Raw material effect on hemicellulose extraction yield and molecular weight during hot pressurized water pretreatment by autohydrolysis

Gianluca Gallina [a], *Álvaro Cabeza* [a], *Henrik Grénman* [b,c*], *Pierdomenico Biasi* [b], *Juan García-Serna* [a*], *Tapio Salmi* [b]

[a] Department of Chemical Engineering and Environmental Technology, High Pressure Processes Group, University of Valladolid, Valladolid, ES-47011, Spain

[b] Johan Gadolin Process Chemistry Centre, Laboratory of Industrial Chemistry and Reaction Engineering, Åbo Akademi University, Biskopsgatan 8, Turku/Åbo, FI-20500, Finland

[c]Molecular process and Materials Technology, Åbo Akademi University, Biskopsgatan 8, Turku/Åbo, FI-20500, Finland

**To whom correspondence should be addressed. E-mail:* jgserna@iq.uva.es (J. García-Serna)

lgrenman@abo.fi (Henrik Grénman)

19

Abstract

A comprehensive study on the hemicellulose extraction from 10 different tree species was performed at 160 ºC using a novel cascade reactor. The aim was to identify which wood species were best candidates to obtain a high concentration, yield and/or molecular weight of hemicelluloses. Hydrothermal extractions at several times (from 5 to 80 min) were performed. We demonstrated that there is a relation between extraction yield (between 9.7 and 40.3%), composition of the raw material and initial structure determined via TGA data. Additionally, a new empirical equation able to estimate the hemicellulose extraction yield from initial composition data was developed. The highest yield was obtained with eucalyptus wood. Molecular weight of the oligomers varied from 3.4 to over 100 kDa. Three trends were observed: molar mass decay with time, maximum and minimum of molar mass. In general, the higher the extraction yield, the lower the molecular mass of the hemicelluloses.

Keywords

Hemicellulose; subcritical; TGA; biomass; hydrothermal pre-treatment; hydrothermal carbonization

1. Introduction

Lignocellulosic biomass, the most abundant on earth, can be obtained from various sources, such as wood, agricultural and municipal waste and other raw materials which do not compete with edible crops for direct human or animal consumption.

Lignocelluloses primarily consist of three pseudo-components, lignin, cellulose and hemicelluloses, combined in a resistant structure; however, the versatile composition, enables the production of various fuels and high value chemicals [1]. Among the various lignocellulosic biomasses, wood is an important renewable resource because of the huge availability of trees and their non-seasonal character; moreover trees do not require intensive use of fertilizers and pesticides to grow, and they generally contain a minor amount of inorganic substances compared to crops [2].

Extraction and applications of hemicellulose

Hemicelluloses can be isolated from biomass in molecular weights above 3 kDa and can be used for multiple applications. The production of films for packaging applications made with hemicelluloses to replace synthetic plastics has been widely studied [3-6]. Another important application is the production of hydrogels used as drug carriers [7, 8] and to adsorb heavy metal ions from aqueous solutions [9] . According to other studies, it seems also that xylans have the potential to be used in medicine as cholesterol depressant, HIV inhibitor and dietary fibers although these studies are still very preliminar [10].

An effective and clean way to extract hemicellulose is to pretreat the lignocellulosic biomass with hot pressurized water [11-15]. At temperatures above 100 ºC, water is able to extract hemicellulose from biomass. The extracted oligomers may undergo hydrolysis in aqueous medium, catalyzed by hydronium ions and acetyl groups originated from hemicellulose [14, 16].

Depending on the type of raw material used, hemicellulose has a different composition: partially acetylated xylans are the predominant hemicelluloses in hardwoods, while galactoglucomannans are the

predominant hemicelluloses in softwoods; hardwoods are therefore an important raw material for obtaining a hemicellulose rich in xylose.

In a previous work carried out by our group, we studied the efficiency of extraction with water at a temperature of 250 ºC from 9 different species of typical trees in the Castilla y Leon region (Spain). In particular, the total yield of extracted sugars, and the inhibitory effects of lignin in the efficiency of the reaction was evaluated [17]. Such temperature was too high, and both hemicellulose and cellulose were co-extracted and degraded.

Optimization and modelling of the extraction process

In this paper, we will focus on the extraction of hemicellulose alone, using optimal conditions defined in other experiments for maximizing yields, without incurring the degradation of sugars [11, 18]. The main hypothesis that we wanted to test was whether there is a clear relationship between biomass structure and composition and the quality of the hemicellulose extracted. Biomass structure was indirectly investigated via TGA model analysis. Both processes (thermal and hydrothermal degradation) involve similar phenomena, like oligomer cleavage [19]. There is therefore a similarity between the change of the structure of the biomass due to hydrothermal and thermal degradation. Furthermore, TGA is a cheap and quick technique that only requires a little amount of sample to be performed (10 mg). Fractionation of wood from 10 different tree species was carried out in a batchwise operated cascade reactor at a constant temperature of 160 ºC with total recirculation. The concentration of hemicellulose extracted from the species was analyzed at different extraction times by calculating and comparing the yields of the extractions. The molecular weights of the oligomers obtained during various extraction times were measured and a direct correlation with the pH of the extracted solution was identified. The content of lignin and cellulose in the various species was also determined to understand if the composition had an influence on the extraction process. With the help of this methodology, an empirical equation for yield of hemicellulose extraction was proposed.

2. Materials and methods

2.1 Materials

Lignocellulosic raw materials used in this study came from urban trees located in Castilla y Leon (Spain).

The tree species studied were: walnut (*Juglans regia*), large leaved linden (*Tilia platyphyllos)*, field elm (*Ulmus minor*), plane (*Platanus x acerifolia*), eucalyptus (*Eucalyptus globulus*), sour cherry (*Prunus cerasus*), catalpa (*Catalpa bignonioides*), maple (*Acer saccharum*), almond (*Prunus dulcis*) and cedar (*Juniperus oxycedrus*). Nine of the wood species were hardwoods, while cedar was the only softwood. Trees had an approximate age of 30-35 years, with an average height of 18-20 m. During a seasonal pruning, the top of the trees was cut and trunk sections with a diameter of about 20 cm and a height of 5 cm were picked up as a raw material for our experiments. Disks were cut in slices and bark was manually removed in order to reduce the content of extractives [20, 21] . The wood was dried, chipped, milled with a Fritsch Universal Cutting Mill Pulverisette 19 (Germany) and sieved to a particle size between 1.25 and 2.00 mm with a Retsch Vibratory Sieve Shaker AS 200 basic.

2.2 Experimental procedure and analytical methods

2.2.1 Determination of pH

The pH of the extracted solution (hydrolysate) was measured with a Phenomenal pH meter using a refillable glass electrode model 221 with a built-in PT 1000 temperature sensor. The measurement was performed at ambient temperature, mixing the solution with a magnetic stirrer.

2.2.2 Liquid samples chemical composition

Liquid samples collected from the experiments were subjected to acid methanolysis [22]. Resorcinol and sorbitol were used as internal standards. First, a certain amount of liquid containing about 0.1 mg of carbohydrates was freeze-dried in vacuum. A methanol-based sugar monomer solution containing a known amount of the sugar monomers, was used to prepare calibration samples.

An equivalent of 2 mL of 2 M HCl/MeOH anhydrous were added to the experimental and calibration samples, heated subsequently to 100 ºC for 3h. After cooling at ambient temperature, 170 µL of pyridine was added to neutralize the excess of acid, together with 1 mL of sorbitol (0.1 mg/mL in MeOH) and 1 mL of resorcinol (0.1 mg/mL in MeOH). The solution was dried under nitrogen gas at 50ºC and then using a vacuum desiccator. The samples were finally silylated using 150 µL of pyridine, 150 µL of hexamethyldisilazane (HMDS) and 70µL of trimethylchlorosilane (TMCS). The derivatised samples were analysed by a gas chromatograph with flame ionization detection [23].

About 1 µL of each silylated sample was injected through a split injector (250 ºC, split ratio 1:25) into a column coated with dimethyl polysiloxane (HP-1, Hewlett Packard). The column length, internal diameter and film thickness were 25 m, 200 µm, and 0.11 µm, respectively. Hydrogen was used as carrier gas with a flow rate of 45 mL/min. The following temperature programme was applied: 100 ºC, 2 ºC/min, 8 min at 170 ºC, 12ºC/min and 7 min at 300 ºC. The identification and quantification of sugars were accomplished through the injection of standard samples and proper calibration.

2.2.3 Molecular weight analysis

Molecular weights of the hemicelluloses extracted were determined by high-performance size-exclusion chromatography (HPSEC) equipped with multiangle laser-light scattering (MALLS) and refractive index (RI) detectors. The columns employed were Ultrahydrogel TM Column, Linear, 10 µm, 7.8 mm X 300 mm, 500 – 10M. The eluent was 0.1M $NaNO_3$ with a flowrate of 0.5 mL/min at 40ºC. Calculations were performed with the software Astra, Wyatt Technology.

2.2.4 TGA analysis

TGA analysis of the raw materials was carried out in a TGA/SDTA RSI analyser of Mettler Toledo. Samples of approximately 10 mg were heated at a rate of 20 ºC/min under N_2 atmosphere (60 N mL/min flow) from a temperature of 50 ºC up to temperatures around 800 ºC.

2.2.5 Raw material characterization

The total amount of extractives, lignin and structural carbohydrates in the raw materials was determined according to the standard methods published by the National Renewable Energy Laboratory (NREL) [24]. Dried biomass was treated with n-hexane in a Soxhlet equipment, in order to remove the extractives. 300 mg of dried and free-extractives solid were hydrolyzed in 3 mL of 72% wt. sulfuric acid solution at 30 °C for 60 min. The mixture obtained was diluted using 84 mL of deionized water and heated at 120 °C for 60 min to hydrolyze oligosaccharides and obtaining their correspondent monomers. Solid was separate from the liquid solution by vacuum filtration, placed in a muffle at 550 °C for 24 h and the remaining residue was weighted before and after this step to calculate the insoluble lignin and the ash content of the sample. A liquid aliquot was analyzed with UV-Vis spectrophotometer at 320 nm with extinction coefficient of 34 $Lg^{-1}cm^{-1}$ [25] to calculate the amount of soluble lignin. Another liquid aliquot was neutralized to pH range 6 to 7, then it was filtered using a 0.2 µm membrane and analyzed by HPLC to determine the carbohydrates composition.

The column used for the separation of the compounds was SUGAR SH-1011 Shodex at 50.0 ºC with a flow rate of 0.80 mL/min, using a solution of 0.01N of sulphuric acid and water Milli-Q as the mobile phase. The sugars and their derivative were identified with Waters IR detector 2414 and Waters dual λ absorbance detector 2,487 (210 nm and 254 nm).

Carbohydrates composition of hemicelluloses contained in the raw material was measured through GC analysis, after subjecting the solids to acid methanolysis: 2 mL of 2M HCl/MeOH anhydrous were added to 10 mg of dry solid and heated to 100 ºC for 5h [22]. Next steps were the same described in paragraph 2.2.2 for liquid samples. As acid methanolysis, at the conditions used in this paper [22], is not strong enough to break cellulose, glucose identified by GC is assumed to proceed from hemicellulose hydrolysis.

Cellulose content in the raw material was calculated related to the glucose content by substracting the glucose detected by GC to the total glucose detected by HPLC.

The calibration reagents used for analysis were: cellobiose (+98%), glucose (+99%), fructose (+99%), glyceraldehyde (95%), pyruvaldehyde (40%), arabinose (+99%), glycolaldehyde (+98%), 5-hydroxymethylfurfural (99%), lactic acid (85%), formic acid (98%), glucuronic acid (99%), mannose (+99%), xylose (+99%), galactose (+99%), rhamnose (+99%), galacturonic acid (+99%), furfural (+99%), acetic acid (+99%), 4-O-methylglucuronic acid (4-O-MeGlcA) (98%) all of them purchased from Sigma-Aldrich and used without further modification.

2.2.6 Experimental set-up and procedure

The experiments were carried out in a batchwise operated cascade reactor composed of five Parr units connected in series, designed and developed at Åbo Akademi [18, 26]. The main advantage of this system was that, during the same experiment, multiple liquid and solid samples at different resident times could be collected, unlike classic batch reactors where only one sample could be collected from one experiment. The nominal volume of each Parr unit was 200 mL. The experimental device is depicted in Figure 1.

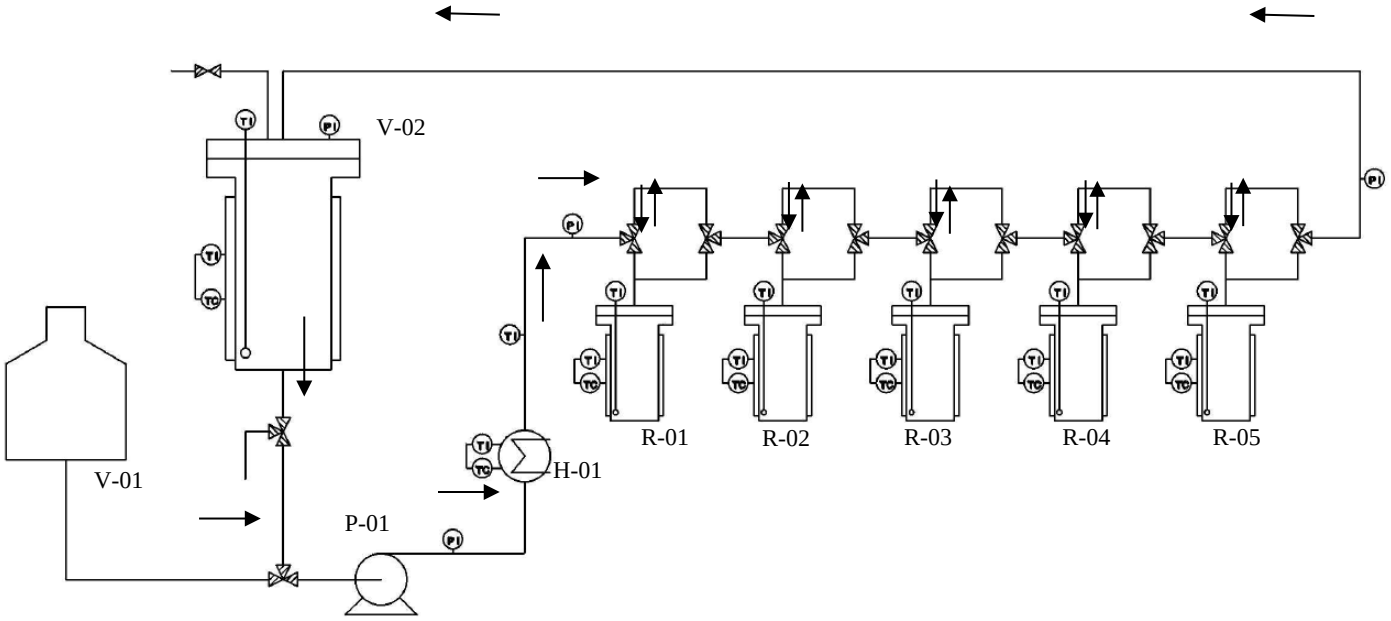


Figure 1. Simplified scheme of the batch cascade reactor used in the experiments. Equipment: V-01 Water tank, V-02 Collector vessel, P-01 Pump, H-01 Heater, R-01/R-05 Reactors.

