

Liquid-Liquid equilibrium Prediction in Fast Pyrolysis Bio-oil Systems: Importance of improving Bio-oil chemical representation

Myriam Rojas Salas ^{a,*}, Frederico Gomes Fonseca ^a, Ana Cristina Corrêa de Araujo ^a, Manuel García Pérez ^b, Axel Funke ^a

a: Karlsruhe Institute of Technology, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany, b: Biological Systems Engineering, Washington State University, L.J. Smith Hall 205, P.O. Box 64120, Pullman, Washington USA

Corresponding author: axel.funke@kit.edu

Abstract

Liquid-liquid extraction, a cornerstone technique for bio-oil refining, suffers from a critical data gap: the lack of reliable partition coefficients (K_{OW}) for key fast pyrolysis bio-oil (FPBO) components. This absence hinders the application of established modeling strategies for optimizing these processes. To address this challenge, our study employs simulations to predict K_{OW} values for crucial FPBO components, employing commercial flowsheeting software and a bio-oil surrogate mixture. The study focuses notably on the unresolved structure of pyrolytic lignin ('pyrolignin'), a major FPBO constituent with great representation variety in the literature. To account for this variability, we explore 20 possible pyrolignin structures (phenolic dimers to tetramers), with the goal of determining the structures for which the experimental data are best described. Two scenarios were considered for accounting for the fraction of pyrolignin in the surrogate FPBO, and the best predictions to the water-oil equilibrium were found when employing more of this fraction. When discussing the K_{ow} of specific compounds, the lowest deviations were met when employing a water-extraction (WE) system for relevant FPBO compounds (levoglucosan, 2-cyclopenten-1-one, and guaiacol). Interestingly, the best KOW predictions for levoglucosan are obtained without including any modeled pyrolignin and modeling it with large molecules or mixtures. Additionally, several specific pyrolignin structures (dimer D2, trimer F1, tetramers I1 and I3) demonstrate promising agreement with experimental data. This work establishes a framework for incorporating bio-oil complexity, particularly the influence of diverse pyrolignin structures, into process modeling. This paves the way for more accurate simulations and ultimately the design of efficient and effective bio-oil refining processes.

Keywords: partition coefficient, levoglucosan, UNIFAC-Dortmund, pyrolytic lignin, model molecules, liquid-liquid equilibrium, pyrolysis bio-oil

Introduction

Fast pyrolysis is a promising pathway for the valorization of lignocellulosic materials. Using this technology, a bio-oil (FPBO) yield over 60 wt.% (of original feedstock) can be expected. This product has a very complex chemical nature with hundreds of oxygenated components covering a wide range of molecular weights and functionalities. A typical FPBO is a viscous liquid emulsion with a high water content (20-32 wt.%), high viscosity and low pH, limiting its usability and upgrading (1).

Figure 1 shows a liquid-liquid extraction scheme commonly used to fractionate FPBO, after extractives and any solids are removed. By adding water, it is possible to tip the balance in the emulsion to produce two main fractions. The water-soluble fraction (WS in Figure 1) is mostly comprised of oxygenated compounds (aldehydes, ketones, carboxylic acids, and sugars, 70-80 wt.% dry), while the water-insoluble fraction (WI in Figure 1) concentrates the majority of the phenolics present in the original mixture, as well as a significant fraction of the furanes (2). A second extraction stage frequently employs organic solvents, such as dichloromethane (DCM in Figure 1), chloroform, or diethylether, and can further elucidate the complexity of each fraction (3). Jeon et al. (3) discuss the effect of different solvents at this stage. While the WS soluble fraction has not received great attention, the extraction products of the WI fraction can be attributed to small phenolics, poly-sugars and hybrid molecules (WIS-DS), while lignin-derived oligomers (LO) populate the altogether insoluble phase (WIS-DIS) (1,3). (2) Understanding the stereochemical interactions between the molecules present in FPBO is crucial to predict phase equilibria. This information is desirable to design and improve FPBO fractionation processes.

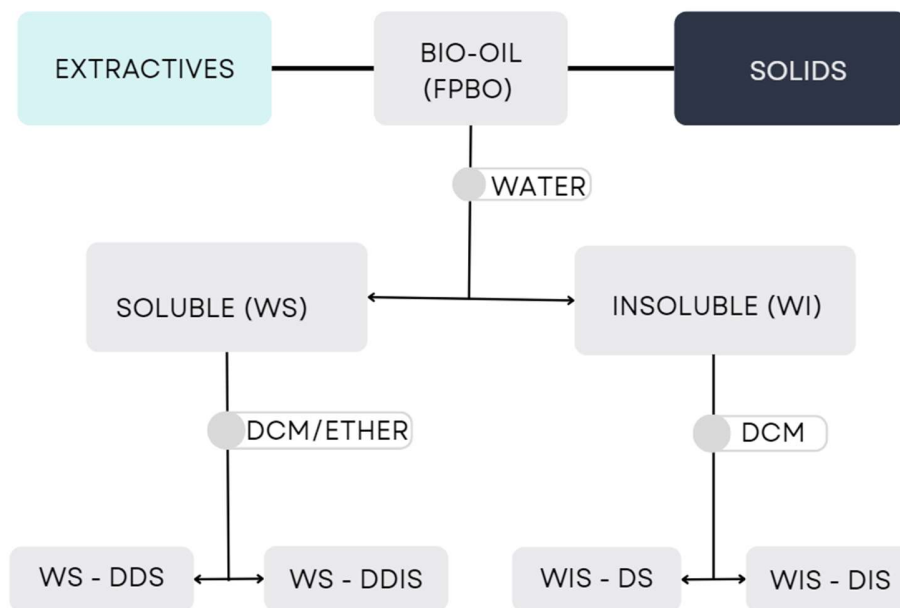


Figure 1: Solvent fractionation scheme. DCM: Dichloromethane (CH_2Cl_2); WS-DDS: Water soluble-DCM+DEE (Diethyl ether, $\text{C}_4\text{H}_{10}\text{O}$) soluble; WS-DDIS: Water soluble-DCM+DEE insoluble; WIS-DIS: Water-insoluble, DCM insoluble; WIS-DS: Water-insoluble, DCM soluble. Adapted from (4).

There are several approaches for FPBO fractionation with the aim to improve quality and/ or recover dedicated fractions. Fractional condensation setups separate compounds based on liquid-vapor equilibrium. The goal of these systems is to remove water and reactive light organics from the heavier phase (5,6). Optimizing temperatures in fractional condensation is commonly used to obtain heavy oils with targeted properties for further upgrading (7). The choice of quenching media used during the recovery of FPBO has a direct influence on phase separation after condensation, and may permit the selective recovery of target components (8). Liquid-liquid extraction can be used for FPBOs or its fractions to separate compounds based on functionalities, such as water washing to remove sugars (9), or light organics to remove phenolics (10). Interesting applications for extracted fractions range from upgrading to biofuels, fermentation/sugar chemistry, extraction of pyrolytic lignin and others. (6)1

A fundamental prerequisite to design fractionation processes is the prediction of phase equilibria, which in turn requires detailed knowledge of the FPBO composition. Apart from elemental, calorific and proximate analyses, GC-MS is the most commonly used chemometric analytics for characterization of FPBOs. A shortcoming of this method is the temperature cutoff (280-350 °C) which prevents identification and quantification of non-volatile molecules. This non-detectable phase is mainly comprised of by-products of lignin decomposition, dehydrated/modified oligo-sugars, and polyaromatics. Lignin depolymerization products are mostly concentrated in the water-insoluble fraction of FPBOs (WI in Figure

1). The WI fraction obtained with the solvent separation scheme shown in Figure 1 is typically known as *pyrolytic lignin* or 'pyrolignin' (term that will be used throughout the rest of the manuscript) (5)

Multiple representations of the nature of the pyrolignin phase have been proposed, but there is little consensus towards its modeling. Fonseca and Funke (6) discuss several proposals and their effectivity in modeling this fraction in vapor-liquid equilibrium (VLE) calculations. It has been shown that disregarding the presence of this fraction during VLE modeling can lead to skewed results and that the choice of surrogate to represent this fraction can significantly influence the outcomes (11–14). Several structures have been proposed for the modeling of the pyrolignin fraction, both by advanced analytical techniques and computational methods (6). The proposed structures constitute overwhelmingly of polyphenolics. Other oligomeric fractions that need to be included in the analysis of pyrolysis oils include dehydrated sugar fractions (humins) and hybrid oligomers (15,16). Although these polar oligomers are mostly concentrated in the water-soluble (WS) fractions, they are also found in the Water-Insoluble (WIS) fraction, which is rich in lignin depolymerization products. The impact of the quality of the explicit representation of FPBO chemistry on phase equilibrium calculations warrants further investigation.

