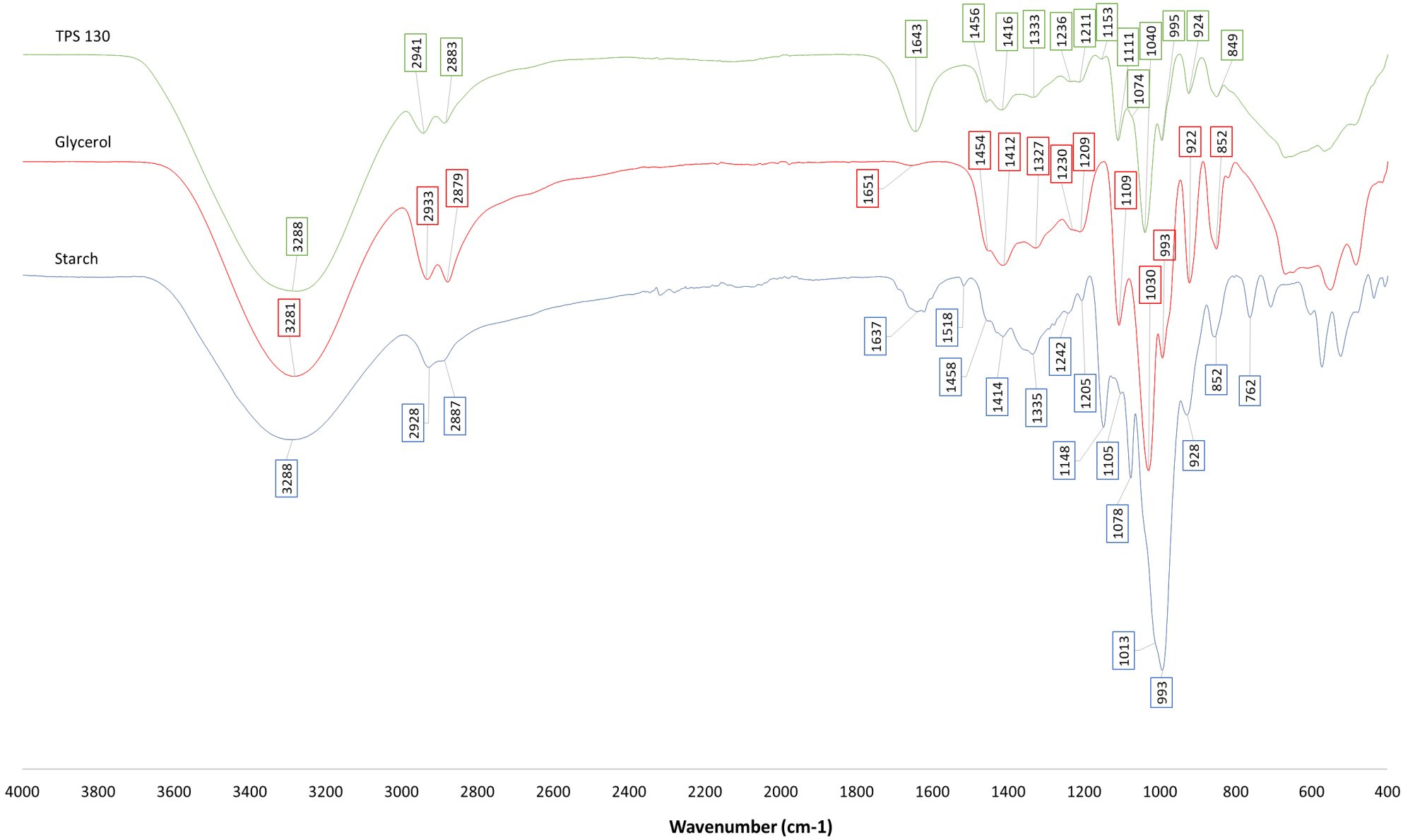


Figure 27: FTIR analysis of a TPS sample and its initial compound

2.4.2. Correlation between process conditions and FTIR spectra

The biocomposites and their corresponding TPS were analyzed with the code presented in appendix 5. The parameters studied were the process temperature, the fiber percentage , the time of heating and the time of treatment, such as defined in the materials and methods (**Section III. 1.3.1**). This code aims to detect correlation between FTIR peaks and process conditions.

For each transformation operated by the code (i.e. normalization, 1st and 2nd derivation), no significant correlation was observed between the FTIR spectra and the studied parameters (R^2 between 0.1 and 0.6). This means that all the studied samples (optimal TPS and corresponding BC) were similar in terms of FTIR spectra for the studied parameters. This shows that the presence of fibers in the biocomposite was not detected by this analysis.

From these results, fibers were isolated from the biocomposites and analyzed by FTIR. These isolated fibers and corresponding biocomposites were also analyzed with the same code and the same parameters. No significant correlation was observed either. This means that all the studied samples (optimal TPS, corresponding BC and isolated fibers) are similar in terms of FTIR spectra for the studied parameters.

It could have been expected that the presence and/or the amount of fibers had an impact on the chemical bonds found in the sample. For example, with the appearance of specific peaks corresponding to cellulose, lignin or hemicelluloses.

The code analysis shows that the fibers' characteristic peaks were not detected and confirms the presence of TPS matrix around them (as seen in **Section III. 2.3**). The ATR-FTIR analysis is a surface analysis and only the outer layer of the sample is analyzed. In this case, it is the TPS matrix around the fibers that was analyzed.

The fiber addition did not seem to have caused chemical modification of the TPS matrix, at the studied surfaces.

These observations show that the FTIR analysis is not adequate to study the fibers modifications occurring during the plasticization.

Analysis of the isolated fibers confirms the presence of TPS matrix around them and supports the good compatibility between them. It could be interesting to study this compatibility with another processing method, such as extrusion, to understand if the microwave-assisted process is more appropriate for this parameter or not.

3. Conclusion and perspectives

In this first part, formulation of thermoplastic starch and biocomposites using microwave-assisted plasticization was studied.

Thermoplastic starch samples were first produced. The lowest tested percentage of starch (20% (w/w)) was considered the most optimal regarding the selection criteria (consistency, color, homogeneity). The highest temperature tested (190°C) gave samples with the highest ability to be molded.

Biocomposites were produced with different natural lignocellulosic fibers (flax, hemp, microcrystalline cellulose) and with the most optimal thermoplastic starch as the matrix. The best results in terms of homogeneity and ability to be molded, were obtained at the highest temperature (190°C) and with the lowest percentages of fibers (5%). Biocomposites made with flax and hemp fibers were similar. The use of microcrystalline cellulose fibers gave better results regarding the same criteria, probably explained by their smaller particles size.

The analysis of these samples through FTIR and optical microscopy revealed the presence of thermoplastic starch matrix surrounding the fibers indicating a good compatibility between these components. No sign of chemical degradation was observed in the matrix and the fibers dimensions were not physically degraded by the process.

Considering all these results, this new processing method seems promising. The homogenization of the matrix and the fibers through this method is the limiting factor and has to be improved.

Other tests could be performed on a larger scale with the addition of an external homogenization system such as suggested before.

The samples produced could be analyzed through tensile tests to evaluate their performances in terms of mechanical properties compared to other processing methods such as extrusion.

Analyses of the fibers surface could help understand their compatibility with the matrix. In fact, fibers surface components are implicated in the interface interactions and will probably influence the strength of the physical or chemical link between the initial components.

IV. Part 2: Modeling of the biocomposites mechanical properties

1. Context

Thermoplastic starch (TPS) have been studied since the start of the 1990's (Ollett et al., 1991; Wiedmann and Strobel, 1991). A few years later, biocomposites made from plasticized starch and natural fibers started to be studied as well (Aichholzer, 1995; Bledzki and Gassan, 1999). For 30 years, research focused on the improvement of these materials properties by numerous means, such as explained in **Section I. 2.3.3.** As this topic has been widely studied, it appears interesting to collect data from literature to create models that could explain the factors influencing the final biocomposites' mechanical properties and by extent (after empirical model validation) predict the mechanical properties of new biocomposites made from starch and natural fibers. Such models could help orientate the research by providing a decisional tool identifying the parameters that have a major impact on the mechanical properties. It could also save time and financial resources as the prediction could reduce the number of tests to be performed, depending on the mechanical properties wanted.

2. Materials and methods

2.1. Database creation

The literature of TPS, biocomposites made from starch and natural fibers and raw natural fibers alone was reviewed to create a database. Only studies using distilled water and/or glycerol as plasticizers for the TPS formulation were selected as they are common plasticizers and were used to formulate biocomposites previously in this work. The data collection also focused on research where the materials were studied through tensile tests. The response measurements were the Young's modulus (YM, in MPa), the tensile strength (TS, in MPa) and the elongation at break (EaB, in %). They are commonly discussed in the literature and give a good representation of the materials' mechanical properties. As explained in **Section I. 3.1.**, the YM represents the ability of a material to be deformed reversibly, the TS represents the maximum strength a sample can bear before breaking and the EaB represents the maximum deformation a material can withstand before breaking. These measurements can allow a prediction of the material's application.

This database was created with input and output variables. The input variables are the formulation and process parameters such as :

- The chemical composition of the fiber (% DW (w/w) cellulose, hemicellulose, lignin) and starch (% amylose, amylopectin)
- The percentages of the TPS matrix compounds (% (w/w) starch, water, glycerol)
- The percentage of fibers expressed as the percentage of the total biocomposite mass

- The fiber length (in mm)
- The temperature (in °C)
- RH (in %)
- Rotor speed (in RPM)

For more clarity, all the variables corresponding to a percentage will be written with the variable name followed by %, such as, starch% or fibers%.

The RH is the relative humidity set during the aging of the sample. The rotor speed was set to 0 when no rotor was used in the process. The output variables are the tensile tests' measurements (YM, TS and EaB).

Many studies did not provide the chemical composition of the fibers and starch used. To complete the database, the percentages of cellulose, hemicellulose and lignin for the fiber and the percentages of amylose and amylopectin for starch were added with standard values for the corresponding initial compounds on the basis of the botanical origin.

The main database created with 72 references was composed of 551 lines. After the missing information concerning the input variables were completed as explained, the lines where information was not found or could not be replaced by standard values, were not considered. This was the case when the speed or the fiber length were not mentioned, for example. In the end, 61 studies with a total of 477 lines referenced all the input variables and at least one output variable. These references are presented in appendix 6.

This database was divided in three, one database for each output variable in order to predict each of the measurements separately. The YM database was built from 50 references out of the 61, the TS database from 57 references and the EaB database from 56 references.

Each of these databases was also used to produce two other ones, one containing only the TPS samples and one with only the biocomposites samples. The goal was to know if it was better to consider the TPS matrix together with the BC samples or if they were different enough that these materials should be studied separately. In total, nine databases were created and tested to create the best model possible for each output variable. Each database is described in **Table 12,** where the length, the input and output variables as well as their validity range are mentioned.

For the TPS+BC database, all the variables referring to the fiber were set to 0 for the TPS samples. When only natural fibers were studied, all the variables referring to TPS were set to 0.

Table 12: Description of the different databases: mean value and validity range of each variable. NA = not applicable

Input variables	Young's modulus TPS	BC	TPS + BC
	Mean (Min-Max)	Mean (Min-Max)	Mean (Min-Max)
Cellulose%	NA	70.47 (0.00 - 100.00)	35.14 (0.00 - 100.00)
Hemicellulose%	NA	9.08 (0.00 - 89.90)	4.53 (0.00 - 89.90)
Lignin%	NA	8.81 (0.00 - 48.40)	4.40 (0.00 - 48.40)
Fiber%	NA	20.16 (0.30 - 100.00)	10.06 (0.00 - 100.00)
Fiber length (mm)	NA	5.55 (0.00 - 150.60)	2.77 (0.00 - 150.60)
Starch%	57.68 (2.88 - 86.96)	55.73 (0.00 - 75.00)	56.71 (0.00 - 86.96)
Amylose%	38.86 (0.00 - 87.00)	19.76 (0.00 - 28.00)	29.33 (0.00 - 87.00)
Amylopectin%	60.84 (13.00 - 100.00)	69.32 (0.00 - 100.00)	65.07 (0.00 - 100.00)
Glycerol%	19.26 (0.00 - 60.00)	20.43 (0.00 - 50.00)	19.84 (0.00 - 60.00)
Water%	22.24 (0.00 - 96.10)	13.91 (0.00 - 96.10)	18.09 (0.00 - 96.10)
Temperature (°C)	139.90 (85.00 - 180.00)	119.50 (0.00 - 200.00)	129.70 (0.00 - 200.00)
Relative humidity (%)	55.76 (7.00 - 95.00)	50.83 (7.00 - 83.00)	53.30 (7.00 - 95.00)
Rotor speed (RPM)	163.20 (0.00 - 2000.00)	200.70 (0.00 - 2000.00)	181.90 (0.00 - 2000.00)
Young's modulus (MPa)	229.62 (0.12 - 3204.00)	907.76 (0.50 - 15000.00)	567.80 (0.12 - 15000.00)
Database length	192	191	383
Variables	**Tensile strength** **TPS**	**BC**	**TPS + BC**
	Mean (Min-Max)	Mean (Min-Max)	Mean (Min-Max)
Cellulose%	NA	44.70 (0.00 - 100.00)	36.43 (0.00 - 100.00)
Hemicellulose%	NA	8.98 (0.00 - 89.90)	4.57 (0.00 - 89.90)
Lignin%	NA	4.25 (0.00 - 48.40)	4.34 (0.00 - 48.40)
Fiber%	NA	18.52 (0.30 - 100.00)	9.41 (0.00 - 100.00)
Fiber length (mm)	NA	5.41 (0.00 - 150.60)	2.75 (0.00 - 150.60)
Starch%	58.43 (2.88 - 95.00)	59.47 (0.00 - 95.00)	58.96 (0.00 - 95.00)
Amylose%	37.31 (0.00 - 87.00)	20.18 (0.00 - 28.00)	28.61 (0.00 - 87.00)
Amylopectin%	62.32 (13.00 - 100.00)	70.16 (0.00 - 100.00)	66.30 (0.00 - 100.00)
Glycerol%	18.11 (0.00 - 50.00)	19.41 (0.00 - 50.00)	18.77 (0.00 - 50.00)
Water%	22.72 (0.00 - 97.00)	12.51 (0.00 - 96.10)	17.54 (0.00 - 97.00)
Temperature (°C)	139.40 (25.00 - 180.00)	126.30 (25.00 - 175.00)	131.90 (0.00 - 200.00)
Relative humidity (%)	53.78 (7.00 - 90.00)	52.82 (7.00 - 83.00)	53.29 (7.00 - 90.00)
Rotor speed (RPM)	171.70 (0.00 - 2000.00)	175.10 (0.00 - 2000.00)	173.40 (0.00 - 2000.00)
Tensile strength (MPa)	7.86 (0.10 - 53.50)	17.04 (0.18 - 550.00)	12.53 (0.10 - 550.00)
Database length	211	218	429
Variables	**Elongation at break** **TPS**	**BC**	**TPS + BC**
	Mean (Min-Max)	Mean (Min-Max)	Mean (Min-Max)
Cellulose%	NA	73.20 (6.70 - 100.00)	29.68 (0.00 - 100.00)
Hemicellulose%	NA	9.23 (0.00 - 89.90)	3.74 (0.00 - 89.90)
Lignin%	NA	9.78 (0.00 - 48.40)	3.97 (0.00 - 48.40)
Fiber%	NA	15.46 (0.30 - 100.00)	6.27 (0.00 - 100.00)
Fiber length (mm)	NA	4.99 (0.00 - 150.60)	2.02 (0.00 - 150.60)
Starch%	56.69 (2.88 - 86.96)	61.16 (0.00 - 75.00)	58.50 (0.00 - 86.96)
Amylose%	32.37 (3.64 - 87.00)	21.80 (0.00 - 28.00)	30.46 (0.00 - 87.00)
Amylopectin%	63.31 (13.00 - 95.00)	72.72 (0.00 - 95.00)	67.12 (0.00 - 95.00)
Glycerol%	21.97 (0.00 - 60.00)	23.36 (0.00 - 50.00)	22.53 (0.00 - 60.00)
Water%	20.70 (0.00 - 97.00)	11.38 (0.00 - 96.10)	16.92 (0.00 - 97.00)
Temperature (°C)	134.70 (25.00 - 180.00)	119.30 (0.00 - 170.00)	128.50 (0.00 - 180.00)
Relative humidity (%)	53.11 (7.00 - 95.00)	50.90 (7.00 - 83.00)	52.21 (7.00 - 95.00)
Rotor speed (RPM)	152.20 (0.00 - 2000.00)	120.80 (0.00 - 2000.00)	139.50 (0.00 - 2000.00)
Elongation at break (%)	58.02 (0.10 - 751.40)	22.50 (0.88 - 105.81)	43.62 (0.10 - 751.40)
Database length	242	165	407

2.2. Models design

For each database, different models were built through a code developed by an LBTV staff member using the RStudio software, linked to Anaconda. For each model, the code (presented in appendix 7) is composed of different steps:

1) Reading and description of the database (**Table 12**)
2) Correlation analysis between the input variables (**Table 14**)
3) Random separation of the database in two: 70% to build the model (called training set) and 30% for the self-validation (to verify the accuracy of the model on new data, called testing set)
4) Design of different models based on transformations such as represented in **Table 13**

Different data transformations were performed by applying a mathematical function on the output variable. The goal is to transform the relationship between the input and output variables into a linear regression as the code evaluate the quality of a linear regression.