A flowrate set at 150 $L \cdot h^{-1}$ was fully recirculated through the units and through a buffer vessel with a volume of 2 L, in order to mimic and maintain a constant mixing inside the system. The liquid/solid ratio was approximately 160. A metallic net was located on the top of each of the 5 units to keep the solid particles inside, avoiding them to be washed away by the liquid during the operation.

The pressure was maintained constant at 9 bara (2.9 bar higher than the thermodynamic phase equilibrium of water at the reaction temperature to assure liquid phase) and it was measured before the first reactor and after the last reactor.

Each reactor unit was equipped with an individual heating jacket and the temperature was measured continuously inside and outside each reactor unit and regulated by PID controllers.

The temperature inside the reaction system was fixed for all the experiments to a constant value of 160 ºC. Each reactor unit was loaded with 5 g of dry wood (25 gr in total), filled with distilled water and kept overnight to pre-wet the raw material (swelling); pipes and the 2 L buffer vessel were also filled

with a known amount of water. Before starting the experiments, water was recirculated through the pipes and the 2 L vessel, by-passing the 200 mL units and preheated until reaching the desired reaction temperature (160 ºC).

After the desired temperature was reached in the bypass mode, the heating was turned on also for the five reactors heat jackets and hot water was let to enter into the units, stabilizing rapidly the temperature in the system.

At pre-established sampling times (5 min, 10 min, 20 min, 40 min and 80 min) reactors were sequentially by-passed, quenched rapidly and detached from the system. Figure S1 in supplementary material (Appendix 2) represents the temperature profile inside the system during the experiments. The system took about 4 minutes to reach a constant temperature of 160 ° C, once the set reaction temperature was reached, temperature variations were less than ±2 ºC. The time needed to stop the reaction by cooling down each unit from 160 ° C to 85 ° C was less than one minute.

A liquid sample was obtained from every single reactor unit, with a total of 5 per experiment. The samples were then analyzed as explained in section 2.2. A total of 10 experiments were performed, at constant operational conditions; wood from a different species of tree were tested in each experiment.

3. Results and discussion

The hypothesis that we wanted to demonstrate was that there is a relationship between the initial raw material and the extraction of hemicelluloses from wood using water as a solvent. Furthermore, we wanted to demonstrate that the differences occur not only between hardwood and softwood, but even among different hardwoods the difference is clear. For that we tested 10 different tree species in the reactor (9 hardword and 1 softwood). Temperature was fixed at 160 °C because it both guarantees high hemicellulose yields and minimizes degradation and undesired side products [18].

3.1 Raw materials characterization

Table 1 represents the composition of the raw materials in dry basis, determined as explained in paragraph 2.2.5.

Table 1. Composition of wooden biomass from 10 different tree species dry basis.

Species	Extractives	Cellulose	Lignin	Hemicellulose	Acetic Acid	Mass balance
	g/g	g/g	g/g	g/g	g/g	% Error
Almond	0.071	0.353	0.306	0.261	0.049	-4.0%
Cedar	0.052	0.314	0.398	0.228	0.058	-5.1%
Sour Cherry	0.021	0.431	0.241	0.304	0.060	-5.6%
Elm	0.022	0.541	0.190	0.248	0.030	-3.1%
Eucalyptus	0.013	0.462	0.252	0.260	0.072	-5.9%
Linden	0.014	0.420	0.278	0.215	0.053	1.9%
Maple	0.012	0.299	0.456	0.239	0.046	-5.2%
Plane	0.023	0.341	0.388	0.243	0.058	-5.3%
Walnut	0.005	0.415	0.330	0.254	0.023	-2.7%
Catalpa	0.002	0.495	0.213	0.251	0.073	-3.4%

The species with the highest content of lignin was the maple (0.46 g/g of wood), while the lowest amount of lignin (0.19 g/g of wood) was contained in elm tree wood. Maple amount of lignin was high respect to other studies [17, 27].

In the experiments carried out in this work, only bark was removed from the wood samples, but sapwood and heartwood were not separated. Also knots, formed when removing branches from the trunk were present in the raw material. It is known that heartwood and branches contain a higher amount of lignin compared to sapwood [28, 29] and this may be the reason of the differences in composition

between our work and others in literature. Also the age of the tree is an important factor, as mature trees contain a higher amount of lignin respect to younger trees [30].

The decision of using all the parts of the wood (except the bark, to minimize the extractive content) was made to have a raw material as close as possible to what could be used in a real biorefinery process.

The highest amount of cellulose was found in elm wood (0.54 g/g wood), while maple wood contained the lowest amount (0.30 g/g of wood).

Table 1 shows that the percentage errors in the mass balance are relatively small, always below 5.9%, comparable with other studies [15, 31].

Compounds constituting hemicelluloses from the different raw materials, determined by GC analysis, are compared in Figure 2. Sour cherry contained the highest amount of hemicellulose (0.30 g/g of wood). Xylose was the most abundant component in all tree species analyzed, with a maximum amount in sour cherry wood (0.19 g/g of wood) and a minimum in Cedar (0.07 g/g of wood).

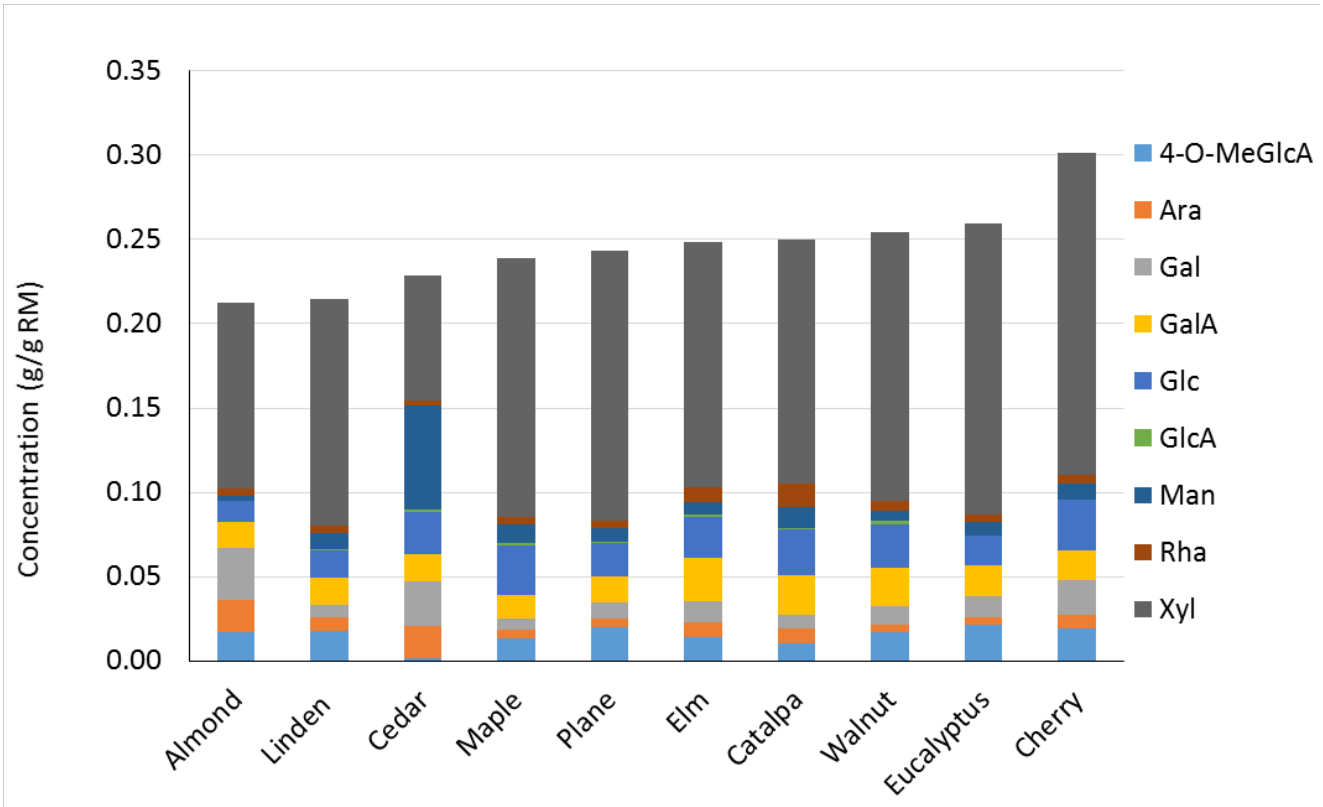


Figure 2. Concentration of 4-O-MeGlcA (4-O-methylglucuronic acid), Ara (arabinose), Gal (galactose), GalA (galacturonic acid), Glc (glucose), GlcA (glucuronic acid), Man (mannose), Rha (rhamnose), Xyl (xylose) in hemicelluloses from 10 different tree species, expressed in g compound/g raw material dry basis.

3.2 Hemicellulose extraction

The yield of hemicellulose extracted after the experiments is represented in Figure 3 as a function of the extraction time. Detailed calculations for the determination of yields and concentration are reported in the supplementary material (Appendix 1): Volumetric concentrations of hemicellulose extracted in liquid phase are represented in Figure S2a in Appendix 2, while a graph representing extracted hemicellulose vs total hemicellulose content is represented in Figure S2b.

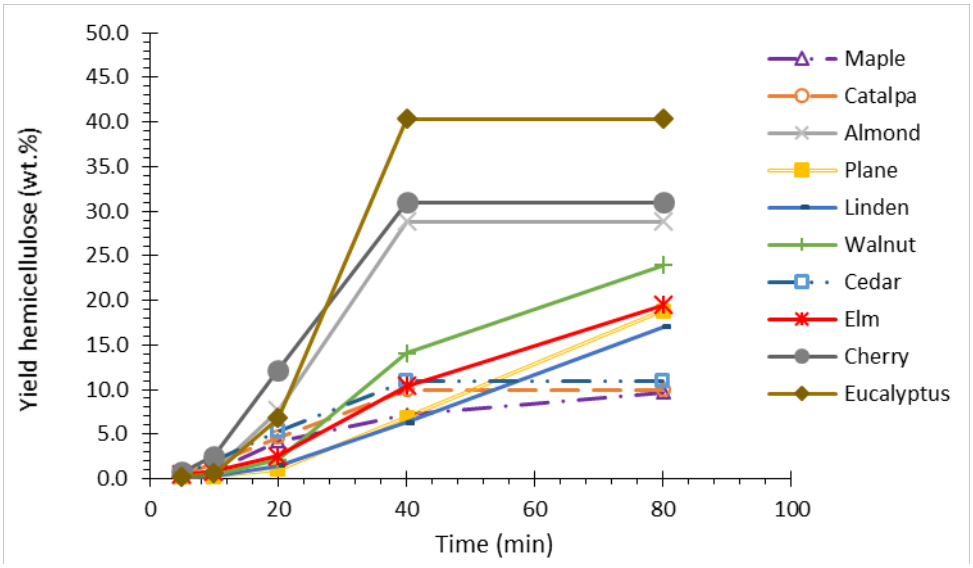


Figure 3. Yield of hemicellulose extracted at different extraction times from the different raw materials.

We found two different behaviors in the extraction curve: (1) walnut, elm, linden and plane showed a continuous almost linear extraction curve during the 80 min, (2) the other species showed slow extraction in the first 20 min, then a rapid increase in the concentration of hemicellulose between time

20 min and 40 min, and finally a plateau from time 40 min till the end. This behavior could be related to the initial sample characteristics and it will be discussed in section 3.4.

The highest yield of hemicellulose extracted and dissolved in the liquid phase was reached using eucalyptus as raw material (40.3%), while the lowest was obtained with maple wood (9.7%). Figure 4 shows that the lowest hemicellulose yields were obtained with species that contained respectively a high amount of cellulose or a high amount of lignin. In particular, catalpa and elm wood contained over 49% of cellulose, and hydrothermal extraction allowed to extract respectively only 9.9% and 12.4% of the whole hemicellulose. On the other hand, maple, cedar and plane wood contained over 38% of lignin, which leaded to hemicellulose yields of 9.7, 10.9 and 18.8%.

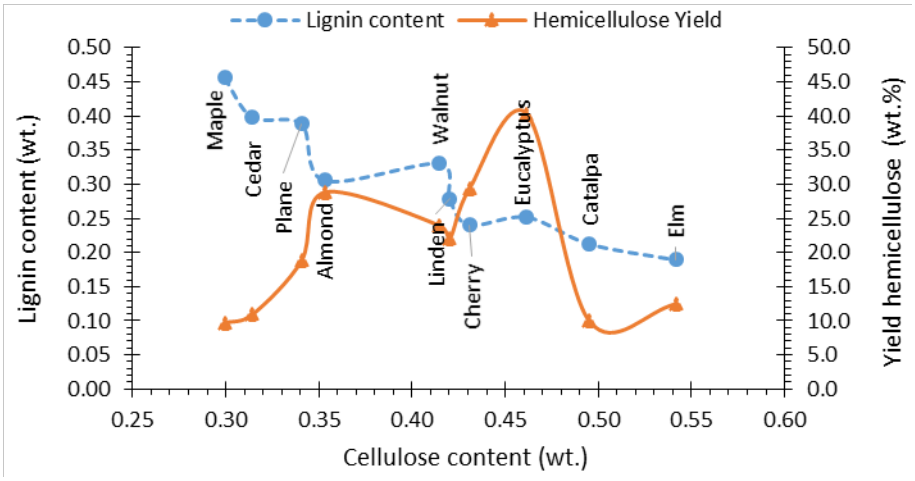


Figure 4. Lignin content in function of the cellulose content; Hemicellulose yield obtained through hydrothermal extraction from 10 tree species with different amount of cellulose and lignin.