While there has been some initial work on predicting VLE of FPBO, studies on liquid-liquid equilibria (LLE) are scarce. Binary/tertiary mixtures have been studied involving typical FPBO components rather than the entire oil (17–19). Oasmaa et al. (1) mention the influence of the pyrolignin phase during LLE, but little has been done to account for its contribution in LLE modeling.

The objective of this study is to investigate the influence of surrogate molecules to represent unknown pyrolytic oligomers for the modeling of LLE to design FPBO extraction processes. Experimental results published by Han et al. (20) will be used as reference to validate the predictive capability of LLE calculations. Data for the separation of all components in the surrogate FPBO mixture are obtained during the data processing, and interesting tendencies are presented for relevant molecules in this work.

Methodology

The study is based on experimental results for FPBO extraction reported elsewhere (20). To reduce the uncertainties by including a large number of compounds with unknown thermophysical properties, the number of different species in the surrogate mixture was lowered according to the heuristics proposed by Fonseca and Funke (6). The original mixture and the two case surrogate mixtures employed are presented in Table 1. The characterization of the original mixture involved quantification of the pyrolyginin in addition to other unknowns from GC analysis which most likely include primarily substituted poly sugars. The compounds used in the surrogate mixture are numbered in the first column; all compounds without number are added to the previous surrogate to determine its fraction (e.g., butyrolactone and 2(5H)-furanone are added to furfural). **Case 1** refers to a surrogate mixture with a direct normalization of all surrogates to meet 100%. **Case 2** refers to a situation in which the unknown fraction of the original mixture (~21 %) is added to the pyrolyginin prior to normalization. 20 possible identities for pyrolyginin were considered to study the influence on LLE prediction, which are presented in Table 2.

Table 1. FPBO composition provided by Han et al. (20), and definition of relevant cases (null, 1, 2)

Number	Compound	Base FPBO (wt.%)	Null Case (\emptyset , wt.%)	Case 1 (wt.%)	Case 2 (wt.%)
1	Water (by KF titration)	32.4	56.92	41.07	35.64
Carboxylic Acids, Ketones, and Furans					
2	Glycolaldehyde	0.68	1.46	1.05	0.91
3	Acetic acid	1.4	3.00	2.16	1.88
4	Acetol	1.21	2.59	1.87	1.62
	Butanedial	0.07			
5	2-cyclopenten-1-one	0.19	0.41	0.29	0.25
6	Furfural	0.31	1.20	0.87	0.75
	Butyrolactone	0.06			
	2(5H)-furanone	0.19			
7	Corylon	0.6	1.29	0.93	0.80
8	Furfural. 5-hydroxymethyl-	0.33	0.71	0.51	0.44
	Other carboxylic acids/ketones/furans	1.09			
Phenolics					
9	Phenol	0.13	0.34	0.24	0.21
10	Phenol. 2-methoxy-	0.76	1.97	1.42	1.23
	Phenol. 2-methyl-	0.1			
	Phenol. 4-methyl-	0.24			
	Phenol. 3-methyl-	0.05			
11	Phenol. 2-methoxy-4-methyl-	0.86	2.22	1.60	1.39
	Phenol. 3,4-dimethyl-	0.12			

	Phenol. 4-ethyl-	0.05			
12	Phenol. 4-ethyl-2-methoxy-	0.46	1.19	0.86	0.75
	Phenol. 2-methoxy-4-vinyl-	0.07			
13	Phenol. 2-methoxy-5-(2-propenyl)-	0.34	2.25	1.62	1.41
	1,2-benzenediol	0.12			
	Phenol. 2,6-dimethoxy-	0.05			
	Phenol. 2-methoxy-4-(1-propenyl)-. (E)-	0.27			
	Phenol. 2-methoxy-4-(1-propenyl)-. (Z)-	0.26			
	Phenol. 4-methoxy-3-(methoxymethyl)-	0.05			
14	Vanillin	0.41	1.06	0.77	0.66
	Phenol. 2-methoxy-4-propyl-	0.24			
15	4-acetyl-guaiacol	0.3	1.22	0.88	0.76
	Guaiacyl-acetone	0.17			
	Other phenols	0.78			
Sugars					
16	Levogluconan	9.57	22.14	15.97	13.86
	Other anhydrosugars	3.03			
Others					
17	Pyrolignin	22		27.89	37.41
Total		78.96	100	100	100

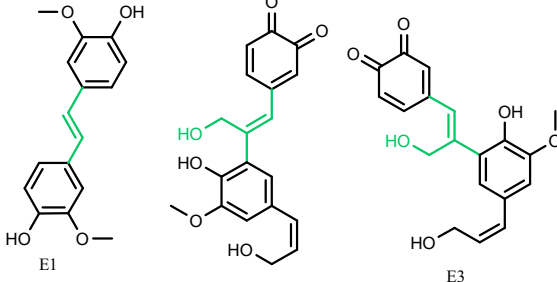
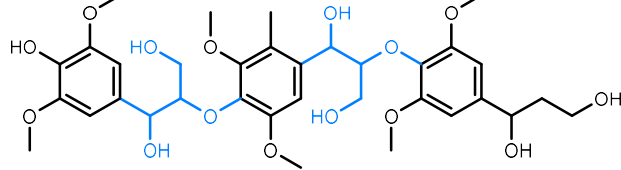
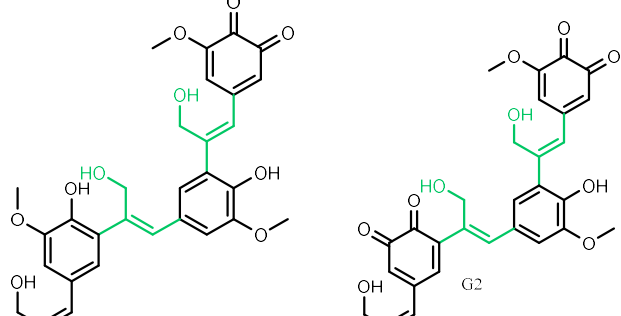
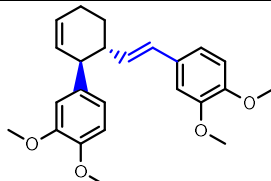
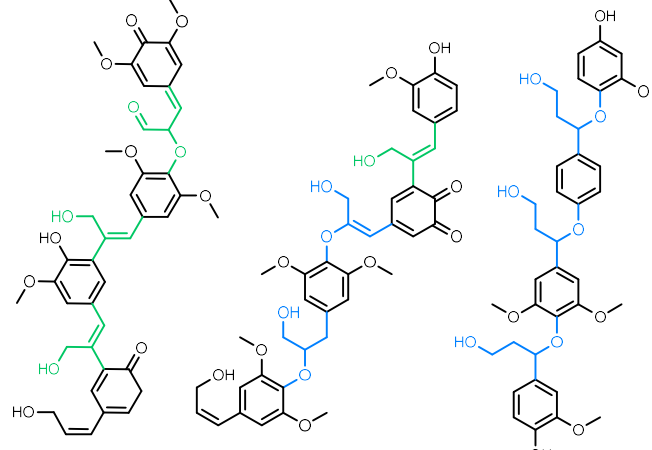
The system was modeled using Aspen Plus™ V14, where the bio-oil is modeled as the mixture presented in Table 1 (case 1 or case 2). Pyrolignin refers to one of the proposals summarized in Table 2, or to a case where no pyrolignin is present. All streams (bio-oil, water, solvent) are modeled as liquid streams at operation temperature (293 K), according to the methodology employed in the experiments that serve as reference datasets for this study (20). The liquid streams are mixed in a mixer unit block and the product is channeled to a decanter unit block. The decanter operates at 293 K and water is defined as the key component of the 2nd liquid, branding it as an aqueous phase. Phase split is modeled by equalizing component fugacities in the two liquids, and liquid-liquid coefficients are estimated from the property method. The property method employed is UNIFAC-Dortmund (DMD), chosen due to its ability to model molecules which little experimental data is available, such as many of the molecules present in the bio-oil surrogate used in this work. The software is able to automatically generate the group distribution based on the inputted molecular structure for different molecules, in order to estimate thermophysical properties such as normal boiling point, critical properties, etc... The division of LLE UNIFAC groups used in this work is detailed in Table S12.