Table 13: Data transformation and corresponding equations tested

Data transformation	Equation
Linear	$y = I + \alpha * x_1 + \beta * x_2 + \cdots + \delta * x_n + \gamma\,(x_1 * x_2) + \cdots + \sigma\,(x_{n-1} * x_n)$
Logarithm	$\log(y) = I + \alpha * x_1 + \beta * x_2 + \cdots + \delta * x_n + \gamma\,(x_1 * x_2) + \cdots + \sigma\,(x_{n-1} * x_n)$
Square root	$\sqrt{y} = I + \alpha * x_1 + \beta * x_2 + \cdots + \delta * x_n + \gamma\,(x_1 * x_2) + \cdots + \sigma\,(x_{n-1} * x_n)$
Inverse	$\frac{1}{y} = I + \alpha * x_1 + \beta * x_2 + \cdots + \delta * x_n + \gamma\,(x_1 * x_2) + \cdots + \sigma\,(x_{n-1} * x_n)$

For each model, the output variable (y) is built from the input variables; independent variables (x_i), and their simple interactions ($x_i * x_j$), which are calculated by the code. The code performs a multilinear regression by varying the coefficient (α, β, …) of each input variable and the intercept (I) to obtain the model with the smallest sum of squared errors between the model and the observations. A coefficient of determination (adjusted R^2) is provided for each model that represents how much the variability of the model is explained by the variables and their interactions. In this work, it was decided that the model with the highest R^2 would be the one selected between the different mathematical transformations.

5) Variable regression coefficients

For each database, the regression coefficients (α, β, …) that were determined for the selected models are computed with their level of significance (see one example in **Table 16**).

6) Variable selection

Once the model is built and that the regression coefficients and their signification are available, it is usually observed that some variables are not significant, especially when interactions are added. A selection of variables can be performed to have a model with only significant regression coefficients (meaning they are significantly different than 0). Only the variables that participate in the model are considered and can be studied.

In this work, the backward stepwise selection method was used. The first variable that is taken out of the model is the one with the highest coefficient p-value, meaning the least significant. Once it is taken out, a new model is created with new coefficients and corresponding significance. From this new model, the new least significant variable is taken out of the model, just as before. A new model is created, and the process continued until only significant coefficients are left. The signification of the p-value is determined manually and in this case p-values < 0.05 were significant.

3. Results and discussion

3.1. Correlation between the input variables

The correlation matrix between the input variables can help understand relationships between these variables within the database. Collinearity may happen when two variables are correlated, meaning there is a linear relationship between them. If more than two variables are correlated, they are considered multicollinear (JMP Statistical Discovery, n.d.).

Each correlation can be explained in different ways, direct correlation such as the two variables are linked by an equation or correlation caused by a bias in the database. A bias can be caused by a correlation randomly present in nature, or when only some combinations of values are present in the database.

If two correlated variables are included in the regression model it can have non-negligible impacts on the model. Many coefficients can be non-significant even if the R^2 of the model is high. The regression coefficients can also have poor estimation or be estimated in the wrong direction. Their standard error can also be higher. Usually, removing one of the correlated variables out of the model can help solving these issues (JMP Statistical Discovery, n.d.).

The correlation matrix between input variables for the YM using the TPS+BC database is presented in **Table 14**.

Table 14: Correlation between the input variables for one database (TPS+BC for YM) as example. Values > |0.4| are in bold.

	Cellulose %	Hemicellulose %	Lignin %	Fiber %	Fiber length (mm)	Starch %	Amylose %	Amylopectin %	Glycerol %	Water %	Temperature (°C)	Relative Humidity (%)	Speed (RPM)
Cellulose %													
Hemicellulose %	0.1119												
Lignin %	0.1810	**0.4723**											
Fiber %	0.3673	0.1625	**0.6074**										
Fiber length (mm)	0.0980	0.0048	**0.4296**	**0.5518**									
Starch %	0.1954	0.0015	-0.2523	-0.2710	-0.2099								
Amylose %	-0.3310	-0.1518	-0.2765	-0.3290	-0.1633	0.0956							
Amylopectin %	0.2620	0.1477	-0.1229	-0.3022	-0.2644	0.1779	**-0.7490**						
Glycerol %	0.0088	0.1516	0.0341	-0.1890	-0.0985	0.3267	-0.1021	0.2567					
Water %	-0.1702	-0.0613	-0.0822	-0.1372	-0.0958	**-0.7590**	0.1047	-0.0414	**-0.6053**				
Temperature (°C)	-0.2657	0.0454	-0.2028	**-0.5051**	-0.3391	**0.4376**	**0.4323**	-0.0505	0.2640	-0.2087			
Relative Humidity (%)	0.1141	-0.2046	-0.1668	-0.0615	-0.0196	-0.1128	0.1540	-0.1309	-0.1496	0.1872	-0.0375		
Speed (RPM)	0.0037	-0.0685	-0.0977	-0.0707	-0.0410	0.1373	-0.0878	0.1140	0.0118	-0.0951	-0.0741	-0.0198	

The correlation coefficient between the percentage of amylose and the one of amylopectin is -0.75 in **Table 14**. This is explained by the fact that the sum of these percentages is equal to 100%. As this correlation is a direct correlation, one of them needs to be taken out of the model. In this case, amylose was removed before creating the models.

In **Table 14**, hemicellulose and lignin have a coefficient of correlation of 0.47. This makes sense as cellulose, hemicellulose and lignin are the main components of fibers. It means that these three variables are somewhat dependent. However, not all combinations are represented in nature and in the database. As example, lignin does not usually exceed 40% and some studied fibers are only made of cellulose. It can thus be interesting to include all of them in the model as they do not have the same ranges of values (see **Table 12**) and have different interactions with the other variables. The importance of these interactions can also be different. This is why these three variables were kept in the model. The same explanation can be used for the composition of the TPS matrix (starch, glycerol, and water%).

Other correlations can be highlighted such as the correlations between the fiber percentage and lignin (0.61) or between the fiber percentage and the fiber length (0.55). These are not direct correlations as no equation can link these variables.

The correlation matrices for the other databases were reviewed but not presented. No other direct correlation was detected.

3.2. Models design and selection

For each database, the model selected and their corresponding adjusted R^2, RMSE (root mean square error) and NRMSE (normalized root mean square error) are represented in **Table 15**. The RMSE represents the difference between the values estimated by the model and the observed values. It can be used to compare the accuracy of different models for one database but not between different databases, as it is dependent on the database size and range. To compare models from different databases, the NRMSE can be used, which is the RMSE divided by the mean value of the output variable observations in the database used (Otto, 2019). NRMSE is presented as a percentage in the table. The different results were compiled before and after the variables selection step.

Table 15: Models selected, before and after variables selection with their corresponding R^2, RMSE and NRMSE

		YM		TS		EaB	
		Before selection	After selection	Before selection	After selection	Before selection	After selection
TPS+BC	Selected model	Linear	Linear	Linear	Linear	Inverse	Log
	R^2	0.9807	0.9629	0.9818	0.9655	0.6223	0.6495
	RMSE	539.74	309.58	15.38	6.89	102.75	5114144.02
	NRMSE	95.06%	54.52%	122.82%	54.99%	235.56%	11724310%
BC	Selected model	Square root	-	Linear	-	Square root	-
	R^2	0.9703		0.9887		0.8167	
	RMSE	5870175.66		85779.19		23591268.44	
	NRMSE	646666%		503399.03%		104850100%	
TPS	Selected model	Linear	Linear	Square root	Log	Square root	Square root
	R^2	0.8366	0.7342	0.7200	0.7267	0.6050	0.5573
	RMSE	385.49	211.78	10.34	512.67	46.16	50.38
	NRMSE	167.89%	92.23%	131.58%	6522.52%	79.57%	86.83%

The models using the BC database could not be calculated after variables selection because of a collinearity detection that could not be solved.

3.2.1. Models comparison before variables selection

For the YM and TS, the adjusted R^2 of the TPS+BC and BC database are greater than 0.97. It means that a great majority of the model variability is explained by the variables studied. For the same output variables, the TPS database presents a smaller R^2. From these observations, it seems that when only the TPS samples were considered, the addition of other variables or interactions could help to explain a larger part of the variability. Concerning the EaB, the largest R^2 was obtained using the BC database and was equal to 0.82. It shows that for this measurement, variables or interactions are missing to explain the entire model variability.

If we compare the NRMSE values, they are considerably higher for the BC database than the other databases. For YM and TS, the smallest NRMSE are obtained using the TPS+BC database while for EaB it was with the TPS database. It seems that for the YM and TS, adding the TPS observations to the BC observations is reducing the model error without changing much the R^2. For the EaB, the error is reduced when TPS observations are added but also the R^2. It seems less evident to explain in this case. It could mean that TPS and BC have different behaviors regarding this measurement and should not be studied together. Important variables could be missing to explain the variability in the TPS database. Also, it is possible that the mathematical transformation used were not suitable to explain the relationship between EaB and the input variables.

Regarding the models with smaller R^2, some hypotheses in terms of missing variables can be formulated. For example, Angellier et al., (2006) showed that the TPS time of aging before tensile tests influences the tensile test responses. The process duration, the crystallinity of starch and the temperature of aging were also studied by a few references (Torres et al., 2007; Van Soest and Borger, 1997) and showed

they had an impact on the mechanical properties of TPS and/or biocomposites. These different parameters were only referenced a few times and thus could not be added to construct the model.

It is also possible that the mathematical transformations or equations used were not adequate and that other transformations could have been more suitable to represent the relationship between the input and output variables.

3.2.2. Models comparison after variable selection

After the variable selection, four out of six models have the same type of mathematical transformation for the models selected. It is the case for the YM using TPS+BC and TPS databases with the linear model, for the TS using the TPS+BC database also with the linear model and for the EaB using the TPS database with the square root model. For the three first cases, the NRMSE were reduced after variables selection and for the last case it was slightly increased. The R^2 were also slightly reduced, when using the TPS+BC database for YM and TS and for the EaB using the TPS database. The R^2 decrease was more important when using the TPS database for YM.

In the two other cases, different types of models were selected, the R^2 were slightly higher but the RMSE were significantly increased. For these cases, the method of variables selection seems less appropriate when the highest R^2 is used to select the best model.

Using the highest R^2 as the model selection parameter can thus be discussed as the model with the highest R^2 does not imply that it has the smallest RMSE, thus the smallest error. For example, before the variables selection, the linear model selected for TS using the TPS+BC database had an R^2 of 0.9818 and a RMSE of 15.39. It means that the model can be well explained by the selected variables but have a large error associated with it. The logarithmic model, also for TS using the same database, had an R^2 of 0.8434 and a RMSE of 6.72, having therefore a better predictive power. This shows that the model quality and accuracy does not depend only on the R^2.

The multilinear regression method can be used for predictive modeling or for explanatory modeling. In the case of predictive modeling, the goal is to predict the output variable the most accurately possible. In explanatory modeling, the focus is to understand which input variable will have the most impact on the output variable and how they will impact it, by analyzing their regression coefficient. Depending on the objective, there must be a compromise between the R^2 and the RMSE to determine which model is more adequate. A high R^2 will show that the selected variables explain a large variability when a small RMSE will translate a small error of the model and a higher accuracy in the prediction. If the model is accurate enough, it can be used to predict the output variable of new samples with a certain error associated with it. The input variables of the new samples must be in the validity range of the model.

Regarding the RMSE obtained in this work, these models will not be used to predict the mechanical properties of new samples. However, the models produced with the highest R^2 can be used as explanatory models. The regression coefficients of one of them will be studied in the next section.

3.3. Regression coefficients analysis

The regression coefficients of the YM using the TPS+BC database are analyzed, regarding the high R^2 associated. The regression coefficients after variables selection are presented in **Table 16**.

Before the variable selection, only 8 coefficients were significant whereas 38 are significant after the selection. All the variables selected participate to the model.

All the non-interaction variables selected are related to the composition of the biocomposite. The process variables do not seem to impact the YM when considered alone but they do when they are studied with a variable related to the composition, such as the interaction between hemicellulose and speed or fiber length and temperature.

Some regression coefficients signs of input variables seem logical such as the fiber percentage that has a positive coefficient, meaning the YM increases with the fiber percentage when all the other variables are fixed. This makes sense as the fiber is used to reinforce the biocomposite and in results the more the fibers, the larger YM in the studied range (Avérous et al., 2001; Gironès et al., 2012).

All the variables related to the TPS composition (starch, glycerol and water%) have a negative coefficient, meaning that for each of these variables, YM decreases when they are increased individually. This is more difficult to explain as these three variables present many interactions with other parameters. The interactions between these variables two by two can be analyzed. It seems that it is better to process starch with water (positive coefficient) than with glycerol (negative coefficient), this is confirmed with the interaction between glycerol and water (negative coefficient). These variables also interact with variables related to the fiber (e.g. Lignin*Starch%, Fiber%*%Glycerol) and also with the process variables (e.g. Glycerol%*Speed, Water%*RH). From all these interactions, it seems that the variables selected should be imperatively studied together when producing biocomposites.

From this analysis, it can be concluded that many variables and their interactions are part of the presented model. The standardization of these coefficients could help understand which ones have the most impact on the final response to be able to focus on these variables or interactions.

Table 16: Regression coefficient of the YM after variables selection. The significance was determined with the P-value as follow : 0 < * < 0.001 < ** < 0.01< * < 0.05**

Young's modulus		
TPS+BC		
Intercept	2793.90	***
Hemicellulose%	-523.04	***
Fiber%	142.49	***
Fiber_length	-1371.60	***
Starch%	-23.80	***
Glycerol%	-56.45	*
Water%	-38.28	*
Cellulose%*Hemicellulose%	0.43	*
Cellulose%*Fiber_length	14.09	***
Cellulose%*Starch%	-0.20	*
Cellulose%*Amylopectin%	0.19	*
Hemicellulose%*Fiber_length	12.17	***
Hemicellulose%*Amylopectin%	1.10	*
Hemicellulose%*RH	6.63	***
Hemicellulose%*Speed	1.34	***
Lignin%*Fiber_length	19.74	***
Lignin%*Starch%	5.74	***
Lignin%*Amylopectin%	-3.37	***
Lignin%*Glycerol%	5.93	***
Lignin%*Water%	5.05	***
Lignin%*RH	-6.23	***
Fiber%*Starch%	-1.05	***
Fiber%*Glycerol%	-2.16	***
Fiber%*Water%	-1.23	***
Fiber_length*Water%	25.13	***
Fiber_length*Temperature	-0.44	***
Starch%*Amylopectin%	0.42	***
Starch%*Glycerol%	-0.63	*
Starch%*Water%	0.32	*
Starch%*Speed	0.04	***
Amylopectin%*Water%	0.19	*
Amylopectin%*RH	-0.54	***
Amylopectin%*Speed	-0.04	***
Glycerol%*Water%	-1.04	*
Glycerol%*RH	1.51	***
Glycerol%*Speed	0.02	*
Water%*RH	0.41	*
Temperature*RH	-0.07	*

3.4. Model improvement

There are two main ways to improve the quality of the different models, by improving the data collection to build the database itself and by changing or adding some mathematical and statistical operations in the code.