It is therefore reasonable to think that the species of trees that contain high amount of lignin or towering amounts of cellulose do not allow achieving high yields in hemicellulose. Lignin and cellulose thus would carry a shield effect that protects hemicellulose and, at the experimental conditions used here, prevents them from breaking through a hydrothermal treatment. Cellulose shield effect was previously observed during the kinetic analysis of biomass isothermal TGA [19], observing that hemicellulose kinetics were enhanced when cellulose degradation started. Regarding lignin, this assumption was confirmed by previous studies, e.g. Yedro et al. [17] also found that a high amount of lignin drove to a

reduction of hemicellulose yield. Moreover, two different kinds of hemicelluloses can be recognized in lignocellulosic biomasses, which differ according to the difficulty of being extracted: one hemicellulose is easy to extract while the other is difficult to recover due to its strong interactions with cellulose and lignin [32-34]. The species that contained intermediate values between those of cellulose and lignin indicated, however, did not show a linear trend in the extraction of hemicellulose, there were indeed fluctuations in the yield values. It can be stated that the composition of the biomass affects the hemicellulose extraction only partially.

It is therefore necessary to study the effect of structure of the plant on the fractionation process, and that is how the three constituent polymers are combined with each other. The study of the histology of the plant is worth studying, although was out of the scope of this study, and for that purpose we envisioned it via a specific modelling tool using thermogravimetric analysis (TGA), as described in section 3.6 [19].

3.3 pH evolution

The evolution of the pH during hot water pretreatments has been analyzed in other works at various temperatures [11, 13, 14, 17]. In this study, we compare the pH evolution in the liquid hydrolysate solutions with hemicellulose extracted from 10 different tree species at 160 ºC. As indicated before, a similar study was carried out by our group, at 250 ºC, where both hemicellulose and cellulose were extracted and in that case degraded [17]. It is well-known that the increase of acidity is directly related to the cleavage of hemicellulose polymers and the release of the structural acetyl group that form acetic acid in the bulk liquid. Futhermore, the protons from acetic acid dissociation catalyze the hydrolysis of hemicellulose oligomers, triggering a chain reaction called autohydrolysis [35, 36].

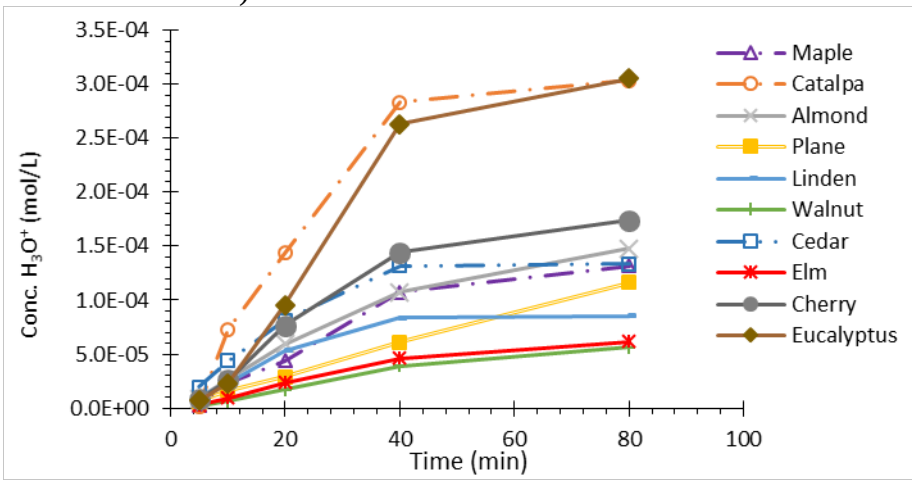
Figure 5a represents the concentration of H_3O^+ ions in the extracted solution as a function of extraction time, for all the experiments. The hydronium concentrations grew rapidly during the first 40 minutes of the process, and then exhibited only minor changes (it is worth remembering that the system works

303 under total recirculation, so the acetic acid released goes through the system during the remaining
304 experiment time).
305 The highest acidity values were achieved by processing wood from catalpa and eucalyptus; the rate of
306 increase in acidity was faster when using catalpa wood. From these results, it can be indirectly deduced
307 that the hemicelluloses of eucalyptus and catalpa contained a high number of acetyl groups, even if the
308 removal from catalpa wood was slightly faster.
309
310 *Acid liberation due to hemicellulose extraction*
311 In this process, we need to consider two different phases, the solid phase and the liquid phase. Many
312 authors have demonstrated that mineral acids enhance hemicellulose hydrolysis. This means that pH of
313 the liquid phase is intentionally decreased (protons increase) by adding an external acid. On the other
314 hand, the solid phase has its own acidic groups, i.e. acetyl groups. We believe, that for our case, where
315 no mineral acid was added, an increase in the amount of hydronium anions in the solution due to the
316 acetyl group cleavage of the extracted hemicellulose cannot be directly correlated to an improvement in
317 the extraction. However, Figure 5b shows very clearly that for every tree species, an increase in the
318 amount of hemicellulose dissolved lead to an increase in the acidity of the solution. This indicates that
319 the increase of acidity in the solution was a result of extraction and hydrolysis of hemicelluloses [26], as
320 acetyl groups bounded to hemicelluloses were cleaved and converted into acetic acid. From our point of
321 view, this indicates that the bonded acetyl groups (still attach to the wood) have a positive effect on the
322 extraction, rather than the protons in solution.
323 To better understand the hydrolysis mechanism, the concentration of acetic acid detected after the
324 characterization of the raw material (representing the total amount of acetyl groups attached to the
325 hemicelluloses) was represented in Table 1 in terms of raw material composition.
326 It is evident that species containing hemicelluloses with the highest amount of acetyl groups (eucalyptus
327 and catalpa) produced a more acidic solution after extracting the hemicelluloses. Conversely, walnut
328 and elm contained the least amount of acetyl groups and released the smallest concentration of H_3O^+

ions after extracting hemicellulose, later we will see that this also affects the molecular weight of dissolved hemicelluloses.

The phenomenon may be due to the fact that the presence of acetyl groups in solid phase catalyzed the hydrolysis of hemicellulose oligomers as long as they solubilized in the liquid phase. Together with the hemicellulose oligomers, also the acetyl groups were detached from the solid, solubilized and converted into acetic acid, lowering the pH of the solution.

a)



b)

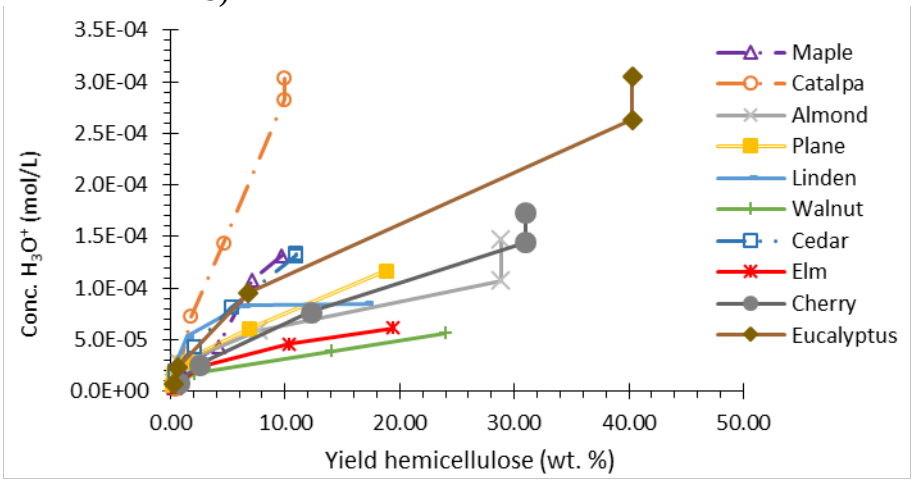


Figure 5. a) Concentration of H_3O^+ ions in the extracted solutions, at different extraction times. b) Hemicellulose yield in function of the concentration of H_3O^+ ions in the extracted solutions.

342

343 **3.4 Molecular weight distribution of the extracted hemicelluloses**

344 Regarding the molar mass of the extracted hemicelluloses we found three different trends, as we
345 demonstrated in Figure 6.

346 *First trend: molar mass decay with time.*

347 This indicates that from the beginning to the end of the process, hemicelluloses were continuously
348 hydrolyzed. Depolymerization was faster during the first 20 minutes and then slowed down (Figure 6a).
349 Hemicellulose yields, on the other hand, increased slowly during the first 20 minutes, then more rapidly
350 between 20 and 40 min and finally tended to stabilize in the last 20 minutes.

351 The possible explication is that, at the beginning of the process, extracted oligomers were long (above
352 30 kDa) and partially soluble (yields below 3%) while, due to the temperature, insolubilized
353 hemicelluloses started to break and became more soluble (first 20 minutes), until they reached a
354 sufficiently low length that they were rapidly solubilized without further hydrolysis (20 to 40 minutes).
355 In the last 20 minutes, almost all the hemicellulose that could be extracted at 160 ºC were already
356 solubilized and its molecular weight decreased slowly due to the protons in the liquid phase.

357

358 *Second trend: maximum of molar mass.*

359 At the beginning, very low molecular masses were produced (below 1 kDa), but as the extraction
360 evolves, the molar mass increases, with a maximum at 20 minutes (Figure 6b). Then hemicellulose
361 chain lengths were reduced until the end of the run.

362 Hemicellulose yields were very low during the first 10 to 20 minutes (a kind of lag time) and then
363 increased linearly.

364 It seems that short molecules had to be removed before large oligomers could be extracted and then
365 depolymerized.

This tendency could be related to the fact that the hemicelluloses of these species had a stronger and more complex structure [32, 37], needing a preliminary cleaving to achieve a sufficiently small size to be solubilized.

For this reason, only a few molecules with small molecular weight were solubilized at the beginning of the reactions. Later on, temperature effect led to the cleavage of the longest oligomers that were subsequently hydrolyzed and solubilized, increasing the concentration in the liquid phase.

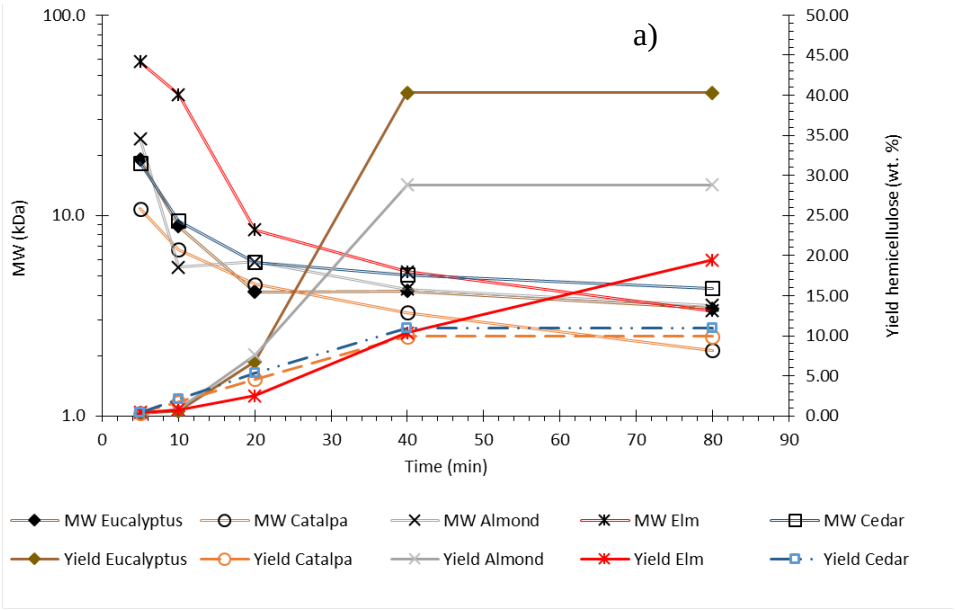
The extraction yield was relatively low in all the cases (below 20% bottom-line).

Third trend: minimum of molar mass.

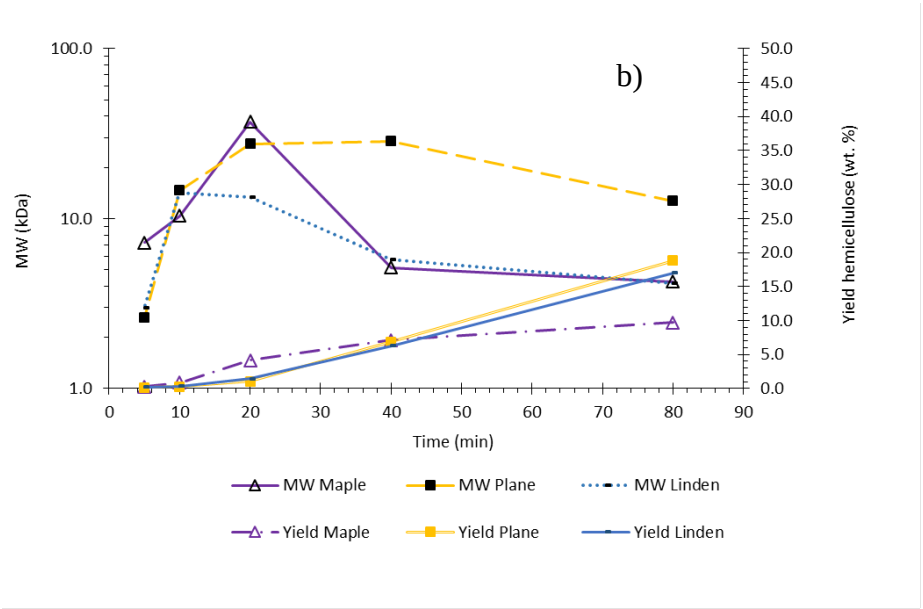
Walnut and cherry revealed a third different behavior (Figure 6c): the molar mass decayed rapidly during the first 20 minutes down to a minimum of around 10 kDa. The curiosity is that from 20 to 80 minutes the average molecular weight grew slightly up to aprox. 15 to 20 kDa.

During the first 10 minutes the extraction yield was very low (below 2-3%), then increased rapidly, and in the final 40 minutes tended to stabilize.