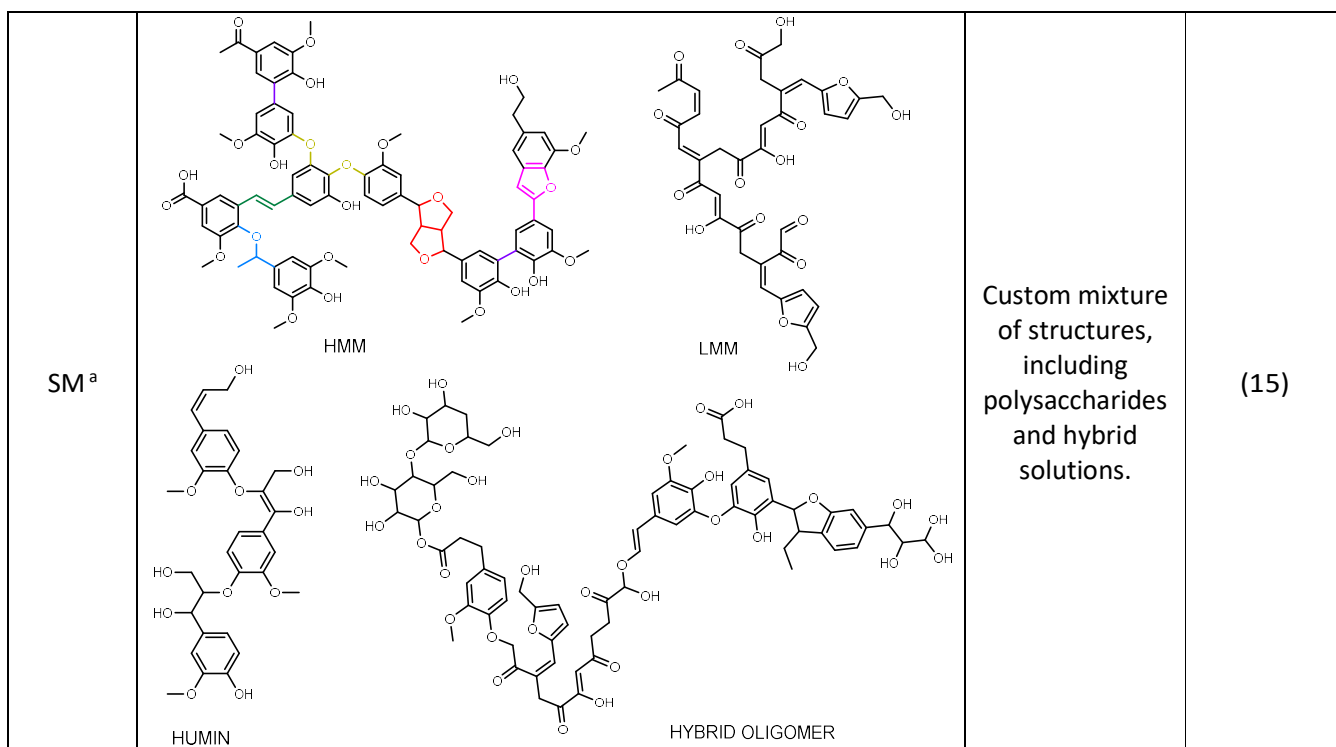
The Aspen Plus™ model was connected to Microsoft Excel™ via the Aspen Simulation Workbook™ to facilitate data collection. Two sets of results are considered based on varying proportions of bio-oil,

water and butanol, according to the experimental setup used for validation (20). Within the BEsystem, the solvent (n-butanol) input was varied six times (BE₁: 6, BE₂:10, BE₃:15, BE₄:22.5, BE₅:30, and BE₆:60 g), while within the WEsystem, the water input was varied also six times (WE₁: 7.5, WE₂:15, WE₃:22.5, WE₄:30, WE₅:45, and WE₆:60 g).

Table 2. Chemical structures of the pyrolignin model molecules.

Group	Structure	Relevant link types	Reference
Dimers			
A	<p style="text-align: center;">A1 A2</p>	Phenylcoumaran β-5	(21)
B	<p style="text-align: center;">B1 B2 B3 B4 B5</p>	Biphenyl 5-5 linkage	(21-23)
C	<p style="text-align: center;">C1</p>	Resinol β-β	(21)
D	<p style="text-align: center;">D1 D2</p>	Alkyl-aryl ether β-O-4	(24),(25)

E	 <p>E1 E2 E3</p>	Stilbene β -5	(26)
Trimers			
F	 <p>F1</p>	Alkyl-aryl ether β -O-4	(27,28)
G	 <p>G1 G2</p>	Stilbene β -5	(26)
H	 <p>H1</p>	Bridging double bond	(29)
Tetramers			
I	 <p>I1 I2 I3</p>	Mixed links: stilbene β -5 (sometimes enolized), alkyl-aryl ether β -O-4	(27,28,30)



a: HMM (28.08%), LMM (25.96%), HUMIN (17.87%), and Hybrid oligomer (28.08%). Structures and proportions adapted from (15).

Data processing, trend analysis, error calculations, and figure creation were performed using Matlab 2024a[®]. Relative deviations to the experimental data (20) were estimated for every species present, as well as a compound entity named the *Bio-oil*, comprised of all species except water and solvent. For each case (1, 2) and system (BE, WE), deviations between calculations and experimental data are reported as mean relative error (MRE).

Han et al. (20) defined parameters for evaluating their experimental results. The mass ratio between the organic and aqueous phase (Ratio O/A) was estimated according to Eq. 1. The partitional coefficient of compound *i* that distributes in both phases ($K_{OW,i}$) is estimated according to Eq. 2.

$$Ratio\ O/A = \frac{m_{org}\ [g]}{m_{aq}\ [g]} \quad 1$$

$$K_{OW,i} = \frac{m_{i,org}\ [g]/m_{org}\ [g]}{m_{i,aq}\ [g]/m_{aq}\ [g]} \quad 2$$

The results presented in figures and data tables in the manuscript represent a choice of different trends and predictive capability. The complete curves and data is included in the supplementary information (SI). The main components evaluated were water, *Bio-oil* (considered all the compounds in Table 1, except water), and solvent (n-butanol).

Results and Discussion

Prediction of phase separation of the system water, *n*-butanol, and bio-oil

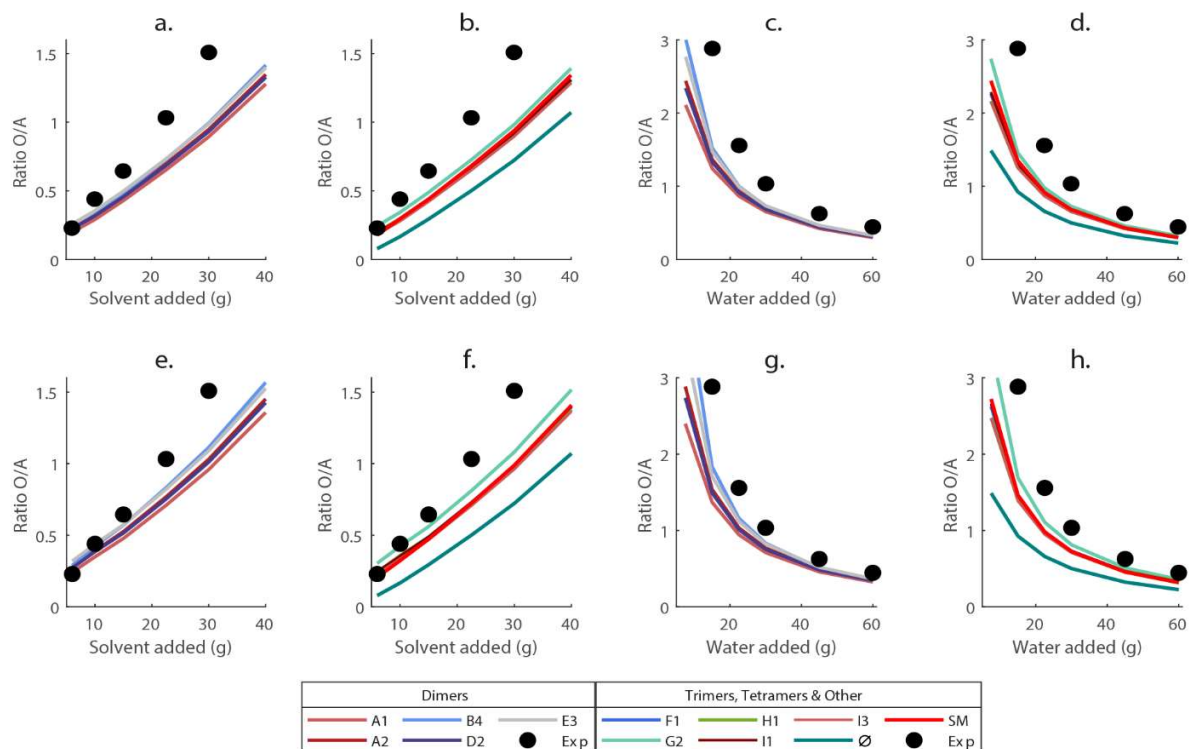


Figure 2. Variation of solvent (BE) and water (BE) and its impact on the MRE of the Ratio O/A for Case 1: BE (a & b), and WE (c & d), and Case 2: WE (e & f), and WE (g & h).