3.4.1. Database creation

The data collection can be improved in many ways to ensure a database closer to the reality and to help reducing the error of the models.

The collection of more data seems evident. The larger the dataset, the larger part of the population is represented.

In numerous papers, the data were presented in graphs and the values were estimated as best as possible. This adds an error to the model. Also, only the mean and standard deviation are usually presented. This means that the variability coming from the samples is not considered in the presented model. Sharing the data for each repetition in a table, in supplementary data for example, could solve this problem. This would, on one hand, increase the size of the database and on the other hand, include the variability of each sample. In the end, only the error coming from the method itself would remain.

As explained before, the addition of other input variables and their interactions, such as other process parameters or aging parameters, could help to explain a larger part of the variability of the model and thus increase the R^2.

Many of these additional parameters are only available in a small number of references which raises an important issue: much information is not communicated through scientific papers. As explained, when the composition of the fiber or starch was not communicated, it was manually added with standard values for these compounds. For the fiber composition, the cellulose, hemicellulose, and lignin percentages had missing values representing 30, 38 and 34% of the original database, respectively. For the amylose and amylopectin percentages, the missing values represented 29% of the original database. These missing values were replaced by standard values to avoid reducing the database by a third.

This analysis clearly showed an impact of the fiber composition on the final composite properties. While this effect is expected (as different types of fibers are added within the TPS matrix in the literature), the missing data were surprising.

3.4.2. Code improvement

As presented in **Table 13**, mathematical transformations were applied on the output variable "y" such as "f(y) = x", with x representing all the input variables. Other transformations could be applied such as "$y^2 = x$" or "$\exp(y) = x$" among others.

Mathematical transformations can also be applied on the input instead of the output variables such as "y = f(x)". The transformation of each input variable and the determination of the most suitable transformation for each one of them is done independently. This method requires more power of calculation as many combinations have to be tested. To understand the relationship between each input variable and the output variable and how it can be optimized, each input variable can be plotted with the output variable separately. For example, as seen in **Figure 28**, depending on the relationship between the input and output variable, the transformations and optimizations will be different.

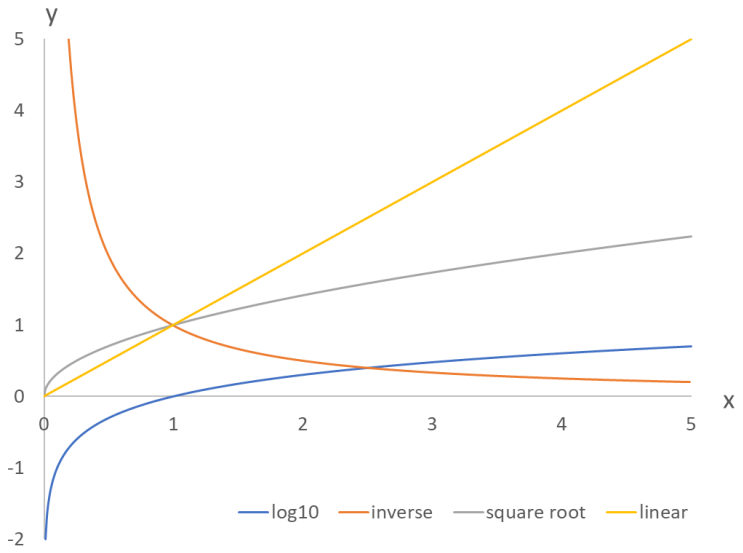


Figure 28: Examples of mathematical functions

As for the validation of the model, many other validation methods exist, and all have strengths and weaknesses.

The model can also be improved by detecting outliers, which are observations that seem to deviate from the others in a dataset. This kind of data can bias the accuracy of the model and different techniques exist to differentiate them from the rest of the data (Santoyo, 2017).

The presence of outliers can from a problem in the database (encoding error). It can also happen that the model is only accurate on a certain range. Before or after a certain threshold for one or several variables, the response can be different, and the observations are far from the others. This type of detection can help improve the model by removing or modifying encoding errors or to select a certain range of values where their respond in the same way.

In this work, simple interactions between the input variables were added to build the models. It could be interesting to add more interactions such as 3 or 4-ways interactions but it can complexify the models. The more interactions the more tedious the selection of variables is.

Many methods of variables selection exist. In this work the backwards method was used but forward or stepwise methods also exist for example. The forward method adds significant variables one at a time until the addition of a new variable is not significant. The stepwise method is a combination of addition and elimination of variables (Choueiry, 2020).

Once the variables are selected and the model is improved as much as possible, the coefficient of each variable can be analyzed. Before comparing these coefficients, they can be standardized by dividing them by the mean or the range of each variable in the database. In this way, each coefficient is equivalent in terms of weight and by comparing them, we can identify the variables that have the most impact on the response.

After the improvement of the model by these different methods, the error coming from the variability of the different methods of measurement (for input and output variables) will still remain.

4. Conclusion and perspectives

In this second part, models were created to explain the mechanical properties of TPS and biocomposites by studying which parameters and interactions are impacting the tensile tests' measurements (Young modulus, tensile strength, elongation at break).

The parameters studied were related to the material's composition and the process parameters. Different databases were created for each measurement, by reviewing the literature. Several mathematical transformations were applied on the response variables and multilinear regressions were performed on each one of them. A selection of variables was performed to obtain models with only significant regression coefficients. The highest R^2 was used to determine the best model for each database.

It appears that higher R^2 were reached for the models created for the Young's modulus and the tensile strength. The variable selection helped decrease the error (RMSE) of most models and also decreased the number of variables included in the model. For the Young's modulus using the TPS+BC database, the selected regression coefficients were provided. Only variables related to the materials compositions were selected alone. These variables as well as variables related to the process parameters were selected as interactions. This analysis indicated that all the studied parameters were important to the model at some point and should be well referenced as it is not always the case. More parameters could be added to explain a larger part of the variability, especially when the R^2 were smaller.

Various ways of improvement have been proposed, by improving the data collection but also by improving the code. The application of these improvements could help to create models with a decreased error, and a higher variability explained. These models could be used to predict the tensile tests' measurements of new samples. Other models could also be generated in order to explain or predict other properties such as gas barrier properties, biodegradability or any other property that can be measured.

V. References

Abdullah, A.H.D., Chalimah, S., Primadona, I., Hanantyo, M.H.G., 2018. Physical and chemical properties of corn, cassava, and potato starchs. IOP Conf. Ser. Earth Environ. Sci. 160. https://doi.org/10.1088/1755-1315/160/1/012003

Aichholzer, W., 1995. 14. Stuttgarter Kunststoff-Kolloquium. Stuttgart.

Altayan, M.M., Al Darouich, T., Karabet, F., 2017. On the Plasticization Process of Potato Starch: Preparation and Characterization. Food Biophys. 12, 397–403. https://doi.org/10.1007/s11483-017-9495-2

American Chemisty Council, n.d. Lifecycle of a Plastic Product [WWW Document]. URL https://plastics.americanchemistry.com/Lifecycle-of-a-Plastic-Product/ (accessed 4.6.20).

Angellier, H., Molina-Boisseau, S., Dole, P., Dufresne, A., 2006. Thermoplastic starch - Waxy maize starch nanocrystals nanocomposites. Biomacromolecules 7, 531–539. https://doi.org/10.1021/bm050797s

Averous, L., Boquillon, N., 2004. Biocomposites based on plasticized starch: Thermal and mechanical behaviours. Carbohydr. Polym. 56, 111–122. https://doi.org/10.1016/j.carbpol.2003.11.015

Avérous, L., Fringant, C., Moro, L., 2001. Plasticized starch-cellulose interactions in polysaccharide composites. Polymer (Guildf). 42, 6565–6572. https://doi.org/10.1016/S0032-3861(01)00125-2

Balakrishnan, P., John, M.J., Pothen, L., Sreekala, M.S., Thomas, S., 2016. Natural fibre and polymer matrix composites and their applications in aerospace engineering, Advanced Composite Materials for Aerospace Engineering. Elsevier Ltd. https://doi.org/10.1016/b978-0-08-100037-3.00012-2

Bîrcă, A., Gherasim, O., Grumezescu, V., Grumezescu, A.M., 2019. Introduction in thermoplastic and thermosetting polymers, in: Materials for Biomedical Engineering: Thermoset and Themoplastic Polymers. pp. 1–28. https://doi.org/10.1016/b978-0-12-816874-5.00001-3

Bledzki, A.., Gassan, J., 1999. Composites reinforced with cellulose based fibers. Prog. Polym. Sci 24, 221–274.

Carvalho, A.J.F., 2013. Starch: Major Sources, Properties and Applications as Thermoplastic Materials., Handbook of Biopolymers and Biodegradable Plastics. Elsevier. https://doi.org/10.1016/B978-1-4557-2834-3.00007-0

Castellani, R., Di Giuseppe, E., Beaugrand, J., Dobosz, S., Berzin, F., Vergnes, B., Budtova, T., 2016. Lignocellulosic fiber breakage in a molten polymer. Part 1. Qualitative analysis using rheo-optical observations. Compos. Part A Appl. Sci. Manuf. 91, 229–237. https://doi.org/10.1016/j.compositesa.2016.10.015

Choueiry, G., 2020. Understand Forward and Backward Stepwise Regression [WWW Document]. URL https://quantifyinghealth.com/stepwise-selection/ (accessed 8.9.20).

Cosucra Socode, n.d. Native pea starch: native Nastar feed [WWW Document]. URL http://www.socode-warcoing.be/en/other-products/native-pea-starch (accessed 7.11.20).

Crawford, C.B., Quinn, B., 2017a. Chap 1. The emergence of plastics, in: Microplastic Pollutants. Elsevier Inc., pp. 1–17. https://doi.org/10.1016/B978-0-12-809406-8.00001-3

Crawford, C.B., Quinn, B., 2017b. Chap 3. Plastic production, waste and legislation, in: Microplastic Pollutants. pp. 39–56. https://doi.org/10.1016/b978-0-12-809406-8.00003-7

Curvelo, A.A.S., De Carvalho, A.J.F., Agnelli, J.A.M., 2001. Thermoplastic starch-cellulosic fibers composites: Preliminary results. Carbohydr. Polym. 45, 183–188. https://doi.org/10.1016/S0144-8617(00)00314-3

Engineering ToolBox, 2008. Dielectric Constants of Liquids [WWW Document]. URL https://www.engineeringtoolbox.com/liquid-dielectric-constants-d_1263.html (accessed 4.17.20).

ETIP Bioenergy, 2020. Starch crops for production of biofuels [WWW Document]. URL http://www.etipbioenergy.eu/value-chains/feedstocks/agriculture/starch-crops (accessed 6.5.20).

European Bioplastics, 2019. What are bioplastics ? [WWW Document]. URL https://www.european-bioplastics.org/bioplastics/ (accessed 3.31.20).

European Commission, 2019. European Platform on Life Cycle Assessment (LCA) [WWW Document]. URL https://ec.europa.eu/environment/ipp/lca.htm (accessed 4.16.20).

Forssell, P.M., Mikkilä, J.M., Moates, G.K., Parker, R., 1997. Phase and glass transition behaviour of concentrated barley starch-glycerol-water mixtures, a model for thermoplastic starch. Carbohydr. Polym. 34, 275–282. https://doi.org/10.1016/s0144-8617(97)00133-1

Fraser-Reid, B.O., Kuniaki, T., Thiem, J. (Eds.), 2008. Glycoscience, Chemistry and Chemical Biology. Springer-Verlag Berlin Heidelberg New York 2008. https://doi.org/10-1007/978-3-540-30429-6

Gassan, J., Gutowski, V.S., Bledzki, A.K., 2000. About the surface characteristics of natural fibres. Macromol. Mater. Eng. 283, 132–139. https://doi.org/10.1002/1439-2054(20001101)283

Genkina, N.K., Wikman, J., Bertoft, E., Yuryev, V.P., 2007. Effects of structural imperfection on gelatinization characteristics of amylopectin starches with A- and B-type crystallinity. Biomacromolecules 8, 2329–2335. https://doi.org/10.1021/bm070349f

Gironès, J., López, J.P., Mutjé, P., Carvalho, A.J.F., Curvelo, A.A.S., Vilaseca, F., 2012. Natural fiber-reinforced thermoplastic starch composites obtained by melt processing. Compos. Sci. Technol. 72, 858–863. https://doi.org/10.1016/j.compscitech.2012.02.019

Janssen, L.P.B.M., Moscicki, L., 2010. Thermoplastic Starch: A Green Material for Various Industries. Wiley-VCH. https://doi.org/10.1002/9783527628216

Jawaid, M., Sapuan, S.M., Alotman, O.Y., 2017. Green Biocomposites Manufacturing and Properties. https://doi.org/10.1007/978-3-319-46610-1

JMP Statistical Discovery, n.d. Multicollinearity [WWW Document]. URL https://www.jmp.com/en_us/statistics-knowledge-portal/what-is-multiple-regression/multicollinearity.html (accessed 8.8.20).

Kachel-Jakubowska, M., Matwijczuk, A., Gagoś, M., 2017. Analysis of the physicochemical properties of post-manufacturing waste derived from production of methyl esters from rapeseed oil. Int. Agrophysics 31, 175–182. https://doi.org/10.1515/intag-2016-0042

Kozasowski, R.M., Mackiewicz-Talarczyk, M., Allam, A.M., 2012. Bast fibres: flax, Handbook of Natural Fibres. Woodhead Publishing Limited. https://doi.org/10.1533/9780857095503.1.56

Leeson, C., 2016. A plastic ocean. UK - Hong Kong.

Leroy, E., Decaen, P., Jacquet, P., Coativy, G., Pontoire, B., Reguerre, A.L., Lourdin, D., 2012. Deep eutectic solvents as functional additives for starch based plastics. Green Chem. 14, 3063–3066. https://doi.org/10.1039/c2gc36107h

López, O. V., García, M.A., 2012. Starch films from a novel (Pachyrhizus ahipa) and conventional sources: Development and characterization. Mater. Sci. Eng. C 32, 1931–1940. https://doi.org/10.1016/j.msec.2012.05.035

Ma, X., Chang, P.R., Yu, J., 2008. Properties of biodegradable thermoplastic pea starch/carboxymethyl cellulose and pea starch/microcrystalline cellulose composites. Carbohydr. Polym. 72, 369–375. https://doi.org/10.1016/j.carbpol.2007.09.002

Malumba, P., Janas, S., Deroanne, C., Masimango, T., Béra, F., 2011. Structure de l'amidon de maïs et principaux phénomènes impliqués dans sa modification thermique. Biotechnol. Agron. Société Environ. 15, 315–326.

Mathes, V., 2018. The composites industry: plenty of opportunities in heterogeneous market. Reinf. Plast. 62, 44–51. https://doi.org/10.1016/j.repl.2017.05.002

Menendez, J.A., Arenillas, A., Fidalgo, B., Fernandez, Y., Zubizarreta, L., Calvo, E.G., 2010. Microwave heating processes involving carbon materials _ Elsevier Enhanced Reader.pdf. Fuel Process. Technol. 91, 1–8.

Mohanty, A.K., Misra, M., Drzal, L.T., 2002. Sustainable Bio-Composites from renewable resources: Opportunities and challenges in the green materials world. J. Polym. Environ. 10, 19–26. https://doi.org/10.1023/A:1021013921916

Morin, S., Bockstal, L., Jacquet, N., Richel, A., 2019. One-step enzymatic grafting of ferulic acid with cellulose to enhance matrices–fibres compatibility in bio-composites. J. Mater. Sci. 54, 13314–13321. https://doi.org/10.1007/s10853-019-03832-x

Morin, S., Lecart, B., Istasse, T., Bailly Maître Grand, C., Meddeb-Mouelhi, F., Beauregard, M., Richel, A., 2020. Effect of a low melting temperature mixture on the surface properties of lignocellulosic flax bast fibers. Int. J. Biol. Macromol. 148, 851–856. https://doi.org/10.1016/j.ijbiomac.2020.01.232

Nafchi, A.M., Moradpour, M., Saeidi, M., Alias, A.K., 2013. Thermoplastic starches: Properties, challenges, and prospects. Starch/Staerke 65, 61–72. https://doi.org/10.1002/star.201200201

Ollett, A.L., Parker, R., Smith, A.C., 1991. Deformation and fracture behaviour of wheat starch plasticized with glucose and water. J. Mater. Sci. 26, 1351–1356. https://doi.org/10.1007/BF00544476

Oniszczuk, T., Combrzyński, M., Matwijczuk, A., Oniszczuk, A., Gładyszewska, B., Podleśny, J., Czernel, G., Karcz, D., Niemczynowicz, A., Wójtowicz, A., 2019. Physical assessment, spectroscopic and chemometric analysis of starch-based foils with selected functional additives. PLoS One 14, 1–19. https://doi.org/10.1371/journal.pone.0212070

Otto, S.A., 2019. HOW TO NORMALIZE THE RMSE [WWW Document]. URL https://www.marinedatascience.co/blog/2019/01/07/normalizing-the-rmse/ (accessed 8.8.20).