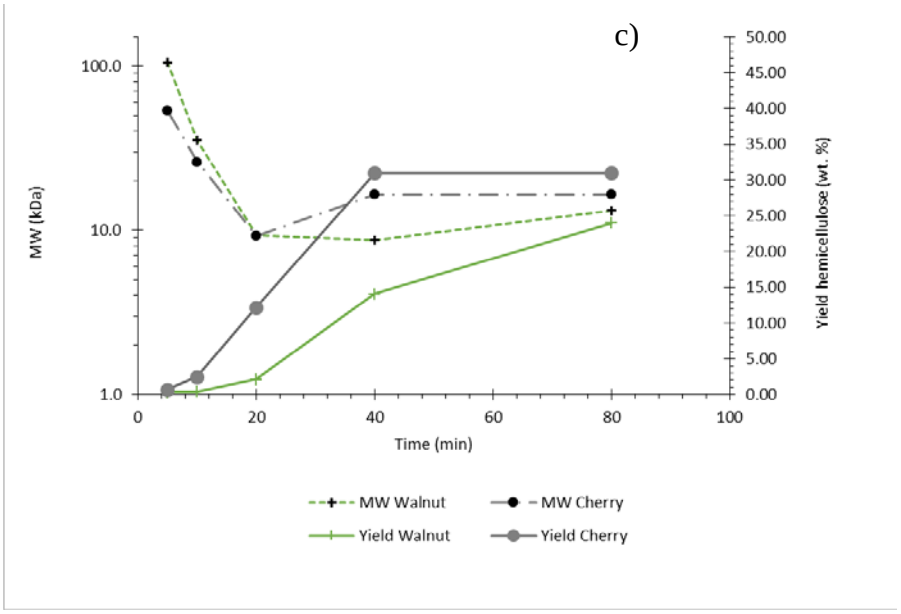
Thus, the behavior was similar to that of Figure 6a at the beginning of the reaction: small amounts of long oligomers were extracted until the longest chains (still attached to the matrix) were hydrolyzed and solubilized. The increase in the molecular weight after 20 minutes could be associated to the presence of non-acetylated hemicelluloses in the matrix, more difficult to hydrolyze and solubilize [37], which appears after the removal of the acetylated hemicelluloses [38-40].


a)
100.0
10.0
1.0
MW (kDa)
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Yield hemicellulose (wt. %)
0
10
20
30
40
50
60
70
80
90
Time (min)
MW Eucalyptus
MW Catalpa
MW Almond
MW Elm
MW Cedar
Yield Eucalyptus
Yield Catalpa
Yield Almond
Yield Elm
Yield Cedar

386


b)
100.0
10.0
1.0
MW (kDa)
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
Yield hemicellulose (wt. %)
0
10
20
30
40
50
60
70
80
90
Time (min)
MW Maple
MW Plane
MW Linden
Yield Maple
Yield Plane
Yield Linden

387


c)
100.0
10.0
1.0
MW (kDa)
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Yield hemicellulose (wt. %)
0
20
40
60
80
100
Time (min)
MW Walnut
MW Cherry
Yield Walnut
Yield Cherry

388

389 **Figure 6.** Molecular weight of hemicellulose oligomers extracted from the different species according
390 to curve shape: a) continuous decay, b) maximum c) minimum
391
392 The hydrolysis process took place initially in the solid phase, where the effect of the temperature broke
393 the bonds between the sugars oligomers, allowing to achieve a sufficiently short length to be solubilized
394 and further hydrolyzed [11].
395 The length of the solubilized oligomers depends on the temperature but also by the concentration of
396 acetyl groups contained in the wood, that catalyze the reduction of hemicelluloses lengths before their
397 solubilization [11].
398 Walnut and elm gave the oligomers with the highest molecular weights and, at the same time, they were
399 the species that contained the least amount of acetyl groups (Table 1). This fact clearly explains how the
400 hydrolysis of hemicelluloses proceeding from these species was not very intense in the solid phase,
401 producing oligomers of large dimensions that were solubilized and subsequently hydrolyzed in the
402 liquid phase.
403 The hemicelluloses belonging to eucalyptus and catalpa, however, were rapidly hydrolyzed in solid
404 phase by the numerous acetyl groups, producing small oligomers that were then solubilized. By the time
405 that we measured the first liquid sample (after 5 min), the hemicelluloses were already short.
406 The size reduction of oligomers in liquid phase was influenced by the increase in the acidity of the
407 extracted solution and the reaction time, as anticipated in paragraph 3.3.
408
409 *Oligomers polydispersity*
410 Dispersity (formerly polydispersity) of extracted oligomers were divided into three groups and
411 represented in Figure S3 in Appendix 2.
412 Hemicelluloses extracted from cedar, cherry and eucalyptus (Figure S3a) displayed the maximum
413 polydispersity at the beginning of the process. In the cases of eucalyptus and cedar, values decreased
414 rapidly during the first 5 minutes and then assumed constant values between 1.5 and 2. Dispersity of

cherry also decreased rapidly during the first 5 minutes (from a value of 6 to 2), it grew to a value around 3 after 20 minutes and then returned to decrease to a value around 1.7.

Dispersity of hemicelluloses extracted from catalpa, almond, plane, linden, walnut and elm (Figure S3b) had a maximum peak at 10 minutes from the beginning of the process, while polydispersity of maple (Figure S3c) had a maximum peak at 20 minutes from the beginning of the process.

Molecular weight distribution of hemicelluloses extracted from all the species are represented in figure S4 in Appendix 2.

3.5 Hemicellulose extraction yield estimation tool development

There is no denying that the initial biomass composition affects hemicelluloses extraction yield as it was demonstrated with the different species in Figure 4. Also molecular weight seems to have a role in this process, as it was explained in section 3.4.

In order to develop an estimation tool, the individual effect of each main biopolymer (hemicelluloses, cellulose and lignin) on the yield must be studied. To do so, the hemicellulose yield was represented versus hemicellulose, cellulose and lignin content (Figure S5 in Appendix 2).

For hemicellulose, a R^2 coefficient of 0.28 was obtained (Figure S5a), which means that there was a weak or no relation between yield of hemicellulose extracted and hemicellulose content in the raw material. However, for cellulose (Figure S5b) and lignin (Figure S5c) the R^2 increased respectively to 0.74 and 0.82 when *catalpa* and *elm* were not considered: these two species are those with the highest content of cellulose.

It appears that the lignin content negatively affected the hemicellulose extraction, while cellulose content promoted it, except when the cellulose content was extremely high.

There should be a relation between extraction yield and the cellulose/lignin content ratio ($w_{C/L}$) as confirmed in Figure 7a. Therefore, the empirical expression to estimate the extraction yield for $w_{C/L}$ would be Eq. 1.

$$y_{HC} = a \cdot ln\left(W_{C/L}\right) + b \quad (1)$$

Nonetheless, this expression was obtained without including the results for *elm* and *catalpa*. To take into account them, Eq.2 was proposed since these two samples had a $w_{C/L}$ >2, which seemed to produce a negative effect on the extraction.

Lignin negative trend could be explained by the fact that it involves the whole biomass, protecting hemicellulose and making the extraction difficult [17].

However, a high amount of cellulose implies a low lignin content, promoting the extraction of hemicellulose. Nevertheless, if lignin content is much low, the hemicellulose could be incorporated and protected among the cellulose fibers [32], which prevent hemicellulose extraction.

$$y_{HC} = a \cdot ln\left(W_{C/L}\right) + b - \frac{c_y}{1 + e^{e \cdot (d - W_{C/L})}} \quad (2)$$

Finally, the differences between the experimental and simulated yields were minimized by the Solver Excel tool to obtain the parameters c_y, d and e. The final values were a=23.33 (dimensionless), b=19.54 (%wt), c_y=31.58 (%wt), d=2.25 (%wt) and e=34.45 (dimensionless). The average error was 12.7 % with a R^2 of 0.83. The comparison between the simulated and experimental values are depicted in Figure 7b. Additionally, the individual discrepancies are arrayed in the Table 1S in appendix 3. It is worth mentioning that the highest errors between calculated and the experimental data were obtained for: *cedar* (28%), *eucalyptus* (16%), *linden* (32%) and *almond* (20%).

Since these discrepancies cannot be explained only by considering differences in composition, molecular weight or protons releasing, another parameter should be considered: the biomass structure. Other studies based on SEM analysis demonstrated that when monomers and oligomers are detached from the biomass, the number of cavities in the matrix increases [13, 41, 42], promoting the removal of further carbohydrates. This kinetics can be assumed as autocatalytic as demonstrated in previous studies

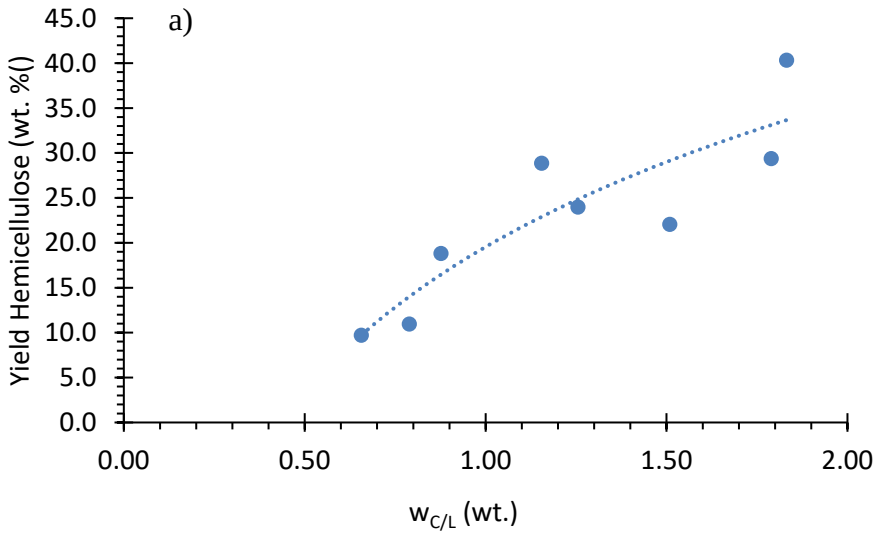
[13, 32]. The same considerations can be applied for the slow pyrolysis process since biomass thermal degradation follows a slow rate until a certain point is reached, where the mass variation becomes abrupt. This behavior can be checked in Figure S6 in Appendix 2, where the thermal degradation of almond wood during a TGA is showed. It can be seen that there is a slow mass change in the sample until, when a temperature of 250 °C is reached, 50% of mass is suddenly lost. This change in tendency was associated to the cleaving of the strongest biopolymer structures of biomass (i.e. cellulose and lignin). Thereby, our assumption is that the structure modifications showed during a TGA can give important information about what happens during the hydrothermal treatment.

Table 2. Cellulose and lignin content

	$w_{C/L}$ [a]	w_T [b]
	-	%wt
Almond	1.16	0.66
Cedar	0.79	0.71
Sour Cherry	1.79	0.67
Elm	2.86	0.73
Eucalyptus	1.83	0.71
Linden	1.51	0.70
Maple	0.66	0.75
Plane	0.88	0.73
Walnut	1.26	0.75
Catalpa	2.33	0.71

[a] Ratio between the cellulose and lignin content: w_C/w_L

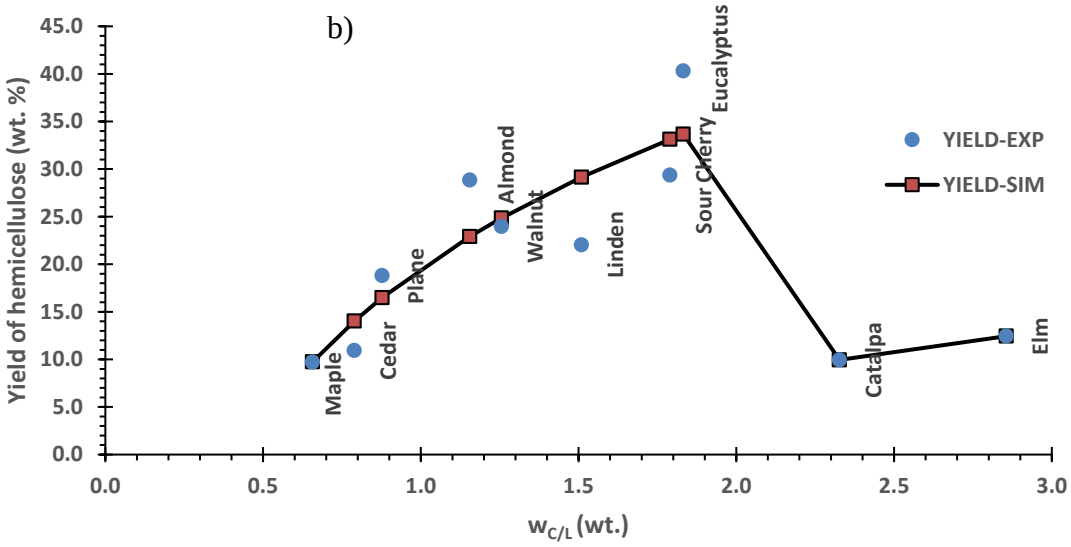
[b] Total amount of lignin and cellulose: $w_C + w_L$



476

477

478



479

480 **Figure 7.** a) Relationship between the hemicellulose extraction yield and the $w_{C/L}$; b) comparison between
481 the experimental and simulated extraction yield.

482

483 **3.5.1 Estimation tool validation**

484 Once the estimation tool has been developed, it was interesting to evaluate its effectiveness with data
485 obtained from experiments carried at the same temperature and residence time as the one presented in
486 this manuscript.

- Holm oak

The yield and composition data were taken from *Yedro et al.*[43]. In this work, holm oak samples were treated in the same cascade reactor that has been used in our study, obtaining at the same conditions a hemicellulose final yield of 28.50 %. In this case, the value of $w_{C/L}$ was 1.74, which drove to an estimated yield of 32.57 % (deviation estimation of +14.2%).

- Extracted grape seeds

In this test the data were collected from *Yedro et al.* [44] who hydrothermally treated grape seeds after a previous extraction with a mixture of ethanol-water. Final hemicellulose yield reached was 3.56%. For this sample $w_{C/L}$ was 0.51 and the estimated yield was 3.95% (deviation estimation of +10.9%).

- Sugarcane bagasse

Data for the hydrothermal extraction of sugarcane bagasse (final yield of 40.4%) were picked up from *Santucci et al.[45]*. The value for $w_{C/L}$ was 1.91 and the calculated yield was 34.67% (deviation estimation of -14.2%).

- Corn straw

In this case the data were taken from [46], being the experimental yield 6.53% ($w_{C/L}$ = 2.41) and the estimated value 8.61% (deviation estimation of +31.9%).