In Figure 2, one can see the evolution of the Ratio O/A with the increasing input of butanol (a, b, e, f) or water (c, d, g, h), for the best performing model molecules. Most of the model molecules present a very similar behavior which can be attributed to their hydrophobic center and possible steric issues, or shortcomings of using UNIFAC-DMD for the prediction of LLE. A clear outlier is the null case (\emptyset) which reinforces the importance of employing a representation for the pyrolygnin phase when modeling LLE phenomena, reinforcing the findings of previous researchers (6,18). The worst MREs are found for the null case (\emptyset), followed by the SM case (details for all curves and data are presented in Tables 1S, 2S and Figures 1S in the Supplementary Information). As SM is a mixture comprised of 4 structures of high molecular weight (15), the bad performance may reflect shortcomings when LLE modeling using UNIFAC-DMD.

The WE system (WE₂-WE₆ as in Figure 2 c, d, g, h; WE₁ consistently lays outside of the value range predicted by the model using any pyrolygnin representative) presents better predictions for the Ratio O/A (Figure 2) for either Case than the BE system (BE₁-BE₅ as in Figure 2 a, b, e, f; BE₆ consistently lays outside of the value range predicted by the model using any pyrolygnin representative). While water is a common

extraction medium for FPBO fractioning (1,3), the nature of the pyrolignin representatives (Table 2) would indicate, at a first glance, a larger tendency to migration to the organic phase.

Despite the worse predictions (Figure 2), the BE system is well predicted up to BE₃, and higher fractions of n-butanol cause the model to present increasingly higher degrees of underestimation of the Ratio O/A, indicating the model predicts a higher migration to the aqueous phase than was observed experimentally. This is interesting due to the relatively high fractions of phenolics and pyrolignin present in the surrogate mixture (Table 1).

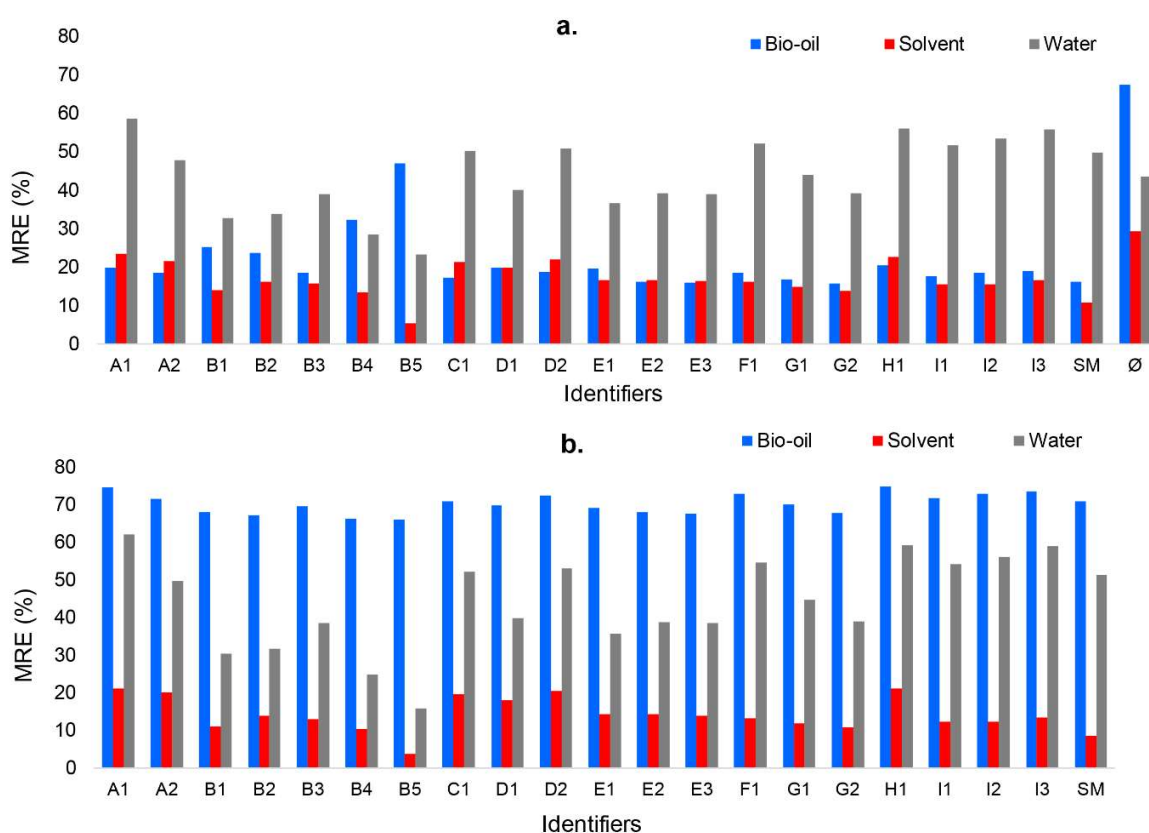


Figure 3. MRE values for the prediction of the K_{ow} for the main fractions within the system WE: Bio-oil, solvent, and water for a) Case 1 and b) Case 2.

Figure 3 presents the deviations in prediction of the partition coefficient (K_{ow}) for the main components of the system WE (*Bio-oil*, water, solvent). Case 1, which excluded water-soluble molecules from the pyrolignin fraction, demonstrated superior performance in predicting the distribution of *Bio-oil* compared to Case 2 (max MRE: 47% for Case 1 vs min MRE: >66% for Case 2). A high MRE was found for the null case (\emptyset , MRE: 67%). These points highlight the importance of accurately accounting for pyrolignin during bio-oil modeling. The system BE was not considered due to lack of experimental data (20). For Case

1, the best results were obtained with G2 (MRE: 15 %), followed by E2, E3, C1, I1 and SM (MRE: <18 %), while the worse results were found for the dimers B5, B4, and B1 (MRE: 47%, 32%, and 25%, respectively)

It is important to keep in mind that the fraction of pyrolignin in the bio-oil was determined by cold water precipitation (20). This means the unknowns present in the base reports of FPBO composition correspond to a water-soluble fraction (poly sugars, polar aromatics, oxygenates of various natures) that is either not present in the database or not volatile at GC/MS operation ranges (6,31). This means that Case 2 (Figure 3 b) may incur in implicit error by including water-soluble molecules in the pyrolignin fraction, which may help explain the more pronounced errors observed when using Case 2

While Case 1 outperformed Case 2 for bio-oil prediction, both methods yielded comparable results for the water fraction (Figure 3). Additionally, the model consistently struggled to accurately predict the K_{ow} of the water fraction (Figure S4), regardless of the chosen case (MRE: 23-58% for Case 1, 15-65% for Case 2). The authors suggest that limitations in the chosen model molecules for sugar oligomers might contribute to this discrepancy. Overall, the best results were found for B5 (MRE: 23%) and the worst for A1 (MRE: 58%), while the null case (\emptyset) results in an MRE of 43% (see also Figure S2, in the SI).

When discussing the K_{ow} of the solvent (n-butanol), unlike for the bio-oil, the B5 structure showed the best approximation to the experimental data (MRE: 5.4%). The null case (\emptyset) also performed the worst in this case (MRE: 29.1%). Interestingly, for all model molecules within Case 2 (Figure 3 b), better predictions were found for the solvent fraction when compared to Case 1 (Figure 3 a). The resulting trend shows that the K_{ow} of n-butanol initially increases and tends to become constant as more water or solvent is added (Figure S3).

The K_{ow} of the *Bio-oil* fraction is increasing when adding water and decreasing when solvent is added (Figure 2S in the SI). This can be explained by the polar nature of the solvent (n-butanol), which is expected to partially distribute into the aqueous phase.

Regarding the modeled pyrolignin fraction (Figures S7-S9), the K_{ow} varied according to each evaluated structure. All structures showed an affinity for the organic phase as expected, with K_{ow} values ranging from 2 to approximately 966,000,000. This wide range indicates the diversity of the structures, with some being entirely transferred to the organic phase ($K_{ow} \gg 10,000$) such as A1, H1, G1, F1, I1 and I3. In the range of 100 to 1,000, showing slightly less refinement, are A2, E2, E3, and D2. The range of 10 to 100 includes B1, B2, E1, and D1. The molecules with the lowest affinity for the organic phase, with $K_{ow} < 10$, are B3, B4, and B5. Although these latter molecules prefer the organic phase, their polarity allows for some low concentration transfer to the aqueous phase.