PerkinElmer, 2005. FT-IR Spectroscopy Attenuated Total Reflectance (ATR).

Pervaiz, M., Panthapulakkal, S., KC, B., Sain, M., Tjong, J., 2016. Emerging Trends in Automotive Lightweighting through Novel Composite Materials. Mater. Sci. Appl. 07, 26–38. https://doi.org/10.4236/msa.2016.71004

PlasticsEurope, The European Association of Plastics Recycling and Recovery Organisations (EPRO), PlasticsEurope's Market Research and Statistics Group (PEMRG), Conversio Market & Strategy GmbH, 2019. Plastics - the Facts 2019.

Ponomarev, D., Rodier, E., Sauceau, M., Nikitine, C., Mizonov, V., Fages, J., 2012. Modelling non-homogeneous flow and residence time distribution in a single-screw extruder by means of Markov chains. J. Math. Chem. 50, 2141–2154. https://doi.org/10.1007/s10910-012-0022-x

Prabhu, T.N., Prashantha, K., 2018. A Review on Present Status and Future Challenges of Starch Based Polymer Films and Their Composites in Food Packaging Applications. Polym. Compos. https://doi.org/10.1002/pc

Prachayawarakorn, J., Hommanee, L., Phosee, D., Chairapaksatien, P., 2010. Property improvement of thermoplastic mung bean starch using cotton fiber and low-density polyethylene. Starch/Staerke 62, 435–443.

Punia, S., Dhull, S.B., Sandhu, K.S., Kaur, M., 2019. Faba bean (Vicia faba) starch: Structure, properties, and in vitro digestibility—A review. Legum. Sci. 1:e18. https://doi.org/10.1002/leg3.18

Ratnayake, W.S., Hoover, R., Warkentin, T., 2002. Pea starch: Composition, structure and properties - A review. Starch/Staerke 54, 217–234. https://doi.org/10.1002/1521-379X(200206)54:6<217::AID-STAR217>3.0.CO;2-R

Ravi, M., Dubey, R.R., Shome, A., Guha, S., Anil Kumar, C., 2018. Effect of surface treatment on Natural fibers composite. IOP Conf. Ser. Mater. Sci. Eng. 376. https://doi.org/10.1088/1757-899X/376/1/012053

Riley, A., 2012. Basics of polymer chemistry for packaging materials, in: Packaging Technology. https://doi.org/10.1533/9780857095701.2.262

Sabetzadeh, M., Bagheri, R., Masoomi, M., 2012. Effect of Corn Starch Content in Thermoplastic Starch/Low-Density Polyethylene Blends on Their Mechanical and Flow Properties. J. Appl. Polym. Sci. 126, E63–E69. https://doi.org/DOI 10.1002/app.36329

Santoyo, S., 2017. A Brief Overview of Outlier Detection Techniques [WWW Document]. URL

https://towardsdatascience.com/a-brief-overview-of-outlier-detection-techniques-1e0b2c19e561 (accessed 8.10.20).

Sharma, S.K., Verma, D.S., Khan, L.U., Kumar, S., Khan, S.B., 2018. Handbook of Materials Characterization, Handbook of Materials Characterization. https://doi.org/10.1007/978-3-319-92955-2

Shrestha, A.K., Halley, P.J., 2014. Starch Modification to Develop Novel Starch-Biopolymer Blends: State of Art and Perspectives, Starch Polymers: From Genetic Engineering to Green Applications. Elsevier B.V. https://doi.org/10.1016/B978-0-444-53730-0.00022-1

Shrivastava, A., 2018a. Chap 1. Introduction to Plastics Engineering, in: Introduction to Plastics Engineering. pp. 1–16. https://doi.org/10.1016/b978-0-323-39500-7.00001-0

Shrivastava, A., 2018b. Chap 7. Environmental Aspects of Plastics. Introd. to Plast. Eng. 207–232. https://doi.org/10.1016/b978-0-323-39500-7.00007-1

Shrivastava, A., 2018c. Chap 3. Plastic Properties and Testing, in: Introduction to Plastics Engineering. pp. 49–110. https://doi.org/10.1016/b978-0-323-39500-7.00003-4

Stagner, J., Alves, V.D., Narayan, R., Beleia, A., 2011. Thermoplasticization of High Amylose Starch by Chemical Modification Using Reactive Extrusion. J. Polym. Environ. 19, 589–597. https://doi.org/10.1007/s10924-011-0307-3

Staiger, M.P., Tucker, N., 2008. 8 - Natural-fibre composites in structural applications, in: K.L. Pickering (Ed.), Properties and Performance of Natural-Fibre Composites. Woodhead Publishing, pp. 269–300. https://doi.org/10.1016/B978-1-84569-267-4.50008-9

Thielen, M., 2014. Bioplastics - Plants and crops raw materials products. Fachagentur Nachwachsende Rohstoffe e.V. (FNR) Agency for Renewable Resources.

Thunwall, M., Kuthanová, V., Boldizar, A., Rigdahl, M., 2008. Film blowing of thermoplastic starch. Carbohydr. Polym. 71, 583–590. https://doi.org/10.1016/j.carbpol.2007.07.001

Torres, F.G., Arroyo, O.H., Gomez, C., 2007. Processing and Mechanical Properties of Natural Fiber Reinforced Thermoplastic Starch Biocomposites. J. Thermoplast. Compos. Mater. 20. https://doi.org/10.1177/0892705707073945

Van Soest, J.J.G., Borger, D.B., 1997. Structure and properties of compression-molded thermoplastic starch materials from normal and high-amylose maize starches. J. Appl. Polym. Sci. 64, 631–644. https://doi.org/10.1002/(SICI)1097-4628(19970425)64:4<631::AID-APP2>3.0.CO;2-O

Verma, D., Sharma, S., 2017. Green biocomposites for structural applications, in: Jawaid, M., Salit, M.S., Alothman, O.Y. (Eds.), Green Biocomposites Design and Applications. Springer, pp. 167–191. https://doi.org/10.1007/978-3-319-49382-4_1

Wang, H., Chu, P.K., 2013. Chapter 4 – Surface Characterization of Biomaterials, in: Characterization of Biomaterials. Elsevier Inc.

Wang, Q., Xiao, S., Shi, S.Q., Cai, L., 2018. Effect of light-delignification on mechanical, hydrophobic, and thermal properties of high-strength molded fiber materials. Sci. Rep. 8, 2–11. https://doi.org/10.1038/s41598-018-19623-4

Wiederhorn, S.M., Fields, R.J., Low, S., Bahng, G.-W., Wehrstedt, A., Hahn, J., 2006. Chap 7. Mechanical properties, in: Czichos, H., Saito, T., Smith, L. (Eds.), Springer Handbook of Materials Measurement Methods.

Wiedmann, W., Strobel, E., 1991. Compounding of Thermoplastic Starch with Twin-screw Extruders. Starch = Stärke 43, 138–145. https://doi.org/https://doi.org/10.1002/star.19910430404

Work, W.J., Horie, K., Hess, M., Stepto, R.F.T., 2004. Definitions of terms related to polymer blends, composites, and multiphase polymeric materials (IUPAC Recommendations 2004). Pure Appl. Chem. 76, 1985–2007. https://doi.org/10.1351/pac200476111985

Yalcin, D., 2016. Effect of Specimen Geometry on Tensile Testing Results [WWW Document]. URL https://www.admet.com/effect-specimen-geometry-tensile-testing-results/ (accessed 3.31.20).

Zakaria, N.H., Muhammad, N., Abdullah, M.M.A.B., 2017. Potential of Starch Nanocomposites for Biomedical Applications. IOP Conf. Ser. Mater. Sci. Eng. 209. https://doi.org/10.1088/1757-899X/209/1/012087

Zhang, Y., Han, J.H., 2006. Mechanical and thermal characteristics of pea starch films plasticized with monosaccharides and polyols. J. Food Sci. 71. https://doi.org/10.1111/j.1365-2621.2006.tb08891.x

Zhang, Y., Liu, Z., Han, J., 2008. Starch-based edible films, in: Chiellini, E. (Ed.), Environmentally Compatible Food Packaging. pp. 108–136. https://doi.org/10.1016/B978-1-84569-194-3.50005-2

Zhang, Y., Rempel, C., Liu, Q., 2014a. Thermoplastic Starch Processing and Characteristics-A Review. Crit. Rev. Food Sci. Nutr. 54, 1353–1370. https://doi.org/10.1080/10408398.2011.636156

Zhang, Y., Rempel, C., McLaren, D., 2014b. Thermoplastic Starch, in: Innovations in Food Packaging: Second Edition. Elsevier Ltd, pp. 391–412. https://doi.org/10.1016/B978-0-12-394601-0.00016-3

VI. Appendixes

1) Process parameters of TPS made from glycerol and/or water from the literature

References	Plant source	Amylose/ amylopectin ratio	Compounds percentages (Glycerol/Water/ Starch)	Process	Temperature	Speed	Processing Time
(Forssell et al., 1997)	Barley	28/72	14/30/56 20/25/55 29/12/59 39/11.5/49.5	Melt mixing	170°C	80 rpm	15 min
(Angellier et al., 2006)	Waxy maize	<1 % amylose	2.7/86.4 /10.9 3.4/86.4/10.2 4/86.4/9.5	Heating in a reactor under pressure with stirring	150°C	ND	10 min
(Ma et al., 2008)	Pea	35/65	23/0/73	Single screw plastic extruder	120-140°C	20 rpm	ND
(Thunwall et al., 2008)	Hydroxy-propylated starch	ND	6/17/77	Extrusion	ND	ND	ND
(Prachayawarakorn et al., 2010)	Mung bean (Vigna radiata)	30/62	50/0/50	Internal mixer	140°C	50 rpm	5 min
(Altayan et al., 2017)	Potato	ND	30/20/50	Twin counter-rotating internal mixer	160°C	60 rpm	7 min

2) Scoring grid of TPS formulated in the microwave with water and glycerol

Sample	Consistency	Browning/ Color	Color homogeneity	Air bubbles	Final score
TPS 1	2	2	2	0	6
TPS 2	0	1	0	2	3
TPS 3	2	1	1	1	5
TPS 4	2	1	1	1	5
TPS 5	2	2	1	1	6
TPS 6	2	2	2	2	8
TPS 7	0	0	0	2	2
TPS 8	1	1	1	1	4
TPS 9	1	1	1	1	4
TPS 10	2	2	2	2	8
TPS 11	0	0	0	2	2
TPS 12	1	1	1	1	4
TPS 13	1	1	1	1	4
TPS 14	2	1	1	2	6
TPS 15	3	1	2	2	8
TPS 16	0	0	0	2	2
TPS 17	3	1	2	2	8
TPS 18	0	1	0	1	2
TPS 19	3	1	2	2	8
TPS 20	0	0	0	2	2
TPS 21	2	1	0	1	4

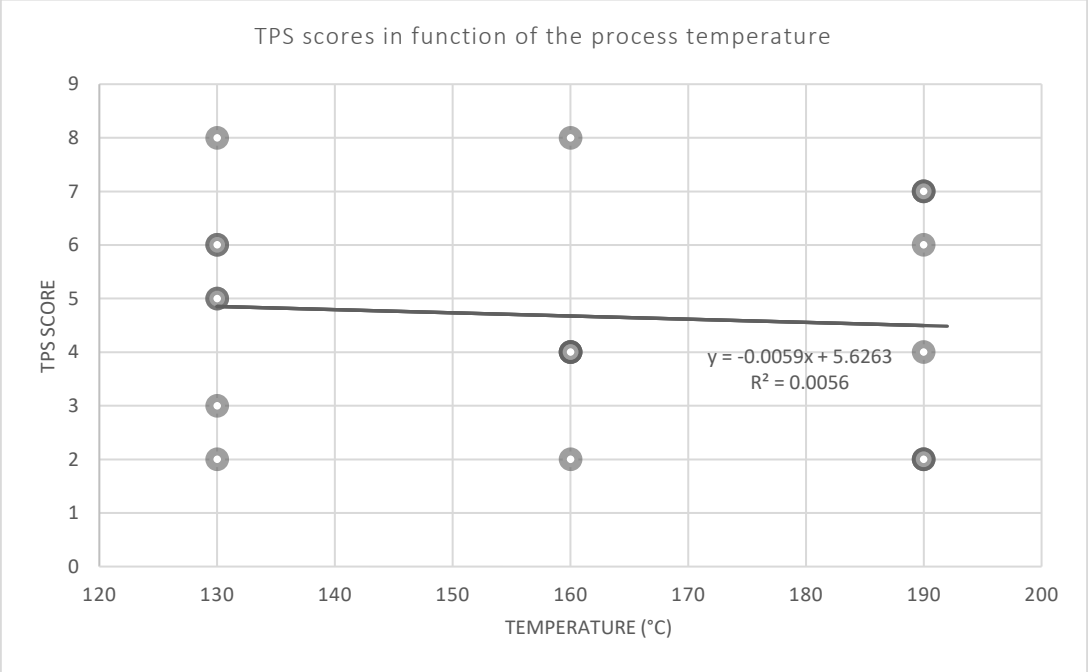
Green = optimal, orange = promising, red = not optimal

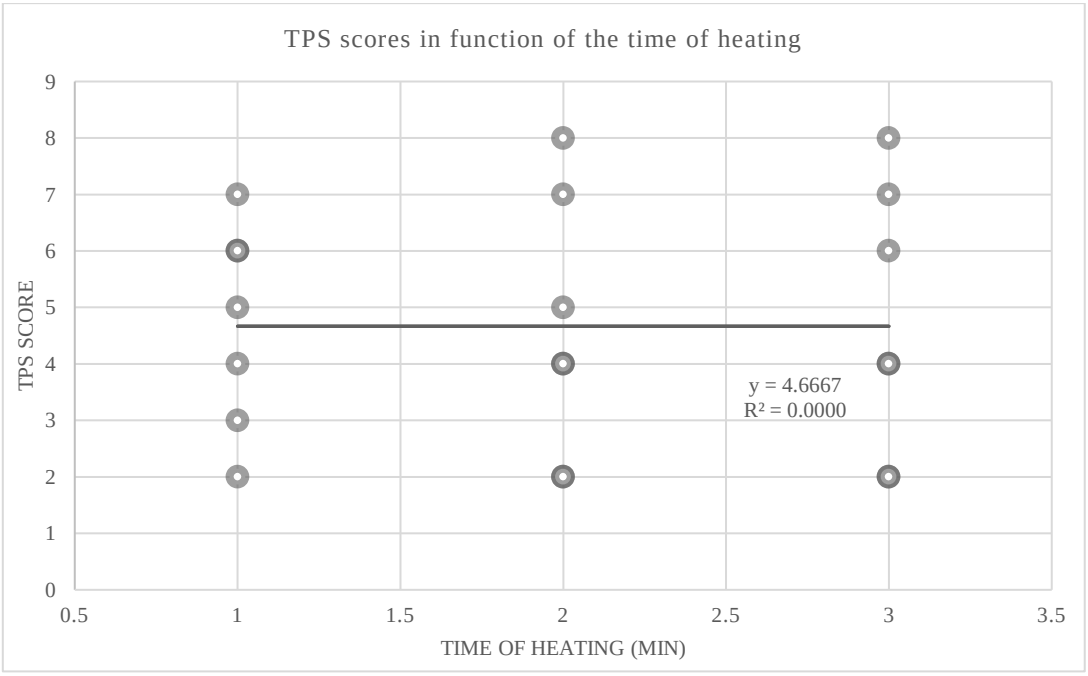
3) Scoring grid of biocomposites formulated in the microwave