- Rice straw

Finally, data for hydrothermal extraction of rice straw were obtained from [47], being the experimental yield 6.25% ($w_{C/L}$ = 2.84) and the estimated value 12.32% (deviation estimation of +97.1%).

In summary, the proposed equation can reproduce the hemicellulose extraction yield at 160 º C for 80 min for several types of biomasses like seeds, woods and agricultural wastes. However, when samples very different from wood are used, bigger discrepancies are obtained, such as for corn and rice straw. Therefore, the proposed expression provides a good estimation of the expected hemicellulose extraction just by using the rough content of lignin and cellulose from the routine analysis.

3.6 Structural effect

As it was mentioned in the previous section, structural differences seem to be the main reason why there was not a perfect relationship between composition and hemicellulose extraction yield. Additionally, it was also introduced the idea that TGA data can be used to understand this structural effect.

With that purpose, TGA was performed for all the biomass samples and fitted by a kinetic model previously developed by our research group [19]. This model (Eq. 3) was obtained applying a transient mass balance for each compound in both phases, liquid and solid. The liquid refers to the water and organic substances that can be present in the sample. The reaction pathway was based on a modification of the Waterloo's mechanism [48-50]. Therefore, it was assumed that each compound present in the sample (e.g. hemicellulose, cellulose and lignin) decomposed into charcoal and gases during the slow pyrolysis process. The charcoal can in turn be volatilized.

$$\frac{dm_j}{dt} = r_j = \sum_{i=1}^{N_r} g_{ij} \cdot r_i \tag{3}$$

Regarding kinetics, two different types were used: one for liquid phase and another for the solid. For the liquid, it was calculated by a conventional mass transfer expression modified to consider the effect of the sample mass reduction (Eq. 4). In this equation, C_j^* is the equilibrium concentration in the gas phase, which was obtained by the assumption of ideal gas behavior and using a modified Antoine's pressure vapor expression.

$$r_i = h \cdot \left(C_j^*\right) \cdot m_j^{nl_i} \tag{4}$$

For the solid, a first order autocatalytic expression was considered (Eq.5). The parameter α_i represent the initial velocity factor and reflects how difficult it is to degrade the sample. Its value was fixed to 0.99.

534 β_i is the acceleration factor and represents how fast degradation is after it has started. This equation was
535 modified for cellulose with another parameter (c) to consider the effect of the heating rate in thermal
536 degradation (Eq. 6). In this work, the value for parameter "c" was fixed at 0.006 because it was the
537 obtained value for those samples where extractives were also present [19], making this study as general
538 as possible.

$$r_i = k_i \cdot e^{-\frac{E_{a_i}}{R \cdot T}} \cdot \mathrm{m_j} \cdot \left(1 - \alpha_i \cdot m_j\right)^{\beta_i} \quad (5)$$

$$r_i = k_\mathrm{i} \cdot e^{-\frac{E_{a_i}}{R \cdot T} + c \cdot T + \ln(T)} \cdot \mathrm{m_j} \cdot \left(1 - \alpha_i \cdot m_j\right)^{\beta_i} \quad (6)$$

539

540 The fitting was done by the Simplex Nelder-Mead's method, solving the system of ordinary differential
541 equations (ODEs) by the Runge-Kutta's method with 8th order of convergence [51]. The objective
542 function selected was the Absolute Average Deviation (AAD) defined as follows:

$$AAD = \sum_{i=1}^{N} \frac{1}{N} \cdot \frac{\left|x_{i_{EXP}} - x_{i_{SIM}}\right|}{x_{i_{EXP}}} \quad (7)$$

543

544 The ADD and the kinetic parameters for each experiment are included in the supplementary material
545 (Appendix 3), respectively in the Table 1S and in the Table 2S. The model was suitable to reproduce the
546 experimental behavior observed during the TGA with an average error of 1.2 %. Since the discrepancy
547 between the experimental and simulated behavior is low, this calculated kinetics can be used to study
548 how hydrothermal extraction yield is affected by sample characteristics (composition and structure).

549

550

551

552

553

3.6.1 Analysis of the TGA kinetics

Slow pyrolysis is affected by a huge set of different variables like the solid and gas residence time, the temperature range, the heating rate, the final temperature, the sample size and the atmosphere type [52, 53]. Moreover, biomass internal structure can also affect it. A good example can be found in *Cabeza et al.* [19] where the pyrolytic behavior of alkaline lignin and lignin extracted from a real biomass were compared. The result was that, while their qualitative response was similar (degradation between 200 °C and 500 °C), the mass variation and the kinetics were completely different, relating this changes to the structural (chemical or physical) differences between the two samples. Following this idea, the kinetics of the samples that in figure 7b showed the highest deviation from the prediction made by the composition (*cedar, eucalyptus*, *linden* and *almond*) were studied in order to check the effect of the biomass structure.

The changes in the individual kinetics (volatilization, char production and char volatilization) for these 4 samples were analyzed. Their values were compared with those other biomasses with a similar cellulose-lignin ratio ($w_{C/L}$) but that were better simulated by the proposed model. Defining a percentage difference as: $\Delta K=(K_{Studied}-K_{Reference})/ K_{Reference} \cdot 100$.

Kinetic of eucalyptus *(*$w_{C/L}$ = 1.83*)* was compared with kinetics of cherry ($w_{C/L}$ = 1.79); kinetic of cedar ($w_{C/L}$ = 0.78) was compared with kinetic of plane ($w_{C/L}$ = 0.88); kinetics of linden ($w_{C/L}$ = 1.51) and kinetic of almond ($w_{C/L}$ = 1.16) were compared with the kinetic of walnut ($w_{C/L}$ = 1.25).

Only reactions with significant deviations have been considered and represented.

Figure 8 shows that for *cherry* and *eucalyptus*, the main differences are in the kinetics of cellulose and lignin volatilization as eucalyptus presents higher values in both cases: lignin and cellulose of eucalyptus are thus easier to volatilize respect to the ones of cherry.

Regarding *cedar* and *plane* (Figure 9), the highest differences were in lignin volatilization: lignin belonging to cedar is easier to volatilize.

Therefore, it seems that the easier lignin volatilization, the bigger hydrothermal extraction yield was obtained.

However, this statement was not fully true for *almond* (Figure 10.a) since it had lower volatilization kinetics than *walnut* and *linden* but its yield was slightly higher (around 5-6% more). This discrepancy may be justified by the char production kinetics, being much lower for *almond* (Figure 10b).

Char production means that the sample is going to have a higher content of carbon, making it very difficult to volatilize during a pyrolysis [53]. Thus, a slower production of char would imply a larger amount of compounds with a lower carbon ratio in the biomass, promoting volatilization although the volatilization kinetics are lower, and explaining why *almond* had a higher yield.

A similar behavior can be observed between *eucalyptus* (maximum yield, 40.30%) and *maple* (minimum yield, 9.70%). In this case, the lignin volatilization for maple was up to 95% higher but char formation kinetic was also 85% bigger, which are similar results to the kinetic differences of *walnut* and *linden* with *almond*.

On the other hand, it should be mentioned that in all the cases the differences between the kinetics were lower and lower when temperature was raised. A behavior that may be explained by the exponential dependence of the kinetics with temperature.

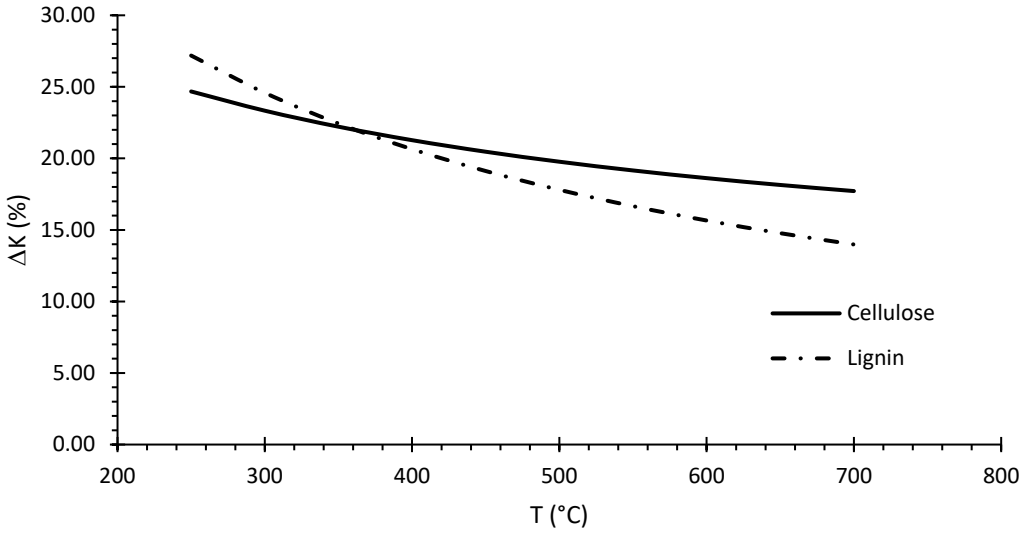


Figure 8. Kinetic differences between *eucalyptus* and *cherry* for cellulose and lignin volatilization. ΔK defined as the difference between the kinetics of eucalyptus and cherry for cellulose and lignin volatilization.

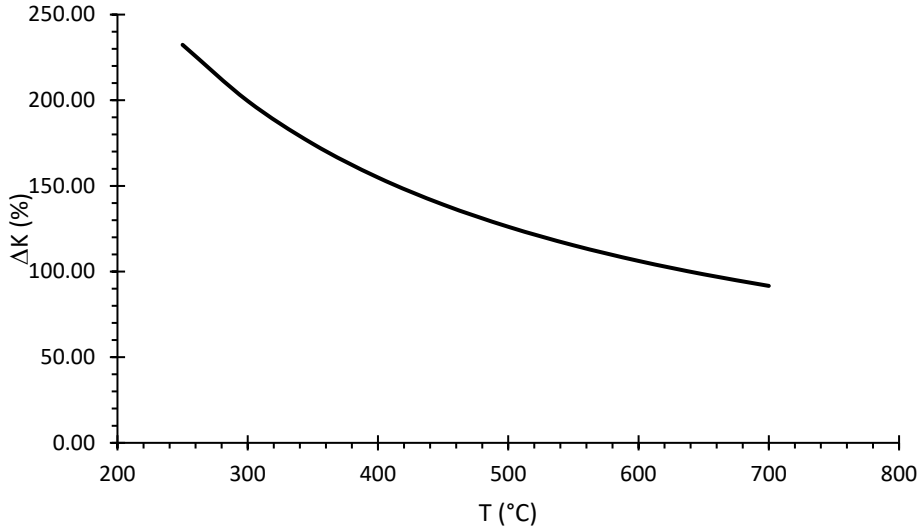
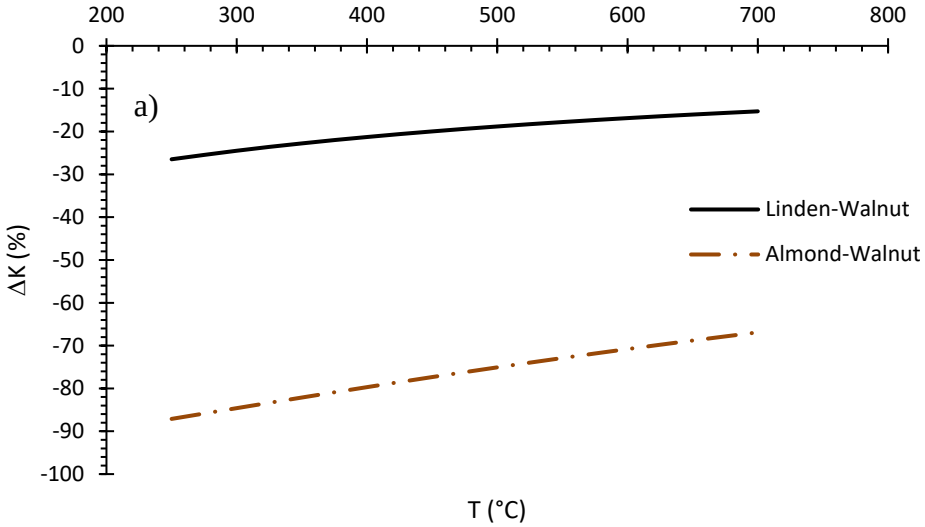


Figure 9. Kinetic differences between *cedar* and *plane* for lignin volatilization. ΔK defined as the difference between the kinetics of *cedar* and *plane* for lignin volatilization.



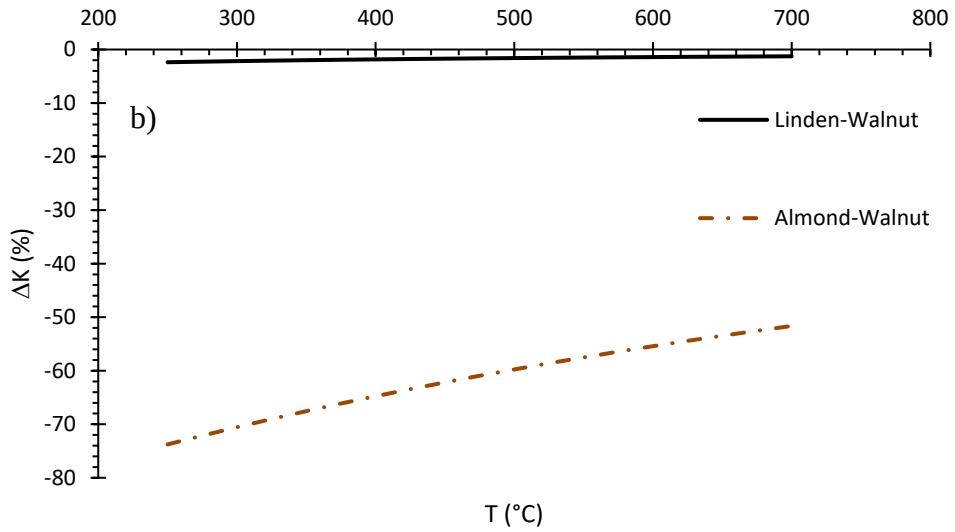


Figure 10. Kinetic differences between *linden and walnut* and *almond and walnut* for lignin volatilization (a) and lignin char production (b) ΔK defined as the difference between the kinetics of *linden and walnut;* and of *almond and walnut* for lignin volatilization (a) and lignin carbonization (b).