Evaluation of relevant individual species

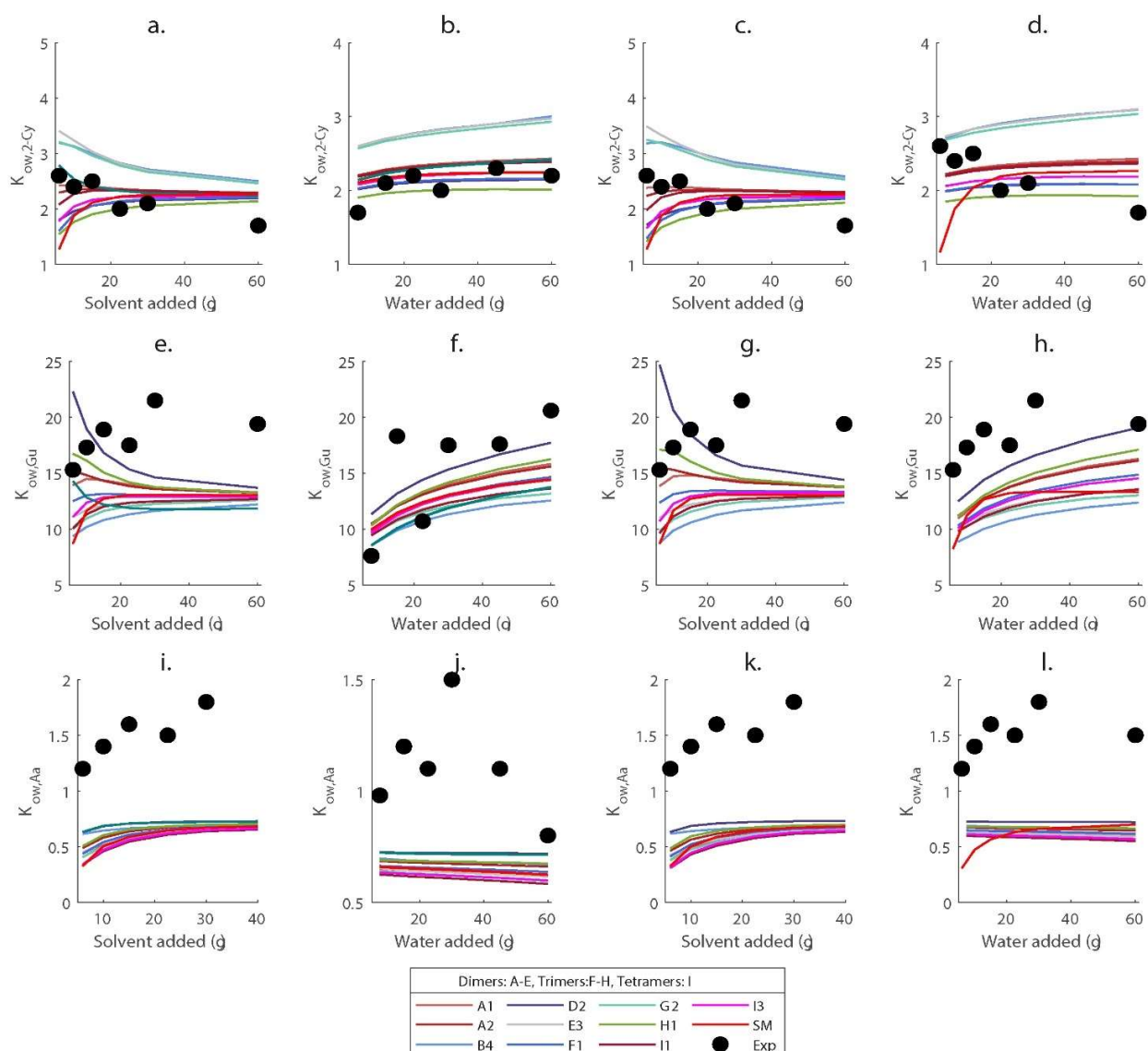


Figure 4. Variation of the predicted K_{ow} for 2-cyclopenten-1-one: a) BE, case 1, b) WE, case 1, c) BE, case 2, d) WE, case 2; guaiacol: e) BE, case1, f) WE, case 1, g) BE, case 2, h) WE, case 2; and acetic acid: i) BE, case1, j) WE, case 1, k) BE, case 2, l) WE, case 2.

Figure 4 presents a comparison between the predicted K_{ow} and the experimental value for three relevant species in FPBO: 2-cyclopenten-1-one, guaiacol and acetic acid. Experimentally, the former presented K_{ow} values between 1.5 and 2.6, lowering for increasing amounts of solvent and water, revealing a moderate affinity to the organic phase (Figure 4 a-d). Case 1 using the WE system (Figure 4 b) presents the best predictions overall, with clear outliers (g) when using E3 and G2. For the remaining cases, the predicted trends plateau and do not accompany the trend of the experimental data. Overall, the best

results are found for D2, F1, H1, I2 and I3 (MRE: 6.6% – 7.4%) while the worse are found for B4, E2 and E3 (MRE: ~ 35%).

The variation with the experimental K_{OW} values for guaiacol do not follow clear trends (Figure 4 e-h), indicating possible issues with the experimental data, but are overall increasingly positive (K_{OW} 7.5-22.5) for increasing additions of solvent and water, for both cases. In terms of prediction, the K_{OW} of this molecule is consistently underestimated for all cases except Case 1 WE_{1,2}. Overall, the best results are found for D2, H1, A2 and A1 (MRE: 23.9% - 25.8%) while the worse are found for G2 and B5 (MRE: 28% and 33 %, respectively).

Regarding acetic acid, the experimental K_{OW} values reflect a better distribution between the organic and aqueous phases (K_{OW} 0.75-1.75), with an even less clear trend than the one seen for guaiacol, due to the high affinity to water at high water additions (Figure 4 j, WE₆). The predictions overwhelmingly underestimate the K_{OW} of acetic acid, estimating an unrealistic concentration in the aqueous phase, which may be attributed to shortcomings of the model (simple decanter) or the UNIFAC-DMD model. Overall, the best results are found for D2 and B1 (MRE: 34.2 % and 34.9%, respectively) while the worse are found for C1 and I1 (MRE: ~ 43.5%).

Trends for all cases, systems and individual components can be consulted in the SI (Tables S5-S11 and Figure S5).