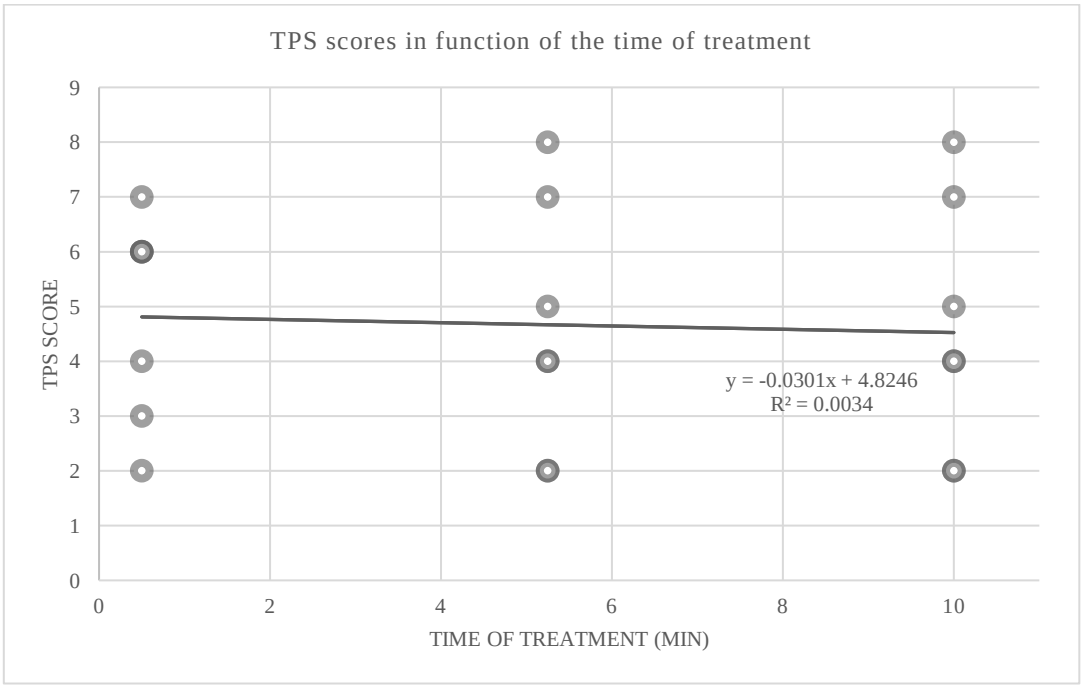
Sample	Consistency	Color homogeneity	Air bubbles	Final decision
BC130C10	1	0	0	1
BC130F10	1	0	1	2
BC130H10	1	0	1	2
BC160C10	1	1	0	2
BC160F10	1	1	1	3
BC160H10	1	1	1	3
BC190C10	2	1	0	3
BC190F10	2	1	1	4
BC190H10	2	1	1	4
BC190C5	3	2	2	7
BC190F5	3	1	1	5
BC190H5	3	2	1	6

Green = optimal, orange = promising, red = not optimal

4) Relationship between the times of heating and of treatment as well as the process temperature with the TPS scores




TPS scores in function of the time of heating
TPS SCORE
TIME OF HEATING (MIN)
y = 4.6667
R² = 0.0000
0
1
2
3
4
5
6
7
8
9
0.5
1
1.5
2
2.5
3
3.5


TPS scores in function of the time of treatment
TPS SCORE
TIME OF TREATMENT (MIN)
y = -0.0301x + 4.8246
R² = 0.0034
0
1
2
3
4
5
6
7
8
9
0
2
4
6
8
10