4. Conclusions

The influence of composition and structure on the hemicellulose extraction yield was confirmed not only by experimental results but also by the analysis of the kinetics from a TGA.

We have demonstrated that the acetyl groups contained in the wood attached to the hemicelluloses cause hydrolysis both in the solid and in the liquid phase. When the acetyl groups are still attached to the solid cause severe degradation of the hemicellulose polymer chain and only short hemicelluloses are produced (e.g. catalpa). On the contrary, cases like walnut and cherry with very low initial acetyl groups release the hemicellulose with higher molecular weight. Thus, hemicelluloses with molecular weights higher than 60 kDa can be obtained from walnut wood after 5 min of extraction; while lower molecular weights than 10 kDa, more feasible for the conversion into monomers, can be obtained from the fractionation of woods such as catalpa and eucalyptus. Additionally, 3 different trends (decreasing, maximum and minimum) for molecular weight were observed.

623

624 With the method use in this work, species with a high content of lignin or a towering cellulose content
625 led to low hemicellulose extraction yields (around 10%). Structural effects were noticed in some
626 species, like eucalyptus, where the extraction yield (40%) is far greater than the expected taking into
627 account the composition alone.
628 Additionally, a tool capable to estimate the value of the extraction yield from initial composition data
629 was proposed (average deviation of 12.7%). Moreover, it could be concluded that those sample with
630 higher volatilization kinetics for lignin during a TGA would imply a bigger hemicellulose extraction if
631 lignin char formation kinetics is not also promoted. In general, what has been said so far about the
632 kinetics obtained through the TGA, while not providing visual data, helps to understand the behavior of
633 the various constituents of the biomass and their interdependence during the fractionation process.
634

635 **Acknowledgements**
636 The authors acknowledge the Spanish Economy and Competitiveness Ministry (MINECO) and FEDER
637 Funds, Project CTQ2015-64892-R (BioFraHynery). MEng. Gianluca Gallina wish to acknowledge the
638 Spanish Economy and Competitiveness Ministry for the scholarship/predoctoral contract BES-2013-
639 063556. Alvaro Cabeza also wish to acknowledge the Spanish Economy and Competitiveness Ministry
640 for the scholarship/predoctoral contract FPU2013/01516.

641 **Abbreviations and symbols**

642 *Acronyms*

643 AAD: average absolute deviation

644 ODE: ordinary differential equation

645 TGA: thermogravimetric analysis

646

Greek letters

α_i: initial reaction rate factor, dimensionless

β_i: acceleration factor, dimensionless

Symbols

a, b, c_y, d, e: parameters of the empirical equation to estimate the extraction yield

c: heating rate correction factor, dimensionless

C_j^*: equilibrium concentration of the compound "j" in the gas phase, mol/L

$\frac{E_{a_i}}{R}$: activation energy for the reaction "i", K

g_{ij}: stoichiometric coefficient of the compound "j" for the reaction "i", $g \cdot g^{-1}$

h: mass transfer coefficient between the liquid and the gas phases, $g \cdot m \cdot min^{-1} \cdot mol^{-1}$

k_i: Arrhenius' pre-exponential factor for the reaction "i", min^{-1}

K_i: Arrhenius' kinetic constant min^{-1}

m_j: mass fraction of the compound "j", g/g

N: number of experiments, dimensionless

N_r: reaction number, dimensionless

nl_i: mass transfer order for the reaction "i", dimensionless

r_i: reaction rate for the reaction "i", $g \cdot g^{-1} \cdot min^{-1}$

r_j: reaction rate of the compound "j", $g \cdot g^{-1} \cdot min^{-1}$

667 t: operating time, min

668 T: operating temperature, K

669 w_C: cellulose content, g·g^{-1}

670 $w_{C/L}$: ratio of cellulose-lignin content, dimensionless

671 w_L: lignin content, g·g^{-1}

672 w_T: lignin and cellulose composition, g·g^{-1}

673 $x_{i_{EXP}}$: experimental value of the variable "X" in the experiment "i"

674 $x_{i_{SIM}}$: experimental value of the variable "X" in the experiment "i"

675 y_{HC}: estimated hemicellulose extraction yield, %wt

676

677

678

679 **References**

680

681 [1] S.D. King, The future of industrial biorefineries, in: W.E. Forum (Ed.), 2010.

682 [2] P. Kilpeläinen, V. Kitunen, J. Hemming, A. Pranovich, H. Ilvesniemi, S. Willför, Pressurized hot
683 water flow-through extraction of birch sawdust - Effects of sawdust density and sawdust size, Nordic Pulp
684 and Paper Research Journal, 29 (2014) 547-556.

685 [3] S. Sabiha-Hanim, A.M. Siti-Norsafurah, Physical Properties of Hemicellulose Films from Sugarcane
686 Bagasse, Procedia Engineering, 42 (2012) 1390-1395.

687 [4] N.M.L. Hansen, D. Plackett, Sustainable Films and Coatings from Hemicelluloses: A Review,
688 Biomacromolecules, 9 (2008) 1493-1505.

689 [5] H. Jiang, Q. Chen, J. Ge, Y. Zhang, Efficient extraction and characterization of polymeric
690 hemicelluloses from hybrid poplar, Carbohydrate Polymers, 101 (2014) 1005-1012.
691 [6] A. Svärd, E. Brännvall, U. Edlund, Rapeseed straw as a renewable source of hemicelluloses:
692 Extraction, characterization and film formation, Carbohydrate Polymers, 133 (2015) 179-186.
693 [7] Q. Ye, X.F. Sun, Z.X. Jing, G.Z. Wang, Y.J. Li, Preparation of pH-sensitive hydrogels based on
694 hemicellulose and its drug release property, Xiandai Huagong/Modern Chemical Industry, 32 (2012) 62-
695 66.
696 [8] A.T. Neffe, C. Wischke, M. Racheva, A. Lendlein, Progress in biopolymer-based biomaterials and
697 their application in controlled drug delivery, Expert Review of Medical Devices, 10 (2013) 813-833.
698 [9] X.-W. Peng, L.-X. Zhong, J.-L. Ren, R.-C. Sun, Highly Effective Adsorption of Heavy Metal Ions
699 from Aqueous Solutions by Macroporous Xylan-Rich Hemicelluloses-Based Hydrogel, Journal of
700 Agricultural and Food Chemistry, 60 (2012) 3909-3916.
701 [10] P. Gatenholm, M. Tenkanen, Hemicelluloses: Science and Technology, in: Hemicelluloses: Science
702 and Technology, American Chemical Society, 2003, pp. i-v.
703 [11] J.V. Rissanen, H. Grénman, C. Xu, S. Willför, D.Y. Murzin, T. Salmi, Obtaining spruce
704 hemicelluloses of desired molar mass by using pressurized hot water extraction, ChemSusChem, 7 (2014)
705 2947-2953.
706 [12] G. Garrote, H. Domínguez, J.C. Parajó, Study on the deacetylation of hemicelluloses during the
707 hydrothermal processing of Eucalyptus wood, Holz als Roh- und Werkstoff, 59 (2001) 53-59.
708 [13] G. Gallina, Á. Cabeza, P. Biasi, J. García-Serna, Optimal conditions for hemicelluloses extraction
709 from Eucalyptus globulus wood: hydrothermal treatment in a semi-continuous reactor, Fuel Processing
710 Technology, 148 (2016) 350-360.
711 [14] J. Krogell, E. Korotkova, K. Eränen, A. Pranovich, T. Salmi, D. Murzin, S. Willför, Intensification
712 of hemicellulose hot-water extraction from spruce wood in a batch extractor – Effects of wood particle
713 size, Bioresource Technology, 143 (2013) 212-220.

714 [15] C.M. Piqueras, Á. Cabeza, G. Gallina, D.A. Cantero, J. García-Serna, M.J. Cocero, Online integrated
715 fractionation-hydrolysis of lignocellulosic biomass using sub- and supercritical water, Chemical
716 Engineering Journal, 308 (2017) 110-125.
717 [16] A. Romaní, G. Garrote, F. López, J.C. Parajó, Eucalyptus globulus wood fractionation by
718 autohydrolysis and organosolv delignification, Bioresource Technology, 102 (2011) 5896-5904.
719 [17] F.M. Yedro, D.A. Cantero, M. Pascual, J. García-Serna, M.J. Cocero, Hydrothermal fractionation of
720 woody biomass: Lignin effect on sugars recovery, Bioresource Technology, 191 (2015) 124-132.
721 [18] J.V. Rissanen, H. Grénman, S. Willför, D.Y. Murzin, T. Salmi, Spruce Hemicellulose for Chemicals
722 Using Aqueous Extraction: Kinetics, Mass Transfer, and Modeling, Industrial & Engineering Chemistry
723 Research, 53 (2014) 6341-6350.
724 [19] A. Cabeza, F. Sobrón, F.M. Yedro, J. García-Serna, Autocatalytic kinetic model for
725 thermogravimetric analysis and composition estimation of biomass and polymeric fractions, Fuel, 148
726 (2015) 212-225.
727 [20] M.C. Morais, H. Pereira, Variation of extractives content in heartwood and sapwood of Eucalyptus
728 globulus trees, Wood Science and Technology, 46 (2012) 709-719.
729 [21] R.M.A. Domingues, G.D.A. Sousa, C.S.R. Freire, A.J.D. Silvestre, C.P. Neto, Eucalyptus globulus
730 biomass residues from pulping industry as a source of high value triterpenic compounds, Industrial Crops
731 and Products, 31 (2010) 65-70.
732 [22] A. Sundberg, K. Sundberg, C. Lillandt, B. Holmbom, Determination of hemicelluloses and pectins
733 in wood and pulp fibres by acid methanolysis and gas chromatography, Nordic Pulp and Paper Research
734 Journal, 11 (1996) 216-219+226.
735 [23] S. Willför, A. Pranovich, T. Tamminen, J. Puls, C. Laine, A. Suurnäkki, B. Saake, K. Uotila, H.
736 Simolin, J. Hemming, B. Holmbom, Carbohydrate analysis of plant materials with uronic acid-containing
737 polysaccharides-A comparison between different hydrolysis and subsequent chromatographic analytical
738 techniques, Industrial Crops and Products, 29 (2009) 571-580.

739 [24] A. Sluiter, B. Hames, R. Ruiz, C. Scarlata, J. Sluiter, D. Templeton, D. Crocker, Determination of
740 structural carbohydrates and lignin in biomass, in: Laboratory Analytical Procedure (LAP), 2008.
741 [25] S.-N. Sun, X.-F. Cao, H.-Y. Li, F. Xu, R.-C. Sun, Structural characterization of residual
742 hemicelluloses from hydrothermal pretreated Eucalyptus fiber, International Journal of Biological
743 Macromolecules, 69 (2014) 158-164.
744 [26] H. Grénman, K. Eränen, J. Krogell, S. Willför, T. Salmi, D.Y. Murzin, Kinetics of Aqueous
745 Extraction of Hemicelluloses from Spruce in an Intensified Reactor System, Industrial & Engineering
746 Chemistry Research, 50 (2011) 3818-3828.
747 [27] W.E. Kaar, D.L. Brink, Summative analysis of nine common north American woods, Journal of
748 Wood Chemistry and Technology, 11 (1991) 479-494.
749 [28] P. Zhang, F. Wu, X. Kang, Chemical properties of wood are under stronger genetic control than
750 growth traits in Populus tomentosa Carr, Annals of Forest Science, 72 (2015) 89-97.
751 [29] B. Košíková, Morphological and chemical characteristics of stem and knot poplar wood, Wood
752 Research, 54 (2009) 117-122.
753 [30] S. Adamopoulos, E. Voulgaridis, C. Passialis, Variation of certain chemical properties within the
754 stemwood of black locust (Robinia pseudoacacia L.), Holz als Roh- und Werkstoff, 63 (2005) 327-333.
755 [31] P.O. Kilpeläinen, S.S. Hautala, O.O. Byman, L.J. Tanner, R.I. Korpinen, M.K.J. Lillandt, A.V.
756 Pranovich, V.H. Kitunen, S.M. Willför, H.S. Ilvesniemi, Pressurized hot water flow-through extraction
757 system scale up from the laboratory to the pilot scale, Green Chemistry, 16 (2014) 3186-3194.
758 [32] A. Cabeza, C.M. Piqueras, F. Sobrón, J. García-Serna, Modeling of biomass fractionation in a lab-
759 scale biorefinery: Solubilization of hemicellulose and cellulose from holm oak wood using subcritical
760 water, Bioresource Technology, 200 (2016) 90-102.
761 [33] C. Wyman, S. Decker, M. Himmel, J. Brady, C. Skopec, L. Viikari, Hydrolysis of Cellulose and
762 Hemicellulose, in: Polysaccharides, CRC Press, 2004.
763 [34] D. Klemm, H.-P. Schmauder, T. Heinze, Cellulose, in: Biopolymers Online, Wiley-VCH Verlag
764 GmbH & Co. KGaA, 2005.

765 [35] G. Garrote, H. Domínguez, J.C. Parajó, Generation of xylose solutions from Eucalyptus globulus
766 wood by autohydrolysis–posthydrolysis processes: posthydrolysis kinetics, Bioresource Technology, 79
767 (2001) 155-164.
768 [36] T.V. Ojumu, B.a.E. AttahDaniel, E. Betiku, B.O. Solomon, Auto-hydrolysis of lignocellulosics under
769 extremely low sulphuric acid and high temperature conditions in batch reactor, Biotechnol. Bioprocess
770 Eng., 8 (2003) 291-293.
771 [37] A. Cabeza, F. Sobrón, F.M. Yedro, J. García-Serna, Two-phase modelling and simulation of the
772 hydrothermal fractionation of holm oak in a packed bed reactor with hot pressurized water, Chemical
773 Engineering Science, 138 (2015) 59-70.
774 [38] B.A. Miller-Chou, J.L. Koenig, A review of polymer dissolution, Progress in Polymer Science
775 (Oxford), 28 (2003) 1223-1270.
776 [39] A. Kruse, E. Dinjus, Hot compressed water as reaction medium and reactant. Properties and synthesis
777 reactions, Journal of Supercritical Fluids, 39 (2007) 362-380.
778 [40] C.C. Teo, S.N. Tan, J.W.H. Yong, C.S. Hew, E.S. Ong, Pressurized hot water extraction (PHWE),
779 Journal of Chromatography A, 1217 (2010) 2484-2494.
780 [41] M.A. Lima, G.B. Lavorente, H.K. da Silva, J. Bragatto, C.A. Rezende, O.D. Bernardinelli, E.R.
781 deAzevedo, L.D. Gomez, S.J. McQueen-Mason, C.A. Labate, I. Polikarpov, Effects of pretreatment on
782 morphology, chemical composition and enzymatic digestibility of eucalyptus bark: a potentially valuable
783 source of fermentable sugars for biofuel production – part 1, Biotechnology for Biofuels, 6 (2013) 75.
784 [42] R. De Oliveira Moutta, M.C. Silva, R.C. Novaes Reis Corrales, M.A. Santos Cerullo, V. Santana
785 Ferreira-Leitão, E. Pinto da Silva Bon, Comparative Response and Structural Characterization of
786 Sugarcane Bagasse, Straw and Bagasse-Straw 1:1 Mixtures Subjected To Hydrothermal Pretreatment and
787 Enzymatic Conversion, Microbial & Biochemical Technology, (2013).
788 [43] F.M. Yedro, H. Grénman, J.V. Rissanen, T. Salmi, J. García-Serna, M.J. Cocero, Chemical
789 composition and extraction kinetics of Holm oak (Quercus ilex) hemicelluloses using subcritical water,
790 The Journal of Supercritical Fluids.