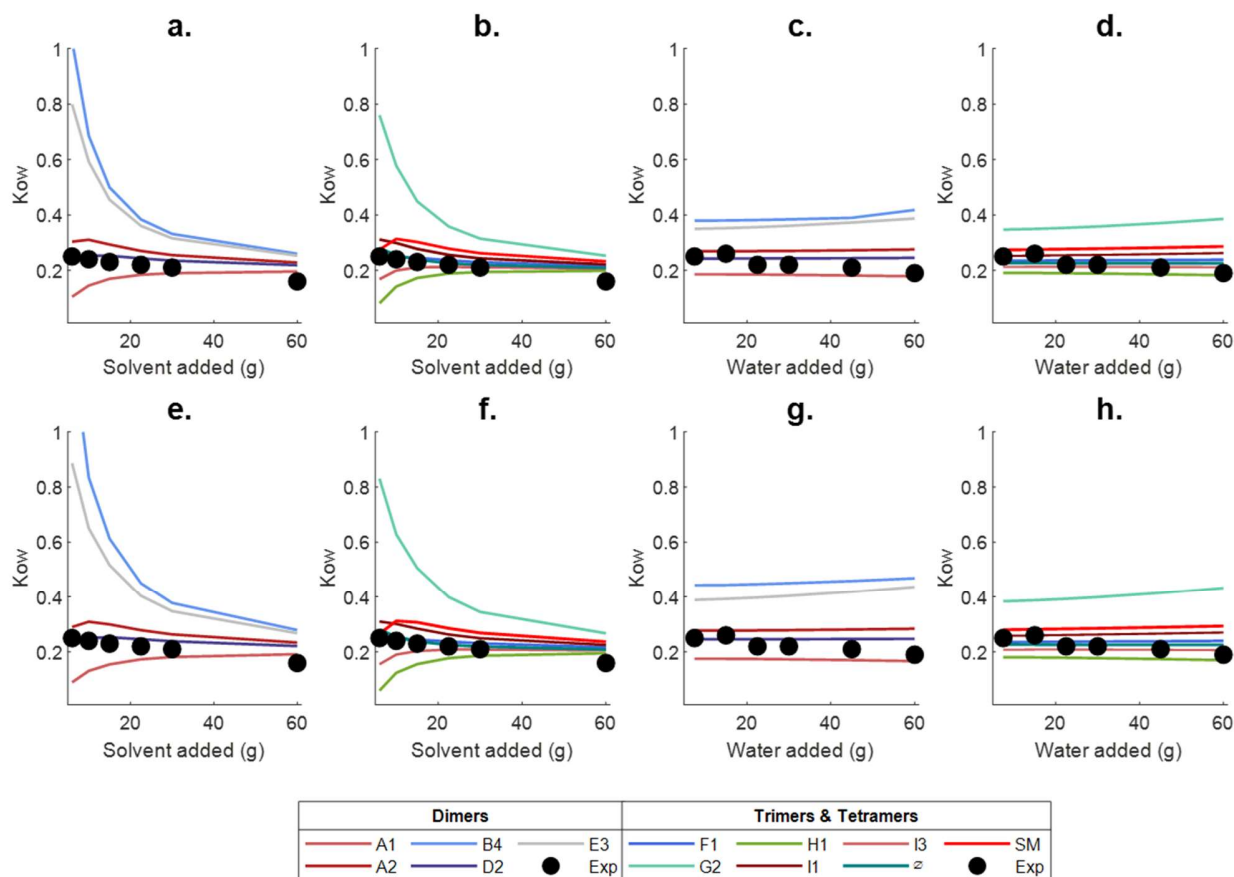


Figure 5. Evolution of the experimental and predicted partition coefficients for levoglucosan with added solvent (BE) and water (WE). Case 1: BE(a & b), and WE(c & d), and Case 2: BE(e & f), and WE(g & h)

Figure 5 mirrors the analysis presented in Figure 4, but focusing on levoglucosan. Experimental results indicate that this sugar distributes preferentially to the aqueous phase ($K_{ow} \sim 0.2$). Interestingly, the K_{ow} tends to further favor the aqueous phase for increasingly higher additions of water and n-butanol. This trend seems to be well predicted in the model, except for some clear outliers, namely B4 and G2 (MRE: 75.2% and 62.6% respectively), while the best performers were I3 and I2 (MRE: 8.7% and 10.5%, respectively). The MREs were slightly larger for Case 2.

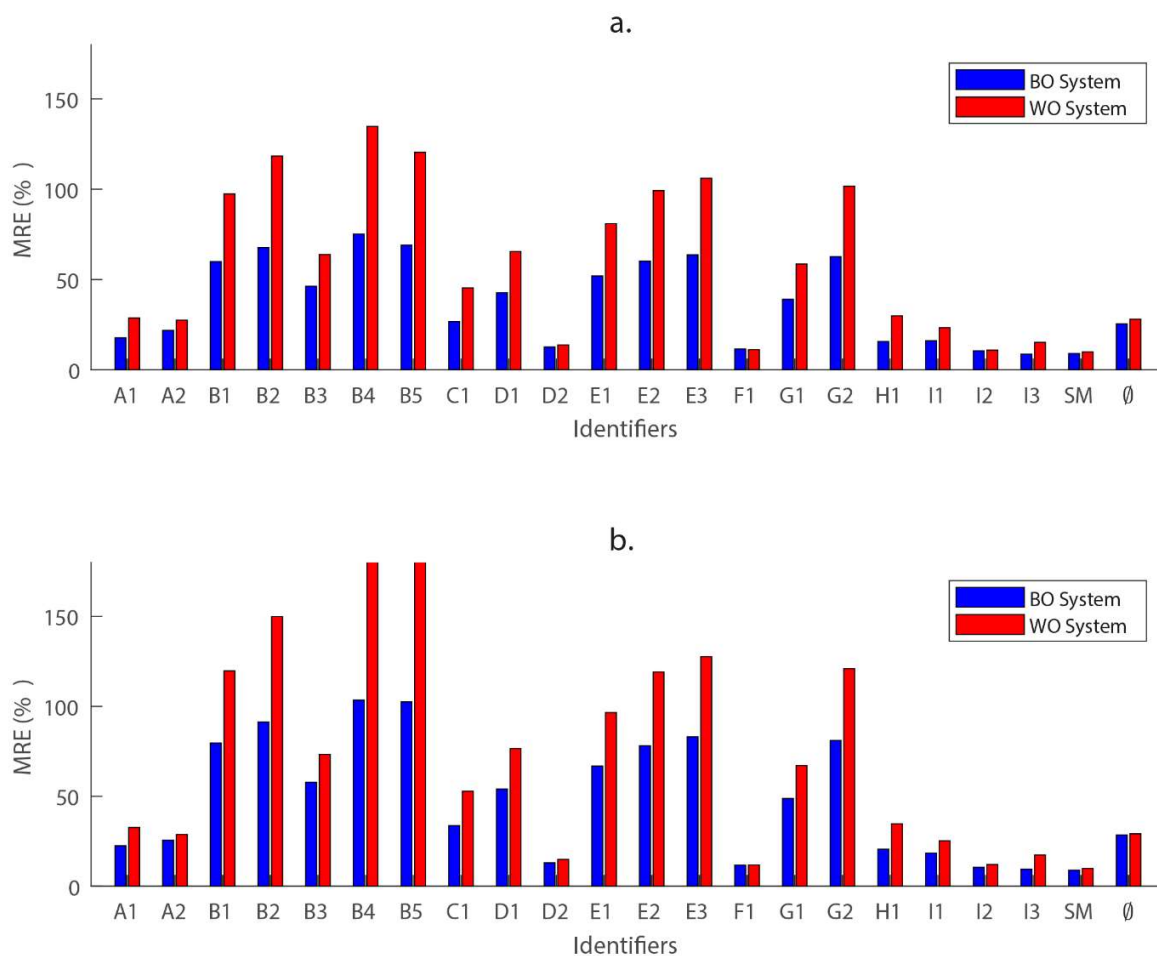


Figure 6. Bar diagrams of MRE values for K_{ow} prediction for levoglucosan for a) Case 1 and b) Case 2.

In Fehler! Verweisquelle konnte nicht gefunden werden., one can see the trends presents in Figure 5 in bar graph form, allowing for a more *in depth* analysis of the performance of each pyrolignin representative. Greatly contrasting with the trends observed in Figure 2, Figure 3, and Figure 4, the best results can be found in WE system for the null case (\emptyset) and the larger representatives SM, I2 and I3 (MRE <10%), as well as some dimers (D2) and trimers (F1) (MRE: 11.2% - 12.6%). The MRE values were slightly higher for the BE system. This may be explained by their complexity and number of different interaction groups presents, which better model potential interactions between levoglucosan and pyrolignin in solution.

It is important to highlight that most of the surrogates evaluated in this study successfully captured the trend of both systems (WE and BE), demonstrating a good approximation to the real system and paving the way for a better understanding of the chemical composition of FPBO and the potential of extraction. Similar results, capturing the trends in levoglucosan phase distribution, have been found by Parku et al.

(32) using conventional water extraction of this compound from already condensed FPBOs, who also reported considerable errors when comparing to modeling results also using UNIFAC-DMD (RMSD \approx 22%). However, it is crucial to consider that the present results have been compared only with a single set of experimental data. This precision could be significantly improved with new experiments and measurements, as well as by considering other sugar oligomer structures to better represent the polar fraction of the system. Candidate sugar oligomers for evaluation include those proposed by Denson et al. (30), which are from monomers until tetramers.

The pyrolygnin representatives from group A (dimers, Table 2) present good results (MRE < 21%), with slight variations between the two group members. The benzofuran ring in A1 contributes to hydrophobic interactions due to its higher affinity, leading to phase separation when mixed with polar solvents such as water, and preferring the organic phase ($K_{OW,A1} > 1000$). Partial saturation of the dihydrobenzofuran ring in A2 increases the polarity ($126 \geq K_{OW,A2} \leq 328$), which may reduce the tendency for phase separation in polar solvents compared to A1. In comparison, group B representatives (methoxy-substituted dimers, Table 2) performed markedly worse (MRE > 45%), but captured the trends observed in Figure 5.

Group D (β -O-4 dimers, Table 2) is comprised of two similar molecules with very different behaviours. The MREs (\geq 50%) observed for D1 may be attributed to due to conformational configurations of these molecules that indubitably affect solubility in either phases, especially in an acid environment such as FPBO. This factor is most likely not well reflected when using the UNIFAC-DMD model. On the other hand, D2 is one of the best performing representatives for the prediction of the partition of levoglucosan (MRE <14%), which may be due to the presence of four methoxy groups and resonance effects between the aryl groups, and possible enol conformations in the hydroxy of the β -O-4 bond. Similar resonance effects are also heavily featured in the structure of F1 (β -O-4 trimer, Table 2), that performs even better than D1.