5) Code used for the FTIR analysis

```
#########################################################################################################################
#### Auteur: Brieuc Lecart, PhD student in Uliège lecart.brieuc@gmail.com ###############################################
#### Testeurs: Sophie Morin ; Lionel Dumoulin ###########################################################################
#### Rédigé par adaptation de mon travail de fin d'étude: "Caractérisation et prédiction de la composition ######
#### biochimique de la biomasse de miscanthus pour un débouché en bioéthanol 2G" ########################################
#### Réalisé à l'INRA d'Estrées-Mons en 2012. Script initial réalisé avec l'aide de l'INRA d'Orléans ############
#########################################################################################################################
setwd('C:/Users/Mathilde/dox/EXCOM700 Biocomposites fibres-TPS/Interpretation R/CalibrationR')
options("scipen"=100)
install.packages("pls")
library(pls)
source('calib_ok.r')
source('carspls_LOO.r')
donnees_decoupees<-read.table("DonneesBrutes/BC/donnees_decoupees.txt",sep="\t",header=TRUE)
donnees_decoupees_norm<-read.table("DonneesBrutes/BC/donnees_decoupees_norm.txt",sep="\t",header=TRUE)
donnees_decoupees_der1<-read.table("DonneesBrutes/BC/donnees_decoupees_der1.txt",sep="\t",header=TRUE)
donnees_decoupees_der2<-read.table("DonneesBrutes/BC/donnees_decoupees_der2.txt",sep="\t",header=TRUE)
donnees_decoupees_der1_norm<-read.table("DonneesBrutes/BC/donnees_decoupees_der1_norm.txt",sep="\t",header=TRUE)
donnees_decoupees_der2_norm<-read.table("DonneesBrutes/BC/donnees_decoupees_der2_norm.txt",sep="\t",header=TRUE)
donnees_decoupees_norm_der1<-read.table("DonneesBrutes/BC/donnees_decoupees_norm_der1.txt",sep="\t",header=TRUE)
donnees_decoupees_norm_der2<-read.table("DonneesBrutes/BC/donnees_decoupees_norm_der2.txt",sep="\t",header=TRUE)
donnees_decoupees_ok<-data.frame('PP'=as.numeric(donnees_decoupees$PP),'signal'=I(as.matrix(donnees_decoupees[,2:
ncol(donnees_decoupees)])))
donnees_decoupees_norm_ok<-data.frame('PP'=as.numeric(donnees_decoupees_norm$PP),
'signal'=I(as.matrix(donnees_decoupees_norm[,2:ncol(donnees_decoupees_norm)])))
donnees_decoupees_der1_ok<-data.frame('PP'=as.numeric(donnees_decoupees_der1$PP),
'signal'=I(as.matrix(donnees_decoupees_der1[,2:ncol(donnees_decoupees_der1)])))
donnees_decoupees_der2_ok<-data.frame('PP'=as.numeric(donnees_decoupees_der2$PP),
'signal'=I(as.matrix(donnees_decoupees_der2[,2:ncol(donnees_decoupees_der2)])))
donnees_decoupees_der1_norm_ok<-data.frame('PP'=as.numeric(donnees_decoupees_der1_norm$PP),
'signal'=I(as.matrix(donnees_decoupees_der1_norm[,2:ncol(donnees_decoupees_der1_norm)])))
donnees_decoupees_der2_norm_ok<-data.frame('PP'=as.numeric(donnees_decoupees_der2_norm$PP),
'signal'=I(as.matrix(donnees_decoupees_der2_norm[,2:ncol(donnees_decoupees_der2_norm)])))
donnees_decoupees_norm_der1_ok<-data.frame('PP'=as.numeric(donnees_decoupees_norm_der1$PP),
'signal'=I(as.matrix(donnees_decoupees_norm_der1[,2:ncol(donnees_decoupees_norm_der1)])))
donnees_decoupees_norm_der2_ok<-data.frame('PP'=as.numeric(donnees_decoupees_norm_der2$PP),
'signal'=I(as.matrix(donnees_decoupees_norm_der2[,2:ncol(donnees_decoupees_norm_der2)])))
donnees_decoupees_transposees<-data.frame('echantillon'=as.character(substring(colnames
(donnees_decoupees_ok$signal),9)),signal=t(donnees_decoupees_ok$signal))
colnames(donnees_decoupees_transposees$signal)<-donnees_decoupees_ok$PP
donnees_decoupees_norm_transposees<-data.frame('echantillon'=as.character(substring
(colnames(donnees_decoupees_norm_ok$signal),9)),signal=t(donnees_decoupees_norm_ok$signal))
colnames(donnees_decoupees_norm_transposees$signal)<-donnees_decoupees_norm_ok$PP
donnees_decoupees_der1_transposees<-data.frame('echantillon'=as.character(substring
(colnames(donnees_decoupees_der1_ok$signal),9)),signal=t(donnees_decoupees_der1_ok$signal))
colnames(donnees_decoupees_der1_transposees$signal)<-donnees_decoupees_der1_ok$PP
donnees_decoupees_der2_transposees<-data.frame('echantillon'=as.character(substring
(colnames(donnees_decoupees_der2_ok$signal),9)),signal=t(donnees_decoupees_der2_ok$signal))
colnames(donnees_decoupees_der2_transposees$signal)<-donnees_decoupees_der2_ok$PP
donnees_decoupees_der1_norm_transposees<-data.frame('echantillon'=as.character
(substring(colnames(donnees_decoupees_der1_norm_ok$signal),9)),signal=t(donnees_decoupees_der1_norm_ok$signal))
colnames(donnees_decoupees_der1_norm_transposees$signal)<-donnees_decoupees_der1_norm_ok$PP
donnees_decoupees_der2_norm_transposees<-data.frame('echantillon'=as.character
(substring(colnames(donnees_decoupees_der2_norm_ok$signal),9)),signal=t(donnees_decoupees_der2_norm_ok$signal))
colnames(donnees_decoupees_der2_norm_transposees$signal)<-donnees_decoupees_der2_norm_ok$PP
donnees_decoupees_norm_der1_transposees<-data.frame('echantillon'=as.character
(substring(colnames(donnees_decoupees_norm_der1_ok$signal),9)),signal=t(donnees_decoupees_norm_der1_ok$signal))
colnames(donnees_decoupees_norm_der1_transposees$signal)<-donnees_decoupees_norm_der1_ok$PP
donnees_decoupees_norm_der2_transposees<-data.frame('echantillon'=as.character
(substring(colnames(donnees_decoupees_norm_der2_ok$signal),9)),signal=t(donnees_decoupees_norm_der2_ok$signal))
colnames(donnees_decoupees_norm_der2_transposees$signal)<-donnees_decoupees_norm_der2_ok$PP
donnees_ref_ok<-read.table("DonneesBrutes/BC/Parametres_BC.txt",sep='\t',header=TRUE)
donnees_calibration<-merge(donnees_ref_ok,donnees_decoupees_transposees,by='echantillon')
donnees_calibration_norm<-merge(donnees_ref_ok,donnees_decoupees_norm_transposees,by='echantillon')
donnees_calibration_der1<-merge(donnees_ref_ok,donnees_decoupees_der1_transposees,by='echantillon')
donnees_calibration_der2<-merge(donnees_ref_ok,donnees_decoupees_der2_transposees,by='echantillon')
donnees_calibration_der1_norm<-merge(donnees_ref_ok,donnees_decoupees_der1_norm_transposees,by='echantillon')
donnees_calibration_der2_norm<-merge(donnees_ref_ok,donnees_decoupees_der2_norm_transposees,by='echantillon')
donnees_calibration_norm_der1<-merge(donnees_ref_ok,donnees_decoupees_norm_der1_transposees,by='echantillon')
donnees_calibration_norm_der2<-merge(donnees_ref_ok,donnees_decoupees_norm_der2_transposees,by='echantillon')
write.table(donnees_calibration,file="Resultats/BC/donnees_decoupees_calib.txt",sep="\t",row.names=FALSE,
col.names=TRUE)
write.table(donnees_calibration_norm,file="Resultats/BC/donnees_decoupees_norm_calib.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(donnees_calibration_der1,file="Resultats/BC/donnees_decoupees_der1_calib.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(donnees_calibration_der2,file="Resultats/BC/donnees_decoupees_der2_calib.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(donnees_calibration_der1_norm,file="Resultats/BC/donnees_decoupees_der1_norm_calib.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(donnees_calibration_der2_norm,file="Resultats/BC/donnees_decoupees_der2_norm_calib.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(donnees_calibration_norm_der1,file="Resultats/BC/donnees_decoupees_norm_der1_calib.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
```

```
write.table(donnees_calibration_norm_der2,file="Resultats/BC/donnees_decoupees_norm_der2_calib.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
chem='fiber.percentage'
calib_data<-drop_outliers_LOO(data_frame=donnees_calibration,trait=chem,maxcomp=20,threshold=0.001,
criterion=1,maxsteps=100)
calib_data_norm<-drop_outliers_LOO(data_frame=donnees_calibration_norm,trait=chem,maxcomp=19,
threshold=0.001,criterion=1,maxsteps=100)
calib_data_der1<-drop_outliers_LOO(data_frame=donnees_calibration_der1,trait=chem,maxcomp=20,
threshold=0.001,criterion=1,maxsteps=100)
calib_data_der2<-drop_outliers_LOO(data_frame=donnees_calibration_der2,trait=chem,maxcomp=20,
threshold=0.001,criterion=1,maxsteps=100)
calib_data_der1_norm<-drop_outliers_LOO(data_frame=donnees_calibration_der1_norm,trait=chem,
maxcomp=20,threshold=0.001,criterion=1,maxsteps=100)
calib_data_der2_norm<-drop_outliers_LOO(data_frame=donnees_calibration_der2_norm,trait=chem,
maxcomp=20,threshold=0.001,criterion=1,maxsteps=100)
calib_data_norm_der1<-drop_outliers_LOO(data_frame=donnees_calibration_norm_der1,trait=chem,
maxcomp=20,threshold=0.001,criterion=1,maxsteps=100)
calib_data_norm_der2<-drop_outliers_LOO(data_frame=donnees_calibration_norm_der2,trait=chem,
maxcomp=20,threshold=0.001,criterion=1,maxsteps=100)
outliers<-list('data'=calib_data$outliers,'data_norm'=calib_data_norm$outliers,'data_der1'
=calib_data_der1$outliers,'data_der2'=calib_data_der2$outliers,
    'data_der1_norm'=calib_data_der1_norm$outliers,'data_der2_norm'=calib_data_der2_norm$outliers,
    'data_norm_der1'=calib_data_norm_der1$outliers,'data_norm_der2'=calib_data_norm_der2$outliers)
print(outliers)
sapply(outliers,length)
sink("Resultats/BC/outliers_decoupees.txt")
print(outliers)
sink()
data_calib_filt<-calib_data$data
data_calib_norm_filt<-calib_data_norm$data
data_calib_der1_filt<-calib_data_der1$data
data_calib_der2_filt<-calib_data_der2$data
data_calib_der1_norm_filt<-calib_data_der1_norm$data
data_calib_der2_norm_filt<-calib_data_der2_norm$data
data_calib_norm_der1_filt<-calib_data_norm_der1$data
data_calib_norm_der2_filt<-calib_data_norm_der2$data
write.table(data_calib_filt,file="Resultats/BC/donnees_decoupees_calib_filt.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(data_calib_norm_filt,file="Resultats/BC/donnees_decoupees_norm_calib_filt.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der1_filt,file="Resultats/BC/donnees_decoupees_der1_calib_filt.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der2_filt,file="Resultats/BC/donnees_decoupees_der2_calib_filt.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der1_norm_filt,file="Resultats/BC/donnees_decoupees_der1_norm_calib_filt.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der2_norm_filt,file="Resultats/BC/donnees_decoupees_der2_norm_calib_filt.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_norm_der1_filt,file="Resultats/BC/donnees_decoupees_norm_der1_calib_filt.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_norm_der2_filt,file="Resultats/BC/donnees_decoupees_norm_der2_calib_filt.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
calib_data_ok<-calib_valid_LOO(data_set=data_calib_filt,trait='trait',maxcomp=10,ratio=3/4,criterion=1)
calib_data_norm_ok<-calib_valid_LOO(data_set=data_calib_norm_filt,trait='trait',maxcomp=10,ratio=3/4,criterion=1)
calib_data_der1_ok<-calib_valid_LOO(data_set=data_calib_der1_filt,trait='trait',maxcomp=10,ratio=3/4,criterion=1)
calib_data_der2_ok<-calib_valid_LOO(data_set=data_calib_der2_filt,trait='trait',maxcomp=10,ratio=3/4,criterion=1)
calib_data_der1_norm_ok<-calib_valid_LOO(data_set=data_calib_der1_norm_filt,trait='trait',maxcomp=10,
ratio=3/4,criterion=1)
calib_data_der2_norm_ok<-calib_valid_LOO(data_set=data_calib_der2_norm_filt,trait='trait',maxcomp=10,
ratio=3/4,criterion=1)
calib_data_norm_der1_ok<-calib_valid_LOO(data_set=data_calib_norm_der1_filt,trait='trait',maxcomp=10,
ratio=3/4,criterion=1)
calib_data_norm_der2_ok<-calib_valid_LOO(data_set=data_calib_norm_der2_filt,trait='trait',maxcomp=10,
ratio=3/4,criterion=1)
out_calib_ok<-data.frame('traitement'=c('Brut','Norm','Der1','Der2','Der1Norm','Der2Norm','NormDer1','NormDer2'),
rbind(calib_data_ok$output,calib_data_norm_ok$output,calib_data_der1_ok$output,calib_data_der2_ok$output,
calib_data_der1_norm_ok$output,calib_data_der2_norm_ok$output,calib_data_norm_der1_ok$output,
calib_data_norm_der2_ok$output),
'nb_outliers'=sapply(outliers,length))
out_calib_ok
sink("Resultats/BC/calib_LOO_decoupees_filtree.txt")
print(out_calib_ok)
sink()
#calib_data_MCCV<-MCCV(data_set=data_calib_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#calib_data_norm_MCCV<-MCCV(data_set=data_calib_norm_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#calib_data_der1_MCCV<-MCCV(data_set=data_calib_der1_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#calib_data_der2_MCCV<-MCCV(data_set=data_calib_der2_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#calib_data_der1_norm_MCCV<-MCCV(data_set=data_calib_der1_norm_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#calib_data_der2_norm_MCCV<-MCCV(data_set=data_calib_der2_norm_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#calib_data_norm_der1_MCCV<-MCCV(data_set=data_calib_norm_der1_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#calib_data_norm_der2_MCCV<-MCCV(data_set=data_calib_norm_der2_filt,trait='trait',maxcomp=10,fold=4,iter=500,criterion=1)
#out_calib_MCCV<-data.frame('traitement'=c('Brut','Norm','Der1','Der2','Der1Norm','Der2Norm','NormDer1','NormDer2'),
#rbind(calib_data_MCCV$output,calib_data_norm_MCCV$output,calib_data_der1_MCCV$output,calib_data_der2_MCCV$output,
#calib_data_der1_norm_MCCV$output,calib_data_der2_norm_MCCV$output,calib_data_norm_der1_MCCV$output,
```

```
calib_data_norm_der2_MCCV$output),
#'nb_outliers'=sapply(outliers,length))
#out_calib_MCCV
#sink("Resultats/calib_MCCV_decoupees_filtree.txt")
#print(out_calib_MCCV)
#sink()
cars_data<-carspls_LOO(X=data_calib_filt$signal,y=data_calib_filt$trait,nLV=10,iteration=100)
cars_data_norm<-carspls_LOO(X=data_calib_norm_filt$signal,y=data_calib_norm_filt$trait,nLV=10,iteration=100)
cars_data_der1<-carspls_LOO(X=data_calib_der1_filt$signal,y=data_calib_der1_filt$trait,nLV=10,iteration=100)
cars_data_der2<-carspls_LOO(X=data_calib_der2_filt$signal,y=data_calib_der2_filt$trait,nLV=10,iteration=100)
cars_data_der1_norm<-carspls_LOO(X=data_calib_der1_norm_filt$signal,y=data_calib_der1_norm_filt$trait,nLV=10,iteration=100)
cars_data_der2_norm<-carspls_LOO(X=data_calib_der2_norm_filt$signal,y=data_calib_der2_norm_filt$trait,nLV=10,iteration=100)
cars_data_norm_der1<-carspls_LOO(X=data_calib_norm_der1_filt$signal,y=data_calib_norm_der1_filt$trait,nLV=10,iteration=100)
cars_data_norm_der2<-carspls_LOO(X=data_calib_norm_der2_filt$signal,y=data_calib_norm_der2_filt$trait,nLV=10,iteration=100)
selected_lambda<-list(data=as.double(colnames(data_calib_filt$signal[,cars_data$SelectedVariables])),data_norm=as.double
(colnames(data_calib_norm_filt$signal[,cars_data_norm$SelectedVariables])),
    data_der1=as.numeric(colnames(data_calib_der1_filt$signal[,cars_data_der1$SelectedVariables])),data_der2=as.numeric
    (colnames(data_calib_der2_filt$signal[,cars_data_der2$SelectedVariables])),
    data_der1_norm=as.numeric(colnames(data_calib_der1_norm_filt$signal[,cars_data_der1_norm$SelectedVariables])),
    data_der2_norm=as.numeric(colnames(data_calib_der2_norm_filt$signal[,cars_data_der2_norm$SelectedVariables])),
    data_norm_der1=as.numeric(colnames(data_calib_norm_der1_filt$signal[,cars_data_norm_der1$SelectedVariables])),
    data_norm_der2=as.numeric(colnames(data_calib_norm_der2_filt$signal[,cars_data_norm_der2$SelectedVariables])))
sink("Resultats/BC/selected_PP_decoupees.txt")
print(selected_lambda)
sink()
data_calib_filt_red<-data_calib_filt
data_calib_filt_red$signal<-data_calib_filt_red$signal[,which(colnames(data_calib_filt_red$signal) %in% selected_lambda$data)]
data_calib_norm_filt_red<-data_calib_norm_filt
data_calib_norm_filt_red$signal<-data_calib_norm_filt_red$signal[,which(colnames(data_calib_norm_filt_red$signal)
 %in% selected_lambda$data_norm)]
data_calib_der1_filt_red<-data_calib_der1_filt
data_calib_der1_filt_red$signal<-data_calib_der1_filt_red$signal[,which(colnames(data_calib_der1_filt_red$signal)
 %in% selected_lambda$data_der1)]
data_calib_der2_filt_red<-data_calib_der2_filt
data_calib_der2_filt_red$signal<-data_calib_der2_filt_red$signal[,which(colnames(data_calib_der2_filt_red$signal)
 %in% selected_lambda$data_der2)]
data_calib_der1_norm_filt_red<-data_calib_der1_norm_filt
data_calib_der1_norm_filt_red$signal<-data_calib_der1_norm_filt_red$signal[,which(colnames
(data_calib_der1_norm_filt_red$signal) %in% selected_lambda$data_der1_norm)]
data_calib_der2_norm_filt_red<-data_calib_der2_norm_filt
data_calib_der2_norm_filt_red$signal<-data_calib_der2_norm_filt_red$signal[,which(colnames
(data_calib_der2_norm_filt_red$signal) %in% selected_lambda$data_der2_norm)]
data_calib_norm_der1_filt_red<-data_calib_norm_der1_filt
data_calib_norm_der1_filt_red$signal<-data_calib_norm_der1_filt_red$signal[,which(colnames
(data_calib_norm_der1_filt_red$signal) %in% selected_lambda$data_norm_der1)]
data_calib_norm_der2_filt_red<-data_calib_norm_der2_filt
data_calib_norm_der2_filt_red$signal<-data_calib_norm_der2_filt_red$signal[,which(colnames
(data_calib_norm_der2_filt_red$signal) %in% selected_lambda$data_norm_der2)]
write.table(data_calib_filt_red,file="Resultats/BC/donnees_decoupees_calib_filt_red.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(data_calib_norm_filt_red,file="Resultats/BC/donnees_decoupees_calib_norm_filt_red.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der1_filt_red,file="Resultats/BC/donnees_decoupees_calib_der1_filt_red.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der2_filt_red,file="Resultats/BC/donnees_decoupees_calib_der2_filt_red.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der1_norm_filt_red,file="Resultats/BC/donnees_decoupees_der1_norm_calib_filt_red.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_der2_norm_filt_red,file="Resultats/BC/donnees_decoupees_der2_norm_calib_filt_red.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_norm_der1_filt_red,file="Resultats/BC/donnees_decoupees_norm_der1_calib_filt_red.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(data_calib_norm_der2_filt_red,file="Resultats/BC/donnees_decoupees_norm_der2_calib_filt_red.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
calib_data_ok_red<-calib_valid_LOO(data_set=data_calib_filt_red,trait='trait',maxcomp=min(c(ncol
(data_calib_filt_red$signal),10)),ratio=3/4,criterion=1)
calib_data_norm_ok_red<-calib_valid_LOO(data_set=data_calib_norm_filt_red,trait='trait',maxcomp=min
(c(ncol(data_calib_norm_filt_red$signal),10)),ratio=3/4,criterion=1)
calib_data_der1_ok_red<-calib_valid_LOO(data_set=data_calib_der1_filt_red,trait='trait',maxcomp=min
(c(ncol(data_calib_der1_filt_red$signal),10)),ratio=3/4,criterion=1)
calib_data_der2_ok_red<-calib_valid_LOO(data_set=data_calib_der2_filt_red,trait='trait',maxcomp=min
(c(ncol(data_calib_der2_filt_red$signal),10)),ratio=3/4,criterion=1)
calib_data_der1_norm_ok_red<-calib_valid_LOO(data_set=data_calib_der1_norm_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_der1_norm_filt_red$signal),10)),ratio=3/4,criterion=1)
calib_data_der2_norm_ok_red<-calib_valid_LOO(data_set=data_calib_der2_norm_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_der2_norm_filt_red$signal),10)),ratio=3/4,criterion=1)
calib_data_norm_der1_ok_red<-calib_valid_LOO(data_set=data_calib_norm_der1_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_norm_der1_filt_red$signal),10)),ratio=3/4,criterion=1)
calib_data_norm_der2_ok_red<-calib_valid_LOO(data_set=data_calib_norm_der2_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_norm_der2_filt_red$signal),10)),ratio=3/4,criterion=1)
out_calib_ok_red<-data.frame('traitement'=c('Brut','Norm','Der1','Der2','Der1Norm','Der2Norm',
'NormDer1','NormDer2'),
rbind(calib_data_ok_red$output,calib_data_norm_ok_red$output,calib_data_der1_ok_red$output,
calib_data_der2_ok_red$output,
calib_data_der1_norm_ok_red$output,calib_data_der2_norm_ok_red$output,calib_data_norm_der1_ok_red$output,
calib_data_norm_der2_ok_red$output),
'nb_outliers'=sapply(outliers,length),
'nb_PP'=sapply(selected_lambda,length))
out_calib_ok_red
sink("Resultats/BC/calib_LOO_decoupees_filtree_reduite.txt")
print(out_calib_ok_red)
sink()
```

```
#calib_data_MCCV_red<-MCCV(data_set=data_calib_filt_red,trait='trait',maxcomp=min(c(ncol
(data_calib_filt_red$signal),10)),fold=4,iter=500,criterion=1)
#calib_data_norm_MCCV_red<-MCCV(data_set=data_calib_norm_filt_red,trait='trait',maxcomp=min
(c(ncol(data_calib_norm_filt_red$signal),10)),fold=4,iter=500,criterion=1)
```

```
#calib_data_der1_MCCV_red<-MCCV(data_set=data_calib_der1_filt_red,trait='trait',maxcomp=min
(c(ncol(data_calib_der1_filt_red$signal),10)),fold=4,iter=500,criterion=1)
#calib_data_der2_MCCV_red<-MCCV(data_set=data_calib_der2_filt_red,trait='trait',maxcomp=min
(c(ncol(data_calib_der2_filt_red$signal),10)),fold=4,iter=500,criterion=1)
#calib_data_der1_norm_MCCV_red<-MCCV(data_set=data_calib_der1_norm_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_der1_norm_filt_red$signal),10)),fold=4,iter=500,criterion=1)
#calib_data_der2_norm_MCCV_red<-MCCV(data_set=data_calib_der2_norm_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_der2_norm_filt_red$signal),10)),fold=4,iter=500,criterion=1)
#calib_data_norm_der1_MCCV_red<-MCCV(data_set=data_calib_norm_der1_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_norm_der1_filt_red$signal),10)),fold=4,iter=500,criterion=1)
#calib_data_norm_der2_MCCV_red<-MCCV(data_set=data_calib_norm_der2_filt_red,trait='trait',
maxcomp=min(c(ncol(data_calib_norm_der2_filt_red$signal),10)),fold=4,iter=500,criterion=1)
#out_calib_MCCV_red<-data.frame('traitement'=c('Brut','Norm','Der1','Der2','Der1Norm',
'Der2Norm','NormDer1','NormDer2'),
#rbind(calib_data_MCCV_red$output,calib_data_norm_MCCV_red$output,
calib_data_der1_MCCV_red$output,calib_data_der2_MCCV_red$output,
#calib_data_der1_norm_MCCV_red$output,calib_data_der2_norm_MCCV_red$output,
calib_data_norm_der1_MCCV_red$output,calib_data_norm_der2_MCCV_red$output),
#'nb_outliers'=sapply(outliers,length),
#'nb_PP'=sapply(selected_lambda,length))
#out_calib_MCCV_red
#sink("Resultats/calib_MCCV_decoupees_filtree_reduite.txt")
#print(out_calib_MCCV_red)
#sink()
#output<-list('trait'=chem,'outliers'=outliers,'Cars_Sel_PP'=selected_lambda,
#    'LOO_CV_AllPP'=out_calib_ok,'MCCV_AllPP'=out_calib_MCCV,
#    'LOO_CV_CarsSelPP'=out_calib_ok_red,'MCCV_CarsSelPP'=out_calib_MCCV_red)
#sink("Resultats/recapitulatif_calibration_decoupees.txt")
#print(output)
#sink()
pretr='norm_'
nbcomp=5
data_ok<-get(paste("data_calib_",pretr,"filt_red",sep=""))
LOO_CV<-plsr(trait~signal,data=data_ok,nbcomp,validation='LOO')
plot(LOO_CV,ncomp=nbcomp,line=T,main=chem,pch=20,col='darkblue')
legend('topleft',c(paste('Nb_Comp=',nbcomp,sep=''),
paste('R2_train=',round(R2(LOO_CV,estimate='train',ncomp=nbcomp)$val[2],2),sep=''),
paste('R2_cv=',round(R2(LOO_CV,estimate='CV',ncomp=nbcomp)$val[2],2),sep=''),
paste('RMSE_cv=',round(RMSEP(LOO_CV,estimate='CV',ncomp=nbcomp)$val[2],2),sep=''),
paste('RPD_cv=',round(sd(data_ok$trait)/RMSEP(LOO_CV,estimate='CV',ncomp=nbcomp)$val[2],2),
sep='')),bty='n')
png(file="Resultats/Calibration_choisie_decoupees.png",
width=600, height=600)
```

```
plot(LOO_CV,ncomp=nbcomp,line=T,main=chem,pch=20,col='darkblue')
legend('topleft',c(paste('Nb_Comp=',nbcomp,sep=''),
paste('R2_train=',round(R2(LOO_CV,estimate='train',ncomp=nbcomp)$val[2],2),sep=''),
paste('R2_cv=',round(R2(LOO_CV,estimate='CV',ncomp=nbcomp)$val[2],2),sep=''),
paste('RMSE_cv=',round(RMSEP(LOO_CV,estimate='CV',ncomp=nbcomp)$val[2],2),sep=''),
paste('RPD_cv=',round(sd(data_ok$trait)/RMSEP(LOO_CV,estimate='CV',ncomp=nbcomp)$val[2],2),
sep='')),bty='n')
dev.off()
sink("Resultats/Coeff_latent_variable_decoupees.txt")
print(LOO_CV$coefficient)
sink()
sink("Resultats/Coeff_interception_decoupees.txt")
print(LOO_CV$Ymeans)
sink()
plot(LOO_CV, plottype = "coef", ncomp=nbcomp, legendpos = "bottomleft", labels = "numbers")
```

```
options("scipen"=100)
install.packages("signal")
library(signal)
setwd('C:/Users/Mathilde/dox/EXCOM700 Biocomposites fibres-TPS/Interpretation R/TraitementR')
donnees_brutes<-read.table("DonneesBrutes/BC/data_FTIR_BC.txt",sep="\t",header=TRUE)
xlabel='Lambda cm-1'
ylabel='Absorbance'
xnondecoupe=c(4100,300)
xdecoupe=c(4000,400)
donnees_brutes_ok<-data.frame('PP'=as.numeric(donnees_brutes$lambda),'signal'=I(as.matrix
(donnees_brutes[,2:ncol(donnees_brutes)])))
windows()
matplot(donnees_brutes_ok$PP,donnees_brutes_ok$signal,type='l',lty=1,xlab=xlabel,ylab=ylabel,
xlim=xnondecoupe,main='donnees_brutes_ok',col=3)
png(file="Resultats/données_brutes_ok.png",
width=600, height=600)
matplot(donnees_brutes_ok$PP,donnees_brutes_ok$signal,type='l',lty=1,xlab=xlabel,ylab=ylabel,
xlim=xnondecoupe,main='donnees_brutes_ok',col=3)
dev.off()
donnees_decoupees<-subset(donnees_brutes_ok,PP >= 410 & PP <= 3900)
write.table(donnees_decoupees,file="Resultats/donnees_decoupees.txt",sep="\t",row.names=FALSE,
col.names=TRUE)
```

```
windows()
matplot(donnees_decoupees$PP,donnees_decoupees$signal,type='l',lty=1,xlab=xlabel,ylab=ylabel,
xlim=xdecoupe,main='donnees_decoupees',col=3)
png(file="Resultats/données_decoupées.png",
width=600, height=600)
matplot(donnees_decoupees$PP,donnees_decoupees$signal,type='l',lty=1,xlab=xlabel,ylab=ylabel,
xlim=xdecoupe,main='donnees_decoupees',col=3)
dev.off()
normalisation<-function(x){(x-mean(x))/sd(x)}
tsf<-(max(donnees_decoupees$PP)-min(donnees_decoupees$PP))/(length(donnees_decoupees$PP)-1)
donnees_decoupees_norm<-data.frame(PP=donnees_decoupees$PP,signal=I(as.matrix(apply
(donnees_decoupees$signal,2,normalisation))))
donnees_decoupees_der1<-data.frame(PP=donnees_decoupees$PP,signal=I(as.matrix(apply
(donnees_decoupees$signal,2,function(x){sgolayfilt(x,p=2,n=37,m=1,ts=tsf)}))))
donnees_decoupees_der2<-data.frame(PP=donnees_decoupees$PP,signal=I(as.matrix(apply
(donnees_decoupees$signal,2,function(x){sgolayfilt(x,p=3,n=61,m=2,ts=tsf)}))))
donnees_decoupees_der1_norm<-data.frame(PP=donnees_decoupees_der1$PP,signal=I(as.matrix
(apply(donnees_decoupees_der1$signal,2,normalisation))))
donnees_decoupees_der2_norm<-data.frame(PP=donnees_decoupees_der2$PP,signal=I(as.matrix
(apply(donnees_decoupees_der2$signal,2,normalisation))))
donnees_decoupees_norm_der1<-data.frame(PP=donnees_decoupees_norm$PP,signal=I(as.matrix
(apply(donnees_decoupees_norm$signal,2,function(x){sgolayfilt(x,p=2,n=37,m=1,ts=tsf)}))))
donnees_decoupees_norm_der2<-data.frame(PP=donnees_decoupees_norm$PP,signal=I(as.matrix
(apply(donnees_decoupees_norm$signal,2,function(x){sgolayfilt(x,p=3,n=61,m=2,ts=tsf)}))))
write.table(donnees_decoupees_norm,file="Resultats/donnees_decoupees_norm.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(donnees_decoupees_der1,file="Resultats/donnees_decoupees_der1.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(donnees_decoupees_der2,file="Resultats/donnees_decoupees_der2.txt",sep="\t",
row.names=FALSE,col.names=TRUE)
write.table(donnees_decoupees_der1_norm,file="Resultats/donnees_decoupees_der1_norm.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(donnees_decoupees_der2_norm,file="Resultats/donnees_decoupees_der2_norm.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(donnees_decoupees_norm_der1,file="Resultats/donnees_decoupees_norm_der1.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
write.table(donnees_decoupees_norm_der2,file="Resultats/donnees_decoupees_norm_der2.txt",
sep="\t",row.names=FALSE,col.names=TRUE)
windows()
matplot(donnees_decoupees_norm$PP,donnees_decoupees_norm$signal,type='l',lty=1,xlab=xlabel,
ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees',col=3)
windows()
matplot(donnees_decoupees_der1$PP,donnees_decoupees_der1$signal,type='l',lty=1,xlab=xlabel,
ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der1',col=3)
windows()
matplot(donnees_decoupees_der2$PP,donnees_decoupees_der2$signal,type='l',lty=1,xlab=xlabel,
ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der2',col=3)
windows()
matplot(donnees_decoupees_der1_norm$PP,donnees_decoupees_der1_norm$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der1_norm',col=3)
windows()
matplot(donnees_decoupees_der2_norm$PP,donnees_decoupees_der2_norm$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der2_norm',col=3)
windows()
matplot(donnees_decoupees_norm_der1$PP,donnees_decoupees_norm_der1$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_norm_der1',col=3)
windows()
matplot(donnees_decoupees_norm_der2$PP,donnees_decoupees_norm_der2$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_norm_der2',col=3)
png(file="Resultats/données_découpées_normalisées.png",
width=600, height=600)
matplot(donnees_decoupees_norm$PP,donnees_decoupees_norm$signal,type='l',lty=1,xlab=xlabel,
ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_norm',col=3)
dev.off()
png(file="Resultats/données_découpées_der1.png",
width=600, height=600)
matplot(donnees_decoupees_der1$PP,donnees_decoupees_der1$signal,type='l',lty=1,xlab=xlabel,
ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der1',col=3)
dev.off()
png(file="Resultats/données_découpées_der2.png",
width=600, height=600)
matplot(donnees_decoupees_der2$PP,donnees_decoupees_der2$signal,type='l',lty=1,xlab=xlabel,
ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der2',col=3)
dev.off()
png(file="Resultats/données_découpées_der1_normalisées.png",
width=600, height=600)
matplot(donnees_decoupees_der1_norm$PP,donnees_decoupees_der1_norm$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der1_norm',col=3)
dev.off()
png(file="Resultats/données_découpées_der2_normalisées.png",
width=600, height=600)
matplot(donnees_decoupees_der2_norm$PP,donnees_decoupees_der2_norm$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_der2_norm',col=3)
dev.off()
```

```
png(file="Resultats/données_découpées_normalisées_der1.png",
width=600, height=600)
matplot(donnees_decoupees_norm_der1$PP,donnees_decoupees_norm_der1$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_norm_der1',col=3)
dev.off()
png(file="Resultats/données_découpées_normalisées_der2.png",
width=600, height=600)
matplot(donnees_decoupees_norm_der2$PP,donnees_decoupees_norm_der2$signal,type='l',lty=1,
xlab=xlabel,ylab=ylabel,xlim=xdecoupe,main='donnees_decoupees_norm_der2',col=3)
dev.off()
```