791 [44] F.M. Yedro, J. García-Serna, D.A. Cantero, F. Sobrón, M.J. Cocero, Hydrothermal fractionation of
792 grape seeds in subcritical water to produce oil extract, sugars and lignin, Catalysis Today, 257, Part 2
793 (2015) 160-168.
794 [45] B.S. Santucci, P. Maziero, S.C. Rabelo, A.A.S. Curvelo, M.T.B. Pimenta, Autohydrolysis of
795 Hemicelluloses from Sugarcane Bagasse During Hydrothermal Pretreatment: a Kinetic Assessment,
796 BioEnergy Research, 8 (2015) 1778-1787.
797 [46] P. Moniz, H. Pereira, T. Quilhó, F. Carvalheiro, Characterisation and hydrothermal processing of
798 corn straw towards the selective fractionation of hemicelluloses, Industrial Crops and Products, 50 (2013)
799 145-153.
800 [47] P. Moniz, H. Pereira, L.C. Duarte, F. Carvalheiro, Hydrothermal production and gel filtration
801 purification of xylo-oligosaccharides from rice straw, Industrial Crops and Products, 62 (2014) 460-465.
802 [48] K. Elyounssi, F.-X. Collard, J.-a.N. Mateke, J. Blin, Improvement of charcoal yield by two-step
803 pyrolysis on eucalyptus wood: A thermogravimetric study, Fuel, 96 (2012) 161-167.
804 [49] M. Van de Velden, J. Baeyens, A. Brems, B. Janssens, R. Dewil, Fundamentals, kinetics and
805 endothermicity of the biomass pyrolysis reaction, Renewable Energy, 35 (2010) 232-242.
806 [50] F. Shafizadeh, Introduction to pyrolysis of biomass, Journal of Analytical and Applied Pyrolysis, 3
807 (1982) 283-305.
808 [51] W. Press, S. Teukolsky, W. Vetterling, B. Flannery, Numerical recipes 3rd edition: The art of
809 scientific computing, (2007).
810 [52] E. Ranzi, A. Cuoci, T. Faravelli, A. Frassoldati, G. Migliavacca, S. Pierucci, S. Sommariva, Chemical
811 Kinetics of Biomass Pyrolysis, Energy & Fuels, 22 (2008) 4292-4300.
812 [53] M. Jahirul, M. Rasul, A. Chowdhury, N. Ashwath, Biofuels Production through Biomass Pyrolysis
813 —A Technological Review, Energies, 5 (2012) 4952.

814

Appendix 1 - Calculation of yields

Reaction yields in the reactors was calculated as follows:

Volume of the system

At the beginning of the experiment (t_0) we measured the volume of the water contained in the system (V_0). At every sampling time, the volume of the system is calculated as:

$V(t)=V_0-\sum Vr_i(t-1)$

Where $Vr_i(t-1)$ is the volume of water contained in the unit removed after the previous sampling time.

So, at a sampling time of 5 min, the volume of the system is V_0; at a sampling time of 10 min, the volume is $V_0-Vr_{(5min)}$ and so on.

Mass of wood in the system

Mass of wood at the beginning of the experiment correspond to the sum of the mass of dry wood contained in every unit.

$Mw_0=\sum Mwr_i$

At every sampling time, the mass of wood in the system is calculated as:

$Mw(t)= \sum Mwr_i(t-1)-Msext(t-1)$

Where $\sum Mwr_i(t-1)$ indicates the mass of wood contained in the reactors at the sampling time, while Msext(t-1) correspond to the total mass of compounds extracted, measured at the previous sampling time.

Mass of hemicellulose extracted

Mass of hemicellulose extracted is calculaed as:

$Mhext(t)= Chext(t)*V(t)- Mhext(t-1)$

Where Chext(t) indicated the concentration of hemicellulose measured in the unit removed at time (t), while Mhext(t-1) indicates the mass of hemicellulose extracted at the previous sampling time.

Mass of hemicellulose in the wood

Concentration of hemicellulose contained in the wood at time 0 (Chw_0) was measured. Mass of hemicellulose contained in the wood at time 0 is calculated as:

$Mhw_0= Chw_0* Mw_0$

At sampling time, mass of hemicellulose extracted is subtracted from the mass of hemicellulose contained in the wood at the previous sampling time.

$Mhw(t)= Mhw(t-1)- Chext(t)$

Concentration of hemicellulose contained in the wood at sampling time is calculated as

$Chw(t) = Mhw(t) / Mw(t)$.

Yield of hemicellulose extracted

Yield of hemicellulose extracted is equal to:

$Yhext = (Mhw_0 - Mhw(t)) / Mhw_0$.

Appendix 2 - Figures

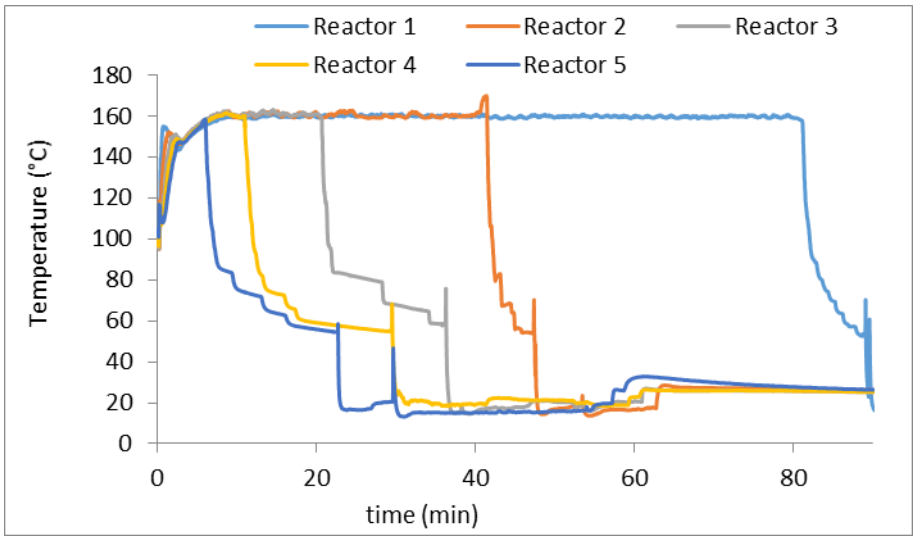
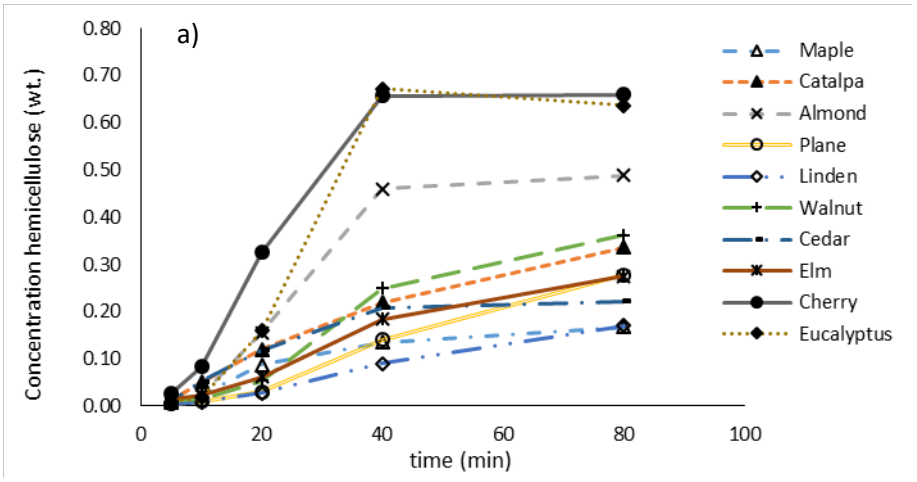


Figure S1. Standard temperature profile followed during the experiments.



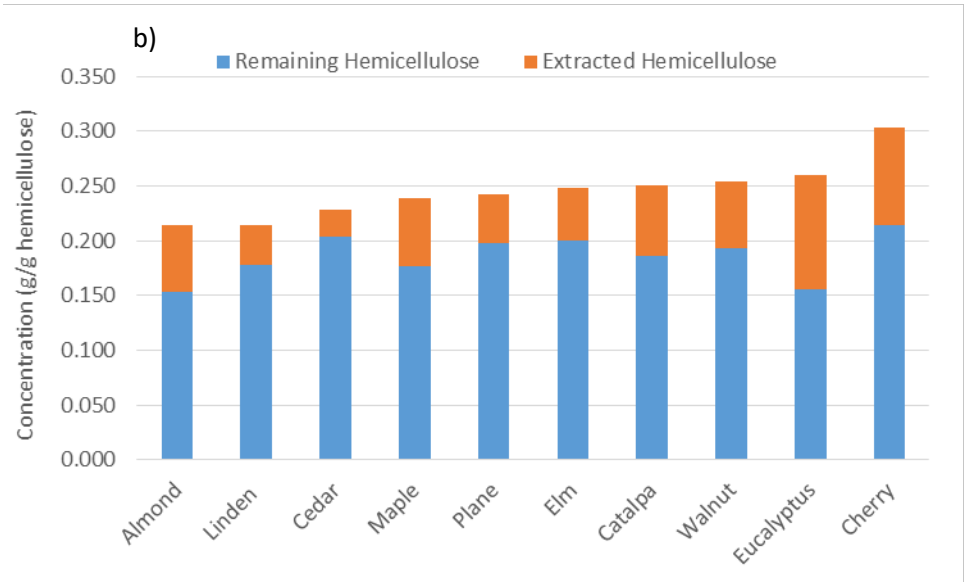


Figure S2. a) Concentration of hemicellulose extracted at different extraction times from the different raw materials. b) extracted hemicellulose vs total hemicellulose content for different raw materials.

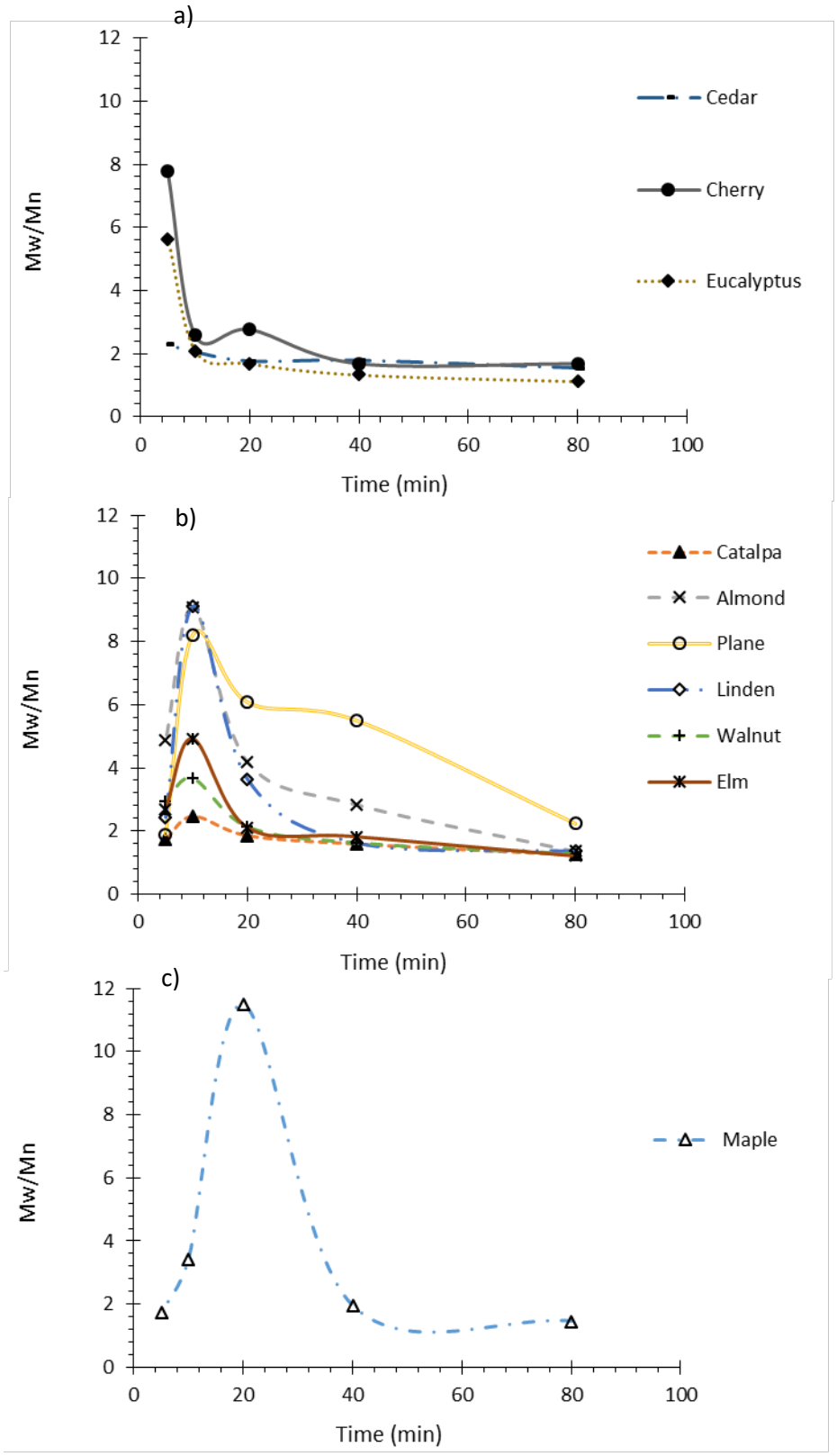


Figure S3. Polidispersity of oligomers extracted from different species of tree.