The difference in performance between G1 and G2 (MRE: <59% and >60%, respectively) within group G (stilbene β -5 trimers, Table 2) further clarify shortcomings in the methodology, namely the inability of the software and the thermodynamic model to account for resonance effects, especially when the difference between the molecules in the presence of a methoxy-hydroxy substituted ring compared to a quinone ring. H1 (saturated intermediate ring, double-ring bond, Table 2) performs comparably with group A, I1 and SM, and may be explained by the saturation of the intermediate ring and the presence of methoxy substituents in the aromatic rings.

Group I (Table 2) features tetramers with different termina and interbonding, but most rings feature methoxy substituents. The best performance between these is found for I2, which may be attributed to only employing stilbene β -5 bonds which are desaturated, compared to other bonds employed by I1 and I3.

Upon reviewing probable trends or correlations between certain characteristics of the evaluated molecules, a considerable correlation was found between the presence of methoxy groups ($-\text{OCH}_3$) in the molecular structure and the accuracy of $K_{\text{OW.LVG}}$ prediction with R^2 values ranging between 0.61 - 0.68. The size of the molecules showed only a slight correlation (R^2 : 0.35 - 0.37), while the number and variety of functional groups present did not exhibit significant trends (Figure S6). Therefore, it can be hypothesized that structures with a greater diversity and quantity of functional groups, especially methoxy groups, would be appropriate for modeling the liquid-liquid phase equilibrium when seeking to extract levoglucosan. This may also be favorable for other sugars of interest.

It is important to highlight that most of the surrogates evaluated in this study successfully captured the trend of both systems (WE and BE), demonstrating a good approximation to the real system and paving the way for a better understanding of the chemical composition of FPBO and the potential of extraction. It is important to acknowledge the inherent limitations of modeling techniques and potential for errors ($\sim 30\%$), which were significantly improved with several structures, resulting in errors of less than 10%. Similar results, capturing the trends in levoglucosan phase distribution, have been found by Parku et al. (32) using conventional water extraction from already condensed FPBOs, who also reported considerable errors when comparing to modeling results also using UNIFAC-DMD ($\text{RMSE} \approx 22\%$). However, it is crucial to consider that both the results presented in this work and the work by Parku et al. (32) have been compared only with a single set of experimental data. This precision could be significantly improved with new experiments and measurements, as well as by considering other sugar oligomer structures to better represent the polar fraction of the system. Candidate sugar oligomers for evaluation include those proposed by Denson et al. (31) (monomers to tetramers).

Evaluation of a lignin representative mixture

As can be observed in Figure 5, the pyrolignin representatives A1 and G2 presented contrasting trends for BE system. We propose that mixtures of these two molecules may give room for optimization as representative of pyrolignin to more accurately predict the K_{OW} of levoglucosan. Mixtures named M1-9 have been tested, corresponding to increasing mass fractions of A1 versus G2 (M1:10%/90%, M2:20%/80%, M3: 30%/70%, M4:40%/60%, M5:50%/50%, M6:60%/40%, M7:70%/30%, M8:80%/20%, and M9:90%/10%). The results indicate only minor improvements of the MRE of the ratio O/A (MRE: BE

28.4% to 34.6%, WE 43.2% to 48.0%) when compared to the studies performed with either A1 (MRE: BE 35.1%, WE 48.6%) or G2 (MRE: BE 28.1%, WE 42.4%).

The distribution of the *bio-oil, solvent, and water* fractions was improved by introducing mixtures M1-M9. Significant enhancement was achieved in the phase separation Ratio OA in the WE system, reducing the MRE from 42% and 48% (for A1 and G2 individually) to 28% with M8 and M9. However, no improvements were observed in the BE system. It was also possible to slightly improve the K_{ow} prediction of the bio-oil and the solvent (less than a 3% decrease), with the water fraction showing better results, achieving a 7% decrease.

Figure 7 presents the evolution of the K_{ow} of levoglucosan when employing the M1-9 as pyrolygnin representatives (also see Table S13). Despite very low MREs (< 10%) when using these mixtures, several of them overestimate the BE_{1-5} points, while all of them overestimate BE_6 , with parallel plateaus. The trends for WE are better, also with parallel plateaus. The best approximations were found for the mixture M2 with MRE of 8.8%, while increasingly worse results are observed for higher fractions of G2.

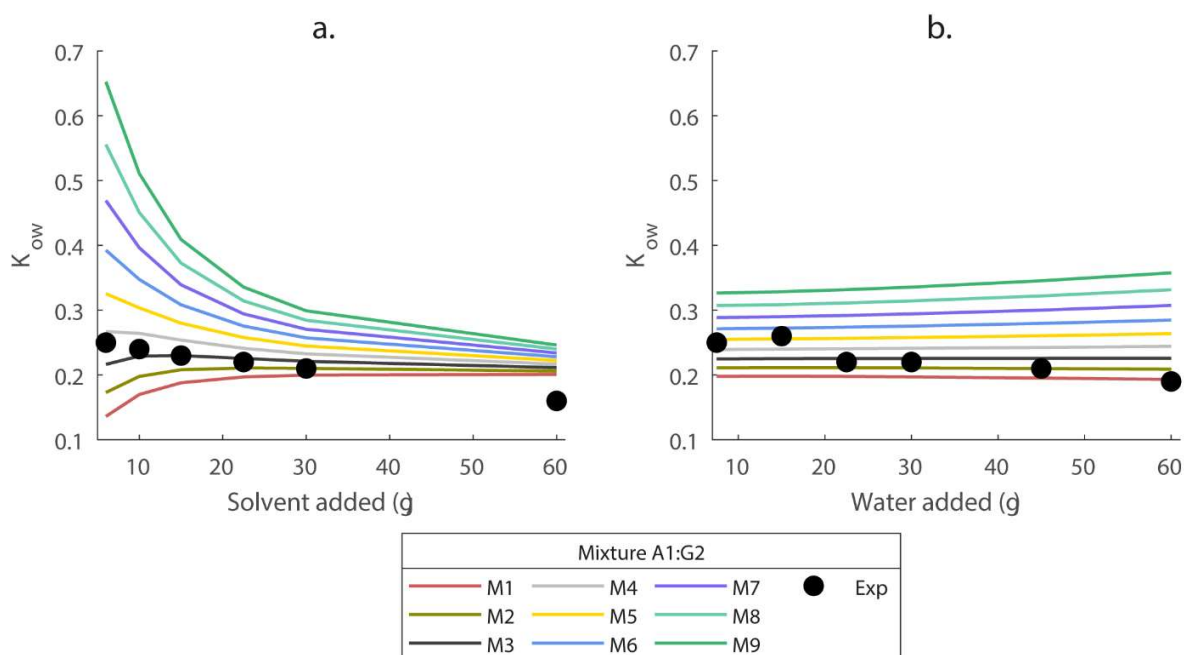


Figure 7. Evolution of the experimental and estimated value for the K_{ow} of levoglucosan for the mixtures M1-M9. a. BE system, b. WE system.

This study suggests that mixing two molecules (dimer and trimer) to represent the pyrolygnin fraction improves the result because it captures a broader range of molecular interactions and physicochemical properties present in the pyrolygnin fraction of FPBO. Different molecules interact with solvents and other components in unique ways and combining two molecules with opposite tendencies

(in K_{ow} levoglucosan prediction) allows the model to simulate these complex interactions more accurately. Complementary properties of different molecules, such as different solubility in the organic and aqueous phases, help achieve a balanced representation.

Conclusions

This work presents a first approach to predict liquid-liquid phase equilibria during extraction of FPBO with n-butanol and water using the UNIFAC-Dortmund model. With a carefully chosen surrogate mixture for the FPBO, lumped parameters like the mass ratio organic:aqueous and partition coefficients of the solvent, lumped organics and individual species are predicted with a reasonable error. Representing the largely unknown oligomers in FPBO with a lignin-derived representative compound ('pyrolignin') achieves an MRE of around 27 % for the ratio O/A and is an important improvement as compared to modeling phase equilibria without a pyrolignin representation (MRE > 57%). The K_{ow} of n-butanol was generally predicted better than that of water. The choice of the pyrolignin representative affects the quality of prediction for the different partition coefficients in opposing ways (i.e., improving one while worsening others), giving little general preference for one representative over the other. Interestingly, 3,4,4'-biphenyltriol (B2 in the context of this work), which has also been shown to be a promising representative for pyrolignin for vapour-liquid phase equilibria calculations, showed the best approximation to the experimental data for the K_{ow} of n-butanol in the aqueous phase (MRE of 5.4%).