6) References used to create the initial database

Abdullah, A.H.D., Putri, O.D., Fikriyyah, A.K., Nissa, R.C., Intadiana, S., 2020. Effect of microcrystalline cellulose on characteristics of cassava starch-based bioplastic. Polym. Technol. Mater. 59, 1250–1258. https://doi.org/10.1080/25740881.2020.1738465

Abera, G., Woldeyes, B., Demash, H.D., Miyake, G., 2020. The effect of plasticizers on thermoplastic starch films developed from the indigenous Ethiopian tuber crop Anchote (Coccinia abyssinica) starch. Int. J. Biol. Macromol. 155, 581–587. https://doi.org/10.1016/j.ijbiomac.2020.03.218

Altayan, M.M., Al Darouich, T., Karabet, F., 2020. Thermoplastic starch from corn and wheat: a comparative study based on amylose content. Polym. Bull. https://doi.org/10.1007/s00289-020-03262-9

Angellier, H., Molina-Boisseau, S., Dole, P., Dufresne, A., 2006. Thermoplastic Starch−Waxy Maize Starch Nanocrystals Nanocomposites. Biomacromolecules 7, 531–539. https://doi.org/10.1021/bm050797s

Averous, L., 2000. Properties of thermoplastic blends: starch–polycaprolactone. Polymer (Guildf). 41, 4157–4167. https://doi.org/10.1016/S0032-3861(99)00636-9

Averous, L., Boquillon, N., 2004. Biocomposites based on plasticized starch: thermal and mechanical behaviours. Carbohydr. Polym. 56, 111–122. https://doi.org/10.1016/j.carbpol.2003.11.015

Avérous, L., Fringant, C., Moro, L., 2001. Plasticized starch–cellulose interactions in polysaccharide composites. Polymer (Guildf). 42, 6565–6572. https://doi.org/10.1016/S0032-3861(01)00125-2

Ayadi, F., Dole, P., 2011. Stoichiometric interpretation of thermoplastic starch water sorption and relation to mechanical behavior. Carbohydr. Polym. 84, 872–880. https://doi.org/10.1016/j.carbpol.2010.12.024

Cheng, G., Zhou, M., Wei, Y.-J., Cheng, F., Zhu, P.-X., 2019. Comparison of mechanical reinforcement effects of cellulose nanocrystal, cellulose nanofiber, and microfibrillated cellulose in starch composites. Polym. Compos. 40, E365–E372. https://doi.org/10.1002/pc.24685

Chung, Y.-L., Ansari, S., Estevez, L., Hayrapetyan, S., Giannelis, E.P., Lai, H.-M., 2010. Preparation and properties of biodegradable starch–clay nanocomposites. Carbohydr. Polym. 79, 391–396. https://doi.org/10.1016/j.carbpol.2009.08.021

Collazo-Bigliardi, S., Ortega-Toro, R., Chiralt Boix, A., 2018. Reinforcement of Thermoplastic Starch Films with Cellulose Fibres Obtained from Rice and Coffee Husks. J. Renew. Mater. 6, 599–610. https://doi.org/10.32604/JRM.2018.00127

Córdoba, A., Cuéllar, N., González, M., Medina, J., 2008. The plasticizing effect of alginate on the thermoplastic starch/glycerin blends. Carbohydr. Polym. 73, 409–416. https://doi.org/10.1016/j.carbpol.2007.12.007

Curvelo, A., 2001. Thermoplastic starch–cellulosic fibers composites: preliminary results. Carbohydr. Polym. 45, 183–188. https://doi.org/10.1016/S0144-8617(00)00314-3

Dai, H., Chang, P.R., Yu, J., Ma, X., 2008. N,N-Bis(2-hydroxyethyl)formamide as a New Plasticizer for Thermoplastic Starch. Starch - Stärke 60, 676–684. https://doi.org/10.1002/star.200800017

De Carvalho, A.J.F., Curvelo, A.A.S., Agnelli, J.A.M., 2002. Wood pulp reinforced thermoplastic starch composites. Int. J. Polym. Mater. 51, 647–660. https://doi.org/10.1080/714975803

De Carvalho, A.J., Curvelo, A.A., Agnelli, J.A., 2001. A first insight on composites of thermoplastic starch and kaolin. Carbohydr. Polym. 45, 189–194. https://doi.org/10.1016/S0144-8617(00)00315-5

Dean, K.M., Do, M.D., Petinakis, E., Yu, L., 2008. Key interactions in biodegradable thermoplastic starch/poly(vinyl alcohol)/montmorillonite micro- and nanocomposites. Compos. Sci. Technol. 68, 1453–1462. https://doi.org/10.1016/j.compscitech.2007.10.037

Dias, A.B., Müller, C.M.O., Larotonda, F.D.S., Laurindo, J.B., 2011. Mechanical and barrier properties of composite films based on rice flour and cellulose fibers. LWT - Food Sci. Technol. 44, 535–542. https://doi.org/10.1016/j.lwt.2010.07.006

Fazeli, M., Florez, J.P., Simão, R.A., 2019. Improvement in adhesion of cellulose fibers to the thermoplastic starch matrix by plasma treatment modification. Compos. Part B Eng. 163, 207–216. https://doi.org/10.1016/j.compositesb.2018.11.048

Fourati, Y., Magnin, A., Putaux, J.-L., Boufi, S., 2020. One-step processing of plasticized starch/cellulose nanofibrils nanocomposites via twin-screw extrusion of starch and cellulose fibers. Carbohydr. Polym. 229, 115554. https://doi.org/10.1016/j.carbpol.2019.115554

González, K., Iturriaga, L., González, A., Eceiza, A., Gabilondo, N., 2020. Improving mechanical and barrier properties of thermoplastic starch and polysaccharide nanocrystals nanocomposites. Eur. Polym. J. 123, 109415. https://doi.org/10.1016/j.eurpolymj.2019.109415

Granda, L.A., Oliver-Ortega, H., Fabra, M.J., Tarrés, Q., Pèlach, M.À., Lagarón, J.M., Méndez, J.A., 2020. Improved Process to Obtain Nanofibrillated Cellulose (CNF) Reinforced Starch Films with Upgraded Mechanical Properties and Barrier Character. Polymers (Basel). 12, 1071. https://doi.org/10.3390/polym12051071

Henrique dos Santos, B., de Souza do Prado, K., Jacinto, A.A., da Silva Spinacé, M.A., 2018. Influence of Sugarcane Bagasse Fiber Size on Biodegradable Composites of Thermoplastic Starch. J. Renew. Mater. 6, 176–182. https://doi.org/10.7569/JRM.2018.634101

Huneault, M.A., Li, H., 2012. Preparation and properties of extruded thermoplastic starch/polymer blends. J. Appl. Polym. Sci. 126, E96–E108. https://doi.org/10.1002/app.36724

Kaewtatip, K., Thongmee, J., 2012. Studies on the structure and properties of thermoplastic starch/luffa fiber composites. Mater. Des. 40, 314–318. https://doi.org/10.1016/j.matdes.2012.03.053

Kaushik, A., Singh, M., Verma, G., 2010. Green nanocomposites based on thermoplastic starch and steam exploded cellulose nanofibrils from wheat straw. Carbohydr. Polym. 82, 337–345. https://doi.org/10.1016/j.carbpol.2010.04.063

Lai, S.-M., Don, T.-M., Huang, Y.-C., 2006. Preparation and properties of biodegradable thermoplastic starch/poly(hydroxy butyrate) blends. J. Appl. Polym. Sci. 100, 2371–2379. https://doi.org/10.1002/app.23085

Liu, W., Wang, Z., Liu, J., Dai, B., Hu, S., Hong, R., Xie, H., Li, Z., Chen, Y., Zeng, G., 2020. Preparation, reinforcement and properties of thermoplastic starch film by film blowing. Food Hydrocoll. 108, 106006. https://doi.org/10.1016/j.foodhyd.2020.106006

Liu, Z.Q., Yi, X.S., Feng, Y., 2001. Effects of glycerin and glycerol monostearate on performance of thermoplastic starch. J. Mater. Sci. 36, 1809–1815.