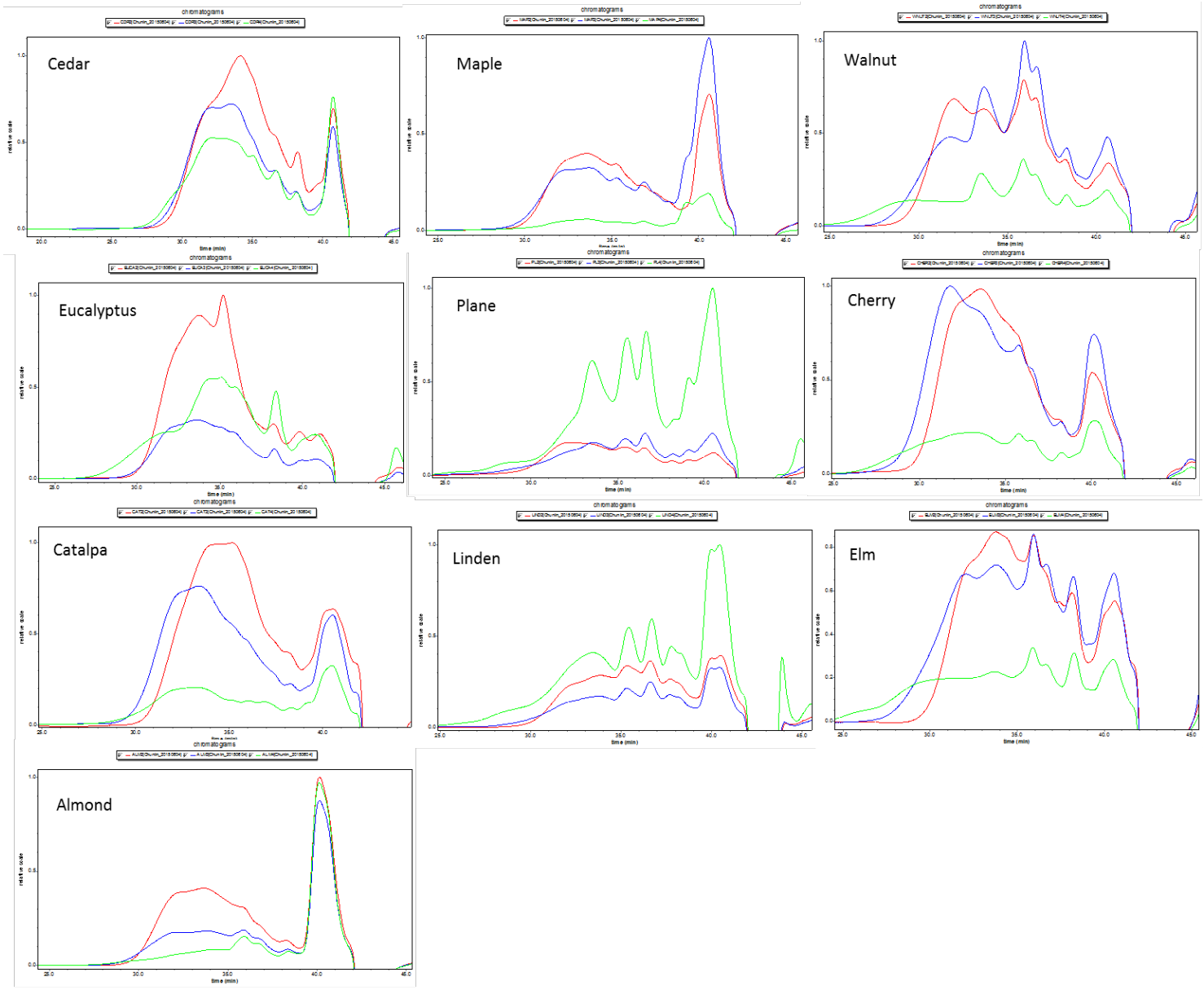


Figure S4. Molecular weight distributions of oligomers extracted from different species at 10 min (green), 20 min (blue) and 40 min (red).

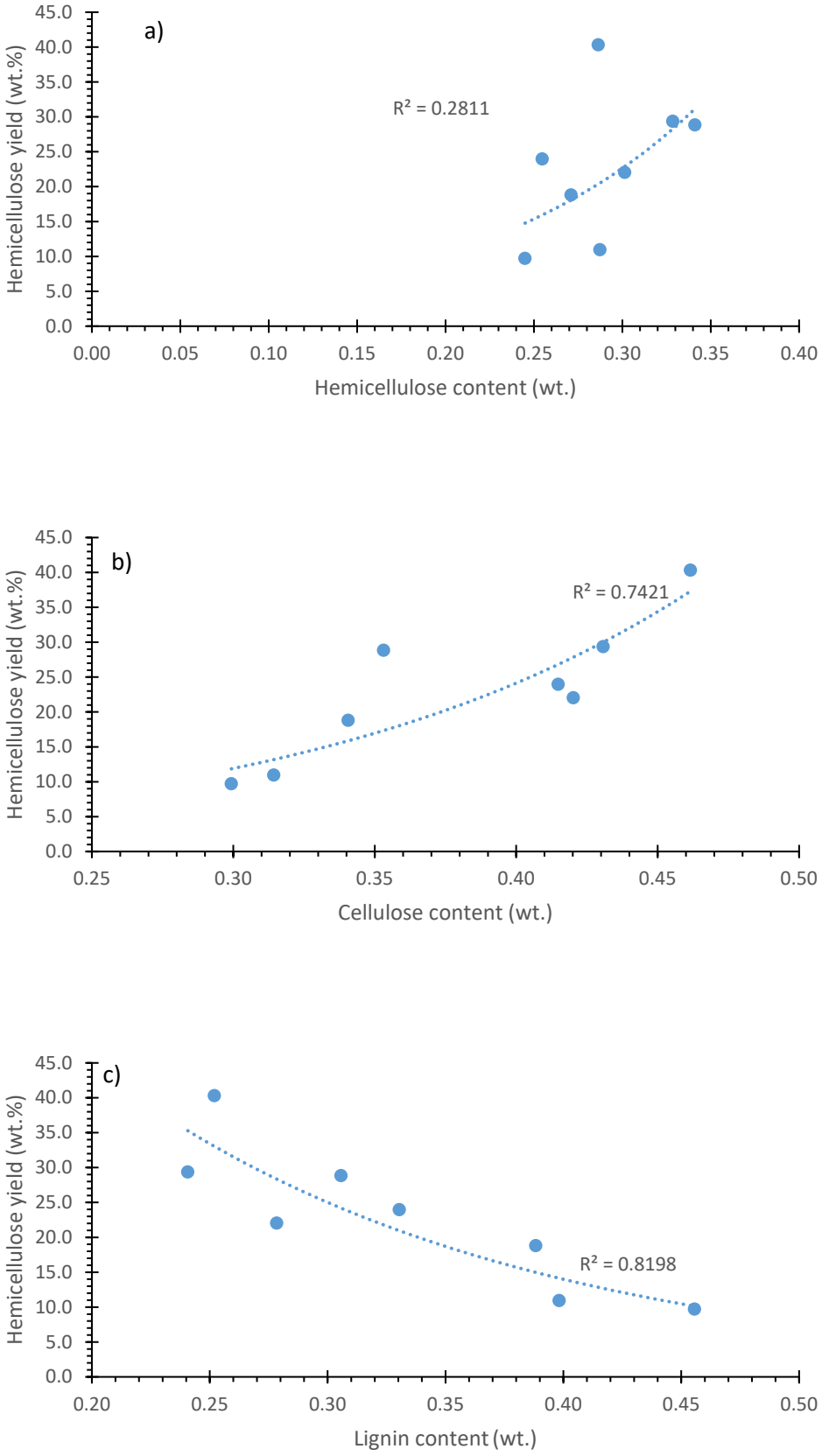


Figure S5. Hemicellulose yield evolution with cellulose (a) and lignin content (b) without *catalpa* and *elm.*

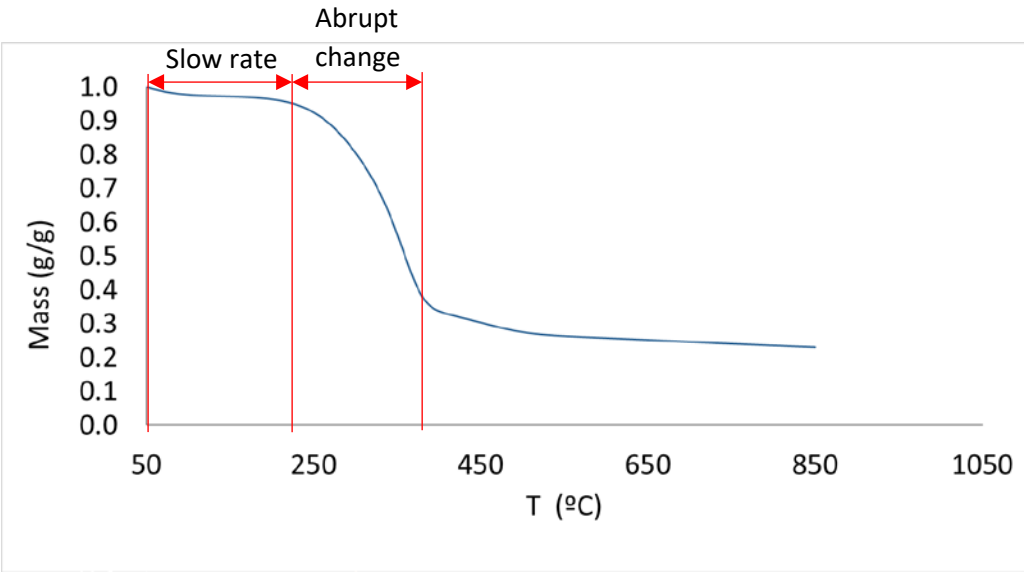


Figure S6. Almond experimental TGA

Appendix 3 – TGA modelling data and results

Table 1S. Absolute deviation for the TGA fittings and hemicellulose extraction yield estimation

	AAD[1] %	AAD[2] %
Walnut	3.70	1.71
Linden	32.27	1.09
Plane	12.31	1.44
Elm	0.01	0.84
Eucalyptus	16.47	0.94
Cherry	12.82	0.85
Cedar	28.20	1.45
Catalpa	0.01	1.56
Maple	0.41	1.19
Almond	20.56	0.82
	12.68	1.19

[1] Yield estimator fitting.

[2] TGA fitting.

Table 2S. Kinetic parameters obtained from the TGA fitting

	k_1*	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	Ea_1/R	Ea_2/R	Ea_3/R	Ea_4/R	Ea_5/R	Ea_6/R	Ea_7/R	Ea_8/R	Ea_9/R	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9
	min^{-1}	min^{-1}	min^{-1}	min^{-1}	min^{-1}	min^{-1}	min^{-1}	min^{-1}	min^{-1}	K	K	K	K	K	K	K	K	K	-	-	-	-	-	-	-	-	-
Walnut-RM**	22,430	9,074	25,377	19,274	0.030	17,769	3,821	0.299	0.398	6,745	12,055	8,119	7,242	347	10,059	8,456	399	109	0.6909	1.477	1.544	1.701	0.524	2.005	2.226	0.839	2.605
Almond	22,041	8,614	25,139	18,977	0.030	17,791	3,923	0.299	0.398	6,546	12,143	9,185	7,934	346	9,995	8,393	399	109	0.689	1.472	1.550	1.702	0.518	2.005	2.226	0.839	2.605
Maple	25,828	5,818	25,897	20,469	0.042	17,656	4,495	0.303	0.385	6,514	11,942	7,166	7,012	349	10,365	8,141	399	111	0.718	1.338	1.532	1.684	0.537	2.006	2.229	0.839	2.607
Catalpa	20,543	6,375	25,467	19,437	0.046	17,435	1,004	0.308	0.431	6,791	12,215	7,945	7,239	349	10,988	10,058	399	110	0.729	1.465	1.526	1.698	0.550	2.010	2.228	0.839	2.605
Cedar	22,364	8,660	25,607	19,330	0.033	17,754	3,851	0.299	0.398	6,485	12,362	7,656	7,247	347	10,100	8,437	399	109	0.687	1.464	1.542	1.702	0.532	2.005	2.226	0.839	2.605
Cherry	22,339	7,913	25,122	18,924	0.030	17,843	3,852	0.298	0.398	6,549	12,083	9,335	8,045	347	9,842	8,437	399	109	0.688	1.436	1.547	1.704	0.525	2.004	2.226	0.839	2.605
Eucalyptus	22,324	8,713	25,209	18,953	0.028	17,831	3,851	0.299	0.398	6,617	12,018	9,211	7,995	347	9,878	8,437	399	109	0.688	1.454	1.547	1.703	0.531	2.004	2.226	0.839	2.605
Elm	24,039	8,850	25,080	18,431	0.005	17,832	3,886	0.299	0.112	6,692	11,814	11,134	8,504	342	9,877	5,172	399	120	0.744	1.452	1.552	1.704	0.499	2.004	2.267	0.839	2.662
Plane	22,467	8,932	25,685	19,448	0.039	17,745	3,821	0.299	0.398	6,656	12,018	7,329	6,895	347	10,129	8,456	399	109	0.690	1.461	1.539	1.701	0.532	2.006	2.226	0.839	2.605
Linden	22,467	9,048	25,348	19,283	0.034	17,769	3,821	0.299	0.398	6,656	12,078	8,279	7,255	347	10,057	8,456	399	109	0.690	1.474	1.545	1.701	0.525	2.005	2.226	0.839	2.605

*Note: sub-index "i" refers to the reaction in which this kinetic parameter is involved. Therefore, 1: hemicellulose gasification, 2: cellulose gasification, 3: lignin gasification, 4: lignin char production, 5: lignin char gasification, 6: cellulose char production, 7: hemicellulose char production, 8: cellulose char gasification and 9: hemicellulose char gasification.

** An overall mass transfer parameters was used for both, water and organic liquid. Its value is 3,000 $g \cdot m \cdot min^{-1} \cdot mol^{-1}$ and 123 $g \cdot m \cdot min^{-1} \cdot mol^{-1}$, respectively.