López, J.P., Mutjé, P., Carvalho, A.J.F., Curvelo, A.A.S., Gironès, J., 2013. Newspaper fiber-reinforced thermoplastic starch biocomposites obtained by melt processing: Evaluation of the mechanical, thermal and water sorption properties. Ind. Crops Prod. 44, 300–305. https://doi.org/10.1016/j.indcrop.2012.11.020

Lopez, O., Garcia, M.A., Villar, M.A., Gentili, A., Rodriguez, M.S., Albertengo, L., 2014. Thermo-compression of biodegradable thermoplastic corn starch films containing chitin and chitosan. LWT - Food Sci. Technol. 57, 106–115. https://doi.org/10.1016/j.lwt.2014.01.024

López, O. V., García, M.A., 2012. Starch films from a novel (Pachyrhizus ahipa) and conventional sources: Development and characterization. Mater. Sci. Eng. C 32, 1931–1940. https://doi.org/10.1016/j.msec.2012.05.035

Lopez-Gil, A., Rodriguez-Perez, M.A., De Saja, J.A., Bellucci, F.S., Ardanuy, M., 2014. Strategies to Improve the Mechanical Properties of Starch-Based Materials: Plasticization and Natural Fibers Reinforcement. Polímeros Ciência e Tecnol. 24, 36–42. https://doi.org/10.4322/polimeros.2014.054

Lu, Y., Tighzert, L., Dole, P., Erre, D., 2005. Preparation and properties of starch thermoplastics modified with waterborne polyurethane from renewable resources. Polymer (Guildf). 46, 9863–9870. https://doi.org/10.1016/j.polymer.2005.08.026

Lu, Y., Weng, L., Cao, X., 2006. Morphological, thermal and mechanical properties of ramie crystallites—reinforced plasticized starch biocomposites. Carbohydr. Polym. 63, 198–204. https://doi.org/10.1016/j.carbpol.2005.08.027

Mahieu, A., Terrié, C., Youssef, B., 2015. Thermoplastic starch films and thermoplastic starch/polycaprolactone blends with oxygen-scavenging properties: Influence of water content. Ind. Crops Prod. 72, 192–199. https://doi.org/10.1016/j.indcrop.2014.11.037

Martins, I.M.G., Magina, S.P., Oliveira, L., Freire, C.S.R., Silvestre, A.J.D., Neto, C.P., Gandini, A., 2009. New biocomposites based on thermoplastic starch and bacterial cellulose. Compos. Sci. Technol. 69, 2163–2168. https://doi.org/10.1016/j.compscitech.2009.05.012

Mehyar, G.F., Han, J.H., 2006. Physical and Mechanical Properties of High-amylose Rice and Pea Starch Films as Affected by Relative Humidity and Plasticizer. J. Food Sci. 69, E449–E454. https://doi.org/10.1111/j.1365-2621.2004.tb09929.x

Mihai, M., Legros, N., Alemdar, A., 2014. Formulation-properties versatility of wood fiber biocomposites based on polylactide and polylactide/thermoplastic starch blends. Polym. Eng. Sci. 54, 1325–1340. https://doi.org/c

Mościcki, L., Mitrus, M., Wójtowicz, A., Oniszczuk, T., Rejak, A., Janssen, L., 2012. Application of extrusion-cooking for processing of thermoplastic starch (TPS). Food Res. Int. 47, 291–299. https://doi.org/10.1016/j.foodres.2011.07.017

Müller, C.M.O., Laurindo, J.B., Yamashita, F., 2012. Composites of thermoplastic starch and nanoclays produced by extrusion and thermopressing. Carbohydr. Polym. 89, 504–510. https://doi.org/10.1016/j.carbpol.2012.03.035

Nasri-Nasrabadi, B., Behzad, T., Bagheri, R., 2014. Preparation and characterization of cellulose nanofiber reinforced thermoplastic starch composites. Fibers Polym. 15, 347–354. https://doi.org/10.1007/s12221-014-0347-0

Nordin, N., Othman, S.H., Kadir Basha, R., Abdul Rashid, S., 2018. Mechanical and thermal properties of starch films reinforced with microcellulose fibres. Food Res. 2, 555–563. https://doi.org/10.26656/fr.2017.2(6).110

Park, H.M., Lee, W.K., Park, C.Y., Cho, W.J., Ha, C.S., 2003. Environmentally friendly polymer hybrids Part I mechanical, thermal, and barrier properties of thermoplastic starch/clay nanocomposites. J. Mater. Sci. 38, 909–915. https://doi.org/https://doi.org/10.1023/A:1022308705231

Park, H.-M., Li, X., Jin, C.-Z., Park, C.-Y., Cho, W.-J., Ha, C.-S., 2002. Preparation and Properties of Biodegradable Thermoplastic Starch/Clay Hybrids. Macromol. Mater. Eng. 287, 553–558. https://doi.org/10.1002/1439-2054(20020801)287:8<553::AID-MAME553>3.0.CO;2-3

Peng, Y., Zha, D., Bin, G., Bengang, L., Panxin, L., 2020. Effect of wheat straw oxidation on thermoplastic starch composites: Mechanical, thermal, and rheological process behaviors. J. Thermoplast. Compos. Mater. 33, 646–658. https://doi.org/10.1177/0892705718809802

Phan, T.D., Debeaufort, F., Luu, D., Voilley, A., 2005. Functional Properties of Edible Agar-Based and Starch-Based Films for Food Quality Preservation. J. Agric. Food Chem. 53, 973–981. https://doi.org/10.1021/jf040309s

Prachayawarakorn, J., Sangnitidej, P., Boonpasith, P., 2010. Properties of thermoplastic rice starch composites reinforced by cotton fiber or low-density polyethylene. Carbohydr. Polym. 81, 425–433. https://doi.org/10.1016/j.carbpol.2010.02.041

Prachayawarakorn, J., Chaiwatyothin, S., Mueangta, S., Hanchana, A., 2013. Effect of jute and kapok fibers on properties of thermoplastic cassava starch composites. Mater. Des. 47, 309–315. https://doi.org/10.1016/j.matdes.2012.12.012

Prachayawarakorn, J., Hommanee, L., Phosee, D., Chairapaksatien, P., 2010. Property improvement of thermoplastic mung bean starch using cotton fiber and low-density polyethylene. Starch - Stärke 62, 435–443. https://doi.org/10.1002/star.201000002

Ríos-Soberanis, C.R., Collí-Pacheco, J.P., Estrada-León, R.J., Moo-Huchin, V.M., Yee-Madeira, H.T., Pérez-Pacheco, E., 2020. Biocomposites based on plasticized starch: thermal, mechanical and morphological characterization. Polym. Bull. https://doi.org/10.1007/s00289-020-03261-w

Sahari, J., Sapuan, S.M., Zainudin, E.S., Maleque, M.A., 2013. Thermo-mechanical behaviors of thermoplastic starch derived from sugar palm tree (Arenga pinnata). Carbohydr. Polym. 92, 1711–1716. https://doi.org/10.1016/j.carbpol.2012.11.031

Shi, R., Liu, Q., Ding, T., Han, Y., Zhang, L., Chen, D., Tian, W., 2007. Ageing of soft thermoplastic starch with high glycerol content. J. Appl. Polym. Sci. 103, 574–586. https://doi.org/10.1002/app.25193

Smitthipong, W., Tantatherdtam, R., Chollakup, R., 2015. Effect of pineapple leaf fiber-reinforced thermoplastic starch/poly(lactic acid) green composite. J. Thermoplast. Compos. Mater. 28, 717–729. https://doi.org/10.1177/0892705713489701

Teixeira, E. de M., Pasquini, D., Curvelo, A.A.S., Corradini, E., Belgacem, M.N., Dufresne, A., 2009. Cassava bagasse cellulose nanofibrils reinforced thermoplastic cassava starch. Carbohydr. Polym. 78, 422–431. https://doi.org/10.1016/j.carbpol.2009.04.034

Thunwall, M., Boldizar, A., Rigdahl, M., 2006. Compression Molding and Tensile Properties of Thermoplastic Potato Starch Materials. Biomacromolecules 7, 981–986. https://doi.org/10.1021/bm050804c

Torres, F.G., Arroyo, O.H., Gomez, C., 2007. Processing and Mechanical Properties of Natural Fiber Reinforced Thermoplastic Starch Biocomposites. J. Thermoplast. Compos. Mater. 20, 207–223. https://doi.org/10.1177/0892705707073945

Van Soest, J.J.G., Borger, D.B., 1997. Structure and properties of compression-molded thermoplastic starch materials from normal and high-amylose maize starches. J. Appl. Polym. Sci. 64, 631–644. https://doi.org/10.1002/(SICI)1097-4628(19970425)64:4<631::AID-APP2>3.0.CO;2-O

Wang, H.-Y., Huang, M.-F., 2007. Preparation, characterization and performances of biodegradable thermoplastic starch. Polym. Adv. Technol. 18, 910–915. https://doi.org/10.1002/pat.930

Wattanakornsiri, A., Pachana, K., Kaewpirom, S., Sawangwong, P., Migliaresi, C., 2011. Green composites of thermoplastic corn starch and recycled paper cellulose fibers. Songklanakarin J. Sci. Technol. 33, 461–467.

Zdanowicz, M., 2020. Starch treatment with deep eutectic solvents, ionic liquids and glycerol. A comparative study. Carbohydr. Polym. 229, 115574. https://doi.org/10.1016/j.carbpol.2019.115574

Zhang, Y., Han, J.H., 2006. Mechanical and Thermal Characteristics of Pea Starch Films Plasticized with Monosaccharides and Polyols. J. Food Sci. 71, E109–E118. https://doi.org/10.1111/j.1365-2621.2006.tb08891.x

Zullo, R., Iannace, S., 2009. The effects of different starch sources and plasticizers on film blowing of thermoplastic starch: Correlation among process, elongational properties and macromolecular structure. Carbohydr. Polym. 77, 376–383. https://doi.org/10.1016/j.carbpol.2009.01.007

7) Code used to design the models

```
###Téléchargement du logiciel : https://www.anaconda.com/producEaB/individual -
Python 3.7 64bit (32 bit non fonctionnel, à vérifier avec le processeur de l'ordinateur)

#Littérature : ce code est inspiré du "Boston housing - price prediction" décrit ci-dessous
#https://tensorflow.rstudio.com/tutorials/beginners/basic-ml/tutorial_basic_regression/
#https://rstudio-pubs-static.s3.amazonaws.com/364346_811c9012a14847428c9b1fc1e956431a.html

###Adaptation du code : Sophie Morin - PhD Student - morin.msophie@gmail.com###
##Testeurs : Louise Delahaye ###



#### Chargement des packages ####
#Enlevez les "#" pour le premier essai du code, puis il ne sera plus nécessaire
d'installer chacun des packages"
#install.packages('keras')
library(keras)
#install_keras()
#install.packages('devtools')
library('devtools')
#install_devtools()
#install.packages("magrittr")
#install.packages("dplyr")
library(magrittr)
library(dplyr)
reticulate::use_condaenv
#install.packages('tfdataseEaB')
library(tfdatasets)
#install.packages('psych')
library(psych)
#install.packages('car')
library(car)
library(tibble)
library(caret)
library(stringr)

#Introduire ici le chemin d'accès de votre base de données#
# Attention : en copiant-collant le chemin d'accès, les "\" doivent être changés en "/"#
#Dossier
setwd('C:/Users/')


#Introduire ici le nom de la base de données de la littérature avec son extension,
qui servira a entraîné le modèle#
#La base de données doit se trouver dans le dossier mentionné précédemment#
#Attention : bien vérifier que le séparateur des décimales est un point et non une virgule#

NomDB_litterature <- readline(prompt="What is the name of the literature database?
(please add the extension of your file):")

###Lecture des bases par le logiciel
##La base de la littérature est affectée à la variable "trainset"#
trainset<-as_tibble(read.table(NomDB_litterature,sep="\t",header=TRUE))
trainset

##Lecture du nom des colonnes des bases de données
column_trainset <- names(trainset)
summarize_trainset <- summary(trainset)
#This is the repartition of your dataset :
summarize_trainset
write.table(summarize_trainset, file= "database_description.txt")
print(paste("Column header of the litterature database", column_trainset))
```

```
#Please introduce in the following code lines the rownumber associated with the
variable you want to study #
studied_parameters <- trainset[,c(4,5,6,7,8,10,11,12,13,14,15,16,17,19)]
studied_parameters

#Go in your folder to load the jpeg image called 'Trainset_plot' representing the
correlation matrices#
jpeg("Trainset_plot.jpg", width = 500, height = 500)
Trainsetplot <- plot(studied_parameters,pch=3)
dev.off()

#The following section will calculate the correlation of the variables of interest
toward the variable to predict#
#Please inter after the '$' sign the name of your column to predict
correlation <- cor(studied_parameters,studied_parameters$EaB)
#The following variable have this coefficient correlation toward the variable to predict#
correlation
#Any feature which is not significant (p<0.05), is not contributing signicantly for the model
#Variable tagged with "True" are significantly contributing to the model
selected_parameters <-  (abs(correlation) > (0.05))
selected_parameters


write.table(correlation, file= "correlation.txt")

#Repartition of the data in the variable to predict. Don't forget to write the name
of your variable of interest after the "$" sign)
jpeg("histogramm_trainset.jpg", width = 1000, height = 500)
histogramme<-hist(trainset$EaB,xlab="Median Value", main="EaB", col="grey")
dev.off()

#set.seed indicate to the computer a time limit to perform the computation
set.seed(30000)
#The algorithm separate randomly here the litterature database onto two subset :
the first one help to create the model, while the second check the accuracy of the model")
inTrain <- createDataPartition(y=studied_parameters$EaB,p=0.70,list=FALSE)
training <- studied_parameters[inTrain,]
testing <- studied_parameters[-inTrain,]
training

###The model is build here. We indicate which kind of generic formula the algorithm
need to perform the simulations
#General linear Model
fit.lm <- lm(EaB~ .,data = training)
coefficients_linear <- data.frame(coef = round(fit.lm$coefficients,2))
#This is the coefficienEaB used to create the model, saved as "coefficienEaB_linear_model.txt"
in your folder
coefficients_linear
write.table(coefficients_linear, file= "coefficients_linear_model.txt")
#This section check the model accuracy on the base of the randomly created subset of the
litterature database
set.seed(30000)
pred.lm <- predict(fit.lm, newdata = testing)
#Please check if the variable indicated after the $ sign is correct
rmse.lm <- sqrt(sum((pred.lm - testing$EaB)^2)/
length(testing$EaB))
validity_lm <- c(RMSE = rmse.lm, R2 = summary(fit.lm)$r.squared)
write.table(validity_lm, file= "validity_linear_model.txt")
validity_lm
summary(fit.lm)

# This section create a new model based on a logarithmic formula
fit.lm1 <- lm(log(EaB)~.,data = training)
coefficients_logarithm <- data.frame(coef = round(fit.lm1$coefficients,2))


coefficients_logarithm
write.table(coefficients_logarithm, file= "coefficients_logarithm_model.txt")
pred.lm1 <- predict(fit.lm1, newdata = testing)
#Please check if the variable indicated after the $ sign is correct
rmse.lm1 <- sqrt(sum((exp(pred.lm1) - testing$EaB)^2)/
length(testing$EaB))
validity_lm1<- c(RMSE = rmse.lm1, R2 = summary(fit.lm1)$r.squared)
write.table(validity_lm1, file= "validity_logarithm_model.txt")
summary(fit.lm1)
validity_lm1
```

```
# You can add more mathematical transformations here. Don't forget to adapt the RMSE calculation

#This section aim to predict the variable of interest on the basis of your personal data
NomDB_perso <-readline(prompt="What is the name of your personnal database?
(please add the extension of your file)")
testset<-as_tibble(read.table(NomDB_perso, sep="\t",header=TRUE))
summary(testset)
testset
column_testset <- names(testset)
print(paste("Column header of your personnal database",column_testset))
#¨Please indicate in the following line which columns number the code have to consider.
These columns should be the same as used for the model building
used_parameters <- testset[,c(3:15)]
prediction_model <- predict(fit.lm5, newdata = used_parameters)
write.table(prediction_model, file = 'Prediction_data_perso.txt', sep="\t")
prediction_model

#This part is used to do a selection of variables using the Backward stepwise method on
the p-values, with a limit at 0.05
#linear model
BWD.fitlm.p <- ols_step_backward_p(fit.lm, prem=0.05, details=TRUE)
BWD.fitlm.p_results <- c(coeff = BWD.fitlm.p$model$coefficients, R²_ajusté = BWD.fitlm.p$adjr,
RMSE =BWD.fitlm.p$rmse)
write.table(BWD.fitlm.p_results, file ='BWD.lineaire.p_results.txt')
BWD.fitlm_interact.p <- ols_step_backward_p(fit.lm_interact,prem=.05, details=TRUE)
BWD.fitlm_interact.p_results <- c(coeff = BWD.fitlm_interact.p$model$coefficients,
R²_ajusté = BWD.fitlm_interact.p$adjr, RMSE =BWD.fitlm_interact.p$rmse)
write.table(BWD.fitlm_interact.p_results, file ='BWD.lineaire_interact_results.txt')
#log model
BWD.fitlm1.p <- ols_step_backward_p(fit.lm1, prem=0.05, details=TRUE)
BWD.fitlm1.p_results <- c(coeff = BWD.fitlm1.p$model$coefficients, R²_ajusté = BWD.fitlm1.p$adjr,
RMSE =exp(BWD.fitlm1.p$rmse))
write.table(BWD.fitlm1.p_results, file ='BWD.log.p_results.txt')
BWD.fitlm1_interact.p <- ols_step_backward_p(fit.lm1_interact,prem=0.05, details=TRUE)
BWD.fitlm1_interact.p_results <- c(coeff = BWD.fitlm1_interact.p$model$coefficients,
R²_ajusté = BWD.fitlm1_interact.p$adjr, RMSE =exp(BWD.fitlm1_interact.p$rmse)) #calcul RMSE...
write.table(BWD.fitlm1_interact.p_results, file ='BWD.log_interact_results.txt')

#This step can be done for any mathematical transformation
```