Similar to VLE prediction, the deviation from experimental results can be very different for individual compounds and is also strongly affected by the choice of pyrolignin representative. In that sense, the general trend of the partition coefficient for levoglucosan is captured very well, which decreases with the addition of n-butanol and increases with water addition. For this specific case, the surrogate mixture without pyrolignin surprisingly provided a very good prediction (MRE <10%) which was also achieved with some pyrolignin representatives (an alkyl-aryl ether β -O-4 trimer and mixed tetramers, all with MRE values between 8 to 12%). This is an indication that sugar oligomers might play an important role during LLE and that the model could be improved by considering representatives for them.

Different to modelling VLE, some pyrolignin representatives show opposing trends when compared to experimental data, allowing for optimization with a mixture of representatives. E.g., combining a dimer (phenylcoumaran β -5) and a trimer (stilbene β -5) significantly improved the model's prediction for levoglucosan extraction with an MRE <10% for the K_{ow} of levoglucosan in both the n-butanol and water extraction. Mixtures of two pyrolignin representatives with average performances and opposite trends were considered to optimize the modeling of the partition of levoglucosan, and presented favorable results compared to the individual representatives. This indicates that carefully curated mixtures of surrogates can enhance the accuracy of phase separation predictions. However, mixtures based on advanced analytics of extracted pyrolytic lignin do not seem to perform particularly well when modeling

LLE systems such as this one, marking an important distinction between surrogates that are thermodynamically and physicochemically interesting.

While the model represents an important first step in describing liquid-liquid phase equilibria for FPBO, further details need to be clarified in order to achieve a model that is suitable from an engineering perspective. This primarily involves the availability of thermophysical data of individual compounds and other relevant parameters to run phase equilibria models since this limits the definition of a surrogate mixture and consequently the quality of prediction. From this study it becomes also clear that experimental data for phase equilibria, specifically in the domain of LLE, is very scarce and further efforts for high quality data are desperately needed to enable detailed optimization of phase equilibrium calculations.

Future work must focus on applying such tests, both experimentally and modeling-wise, to broadly different bio-oils obtained from different feedstocks and process conditions, in order to estimate the level of error expected from this method. It is also recommended that the modeling refines the range of possible representatives for the pyrolytic lignin fraction, but also employs representatives for sugar oligomers (e.g., humins), to further improve the accuracy of phase separation predictions.

Author Contributions

Myriam Rojas: Conceptualization, data gathering, data processing, data analysis, manuscript writing, and preparation of supporting information;

Frederico G. Fonseca: Conceptualization, model setup, data gathering, data processing, manuscript writing, preparation of supporting information, review of manuscript with significant changes;

Ana C. C. Araujo: Conceptualization, data gathering, manuscript writing

Axel Funke: Conceptualization, review of manuscript with significant changes

Manuel Garcia-Perez: Conceptualization, review of manuscript with changes

Current addresses

M. Rojas:

F.G. Fonseca: Institute for Low Carbon Processes, German Aerospace Agency, Walther-Pauer-Straße 5, 03046 Cottbus

| : frederico.gomesfonseca@dlr.de

References

1. Oasmaa A, Fonts I, Pelaez-Samaniego MR, Garcia-Perez ME, Garcia-Perez M. Pyrolysis Oil Multiphase Behavior and Phase Stability: A Review. *Energy and Fuels* [Internet]. 2016 Aug 18 [cited 2024 Jun 27];30(8):6179–200. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.energyfuels.6b01287>
2. Oasmaa A, Sundqvist T, Kuoppala E, Garcia-Perez M, Solantausta Y, Lindfors C, et al. Controlling the Phase Stability of Biomass Fast Pyrolysis Bio-oils. *Energy and Fuels* [Internet]. 2015 Jul 16 [cited 2023 Oct 25];29(7):4373–81. Available from: <https://pubs.acs.org/doi/abs/10.1021/acs.energyfuels.5b00607>
3. Jeon H, Park JY, Lee JW, Oh CH, Kim JK, Yoon J, et al. Fractional Composition Analysis for Upgrading of Fast Pyrolysis Bio-Oil Produced from Sawdust. *Energies* 2022, Vol 15, Page 2054 [Internet]. 2022 Mar 11 [cited 2024 Jul 12];15(6):2054. Available from: <https://www.mdpi.com/1996-1073/15/6/2054/htm>
4. Schmitt CC, Boscagli C, Rapp M, Raffelt K, Dahmen N. Characterization of Light and Heavy Phase of Pyrolysis-Oil from Distinct Biomass for Further Upgrading Reactions. *European Biomass Conference and Exhibition Proceedings* [Internet]. 2017 [cited 2024 Jun 27];1143–7. Available from: <http://www.etaflorence.it/proceedings/?detail=13629>
5. Johansson AC, Lisa K, Sandström L, Ben H, Pilath H, Deutch S, et al. Fractional condensation of pyrolysis vapors produced from Nordic feedstocks in cyclone pyrolysis. *J Anal Appl Pyrolysis*. 2017 Jan 1;123:244–54.
6. G Fonseca F, Funke A. Modeling of liquid-vapor phase equilibria of pyrolysis bio-oils: a review. *Ind Eng Chem Res*. 2024;
7. Parku GK, Krutof A, Funke A, Richter D, Dahmen N. Using Fractional Condensation to Optimize Aqueous Pyrolysis Condensates for Downstream Microbial Conversion. *Ind Eng Chem Res* [Internet]. 2023 Feb 15 [cited 2024 Jan 15];62(6):2792–803. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.iecr.2c03598>
8. Parku GK, Funke A, Dahmen N. Influence of selected quench media used for direct contact condensation on yield and composition of fast pyrolysis bio-oils aided by thermodynamic phase equilibria modelling. *Sep Purif Technol*. 2024 Aug 9;341:126873.

9. Rover MR, Johnston PA, Jin T, Smith RG, Brown RC, Jarboe L. Production of Clean Pyrolytic Sugars for Fermentation. *ChemSusChem* [Internet]. 2014 Jun 1 [cited 2024 Jul 3];7(6):1662–8. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1002/cssc.201301259>
10. Basafa M, Hawboldt K. A review on sources and extraction of phenolic compounds as precursors for bio-based phenolic resins. *Biomass Conversion and Biorefinery* 2021 13:6 [Internet]. 2021 Mar 2 [cited 2024 Jul 2];13(6):4463–75. Available from: <https://link.springer.com/article/10.1007/s13399-021-01408-x>
11. Fonseca FG, Funke A, Niebel A, Soares Dias AP, Dahmen N. Moisture content as a design and operational parameter for fast pyrolysis. *J Anal Appl Pyrolysis* [Internet]. 2019 May;139:73–86. Available from: <https://doi.org/10.1016/j.jaap.2019.01.012>
12. Fonseca FG, Funke A, Niebel A, Dias APS, Dahmen N. Moisture content as a design and operational parameter for fast pyrolysis. In: 1 Deutsches Doktorandenkolloquium Bioenergie. DBFZ Deutsches Biomasseforschungszentrum gemeinnützige GmbH; 2018. p. 286.
13. Şen AU, Fonseca FG, Funke A, Pereira H, Lemos F. Pyrolysis kinetics and estimation of chemical composition of *Quercus cerris* cork. *Biomass Convers Biorefin* [Internet]. 2020 Aug 19; Available from: <http://link.springer.com/10.1007/s13399-020-00964-y>
14. Gustavsson C, Nilsson L. Co-production of pyrolysis oil in district heating plants: Systems analysis of dual fluidized-bed pyrolysis with sequential vapor condensation. *Energy and Fuels*. 2013;27(9):5313–9.
15. Fonts I, Atienza-Martínez M, Carstensen HH, Benés M, Pires APP, Garcia-Perez M, et al. Thermodynamic and physical property estimation of compounds derived from the fast pyrolysis of lignocellulosic materials. *Energy and Fuels* [Internet]. 2021 Nov 4 [cited 2024 May 13];35(21):17114–37. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.energyfuels.1c01709>
16. Pinheiro Pires AP, Arauzo J, Fonts I, Domine ME, Fernández Arroyo A, Garcia-Perez ME, et al. Challenges and opportunities for bio-oil refining: A review. *Energy and Fuels* [Internet]. 2019 Jun 20 [cited 2024 Jul 5];33(6):4683–720. Available from: <https://pubs.acs.org/sharingguidelines>
17. Stephan C, Dicko M, Stringari P, Coquelet C. Liquid-liquid equilibria of water + solutes (acetic acid/acetol/furfural/guaiacol/methanol/phenol/propanal) + solvents (isopropyl acetate/toluene) ternary systems for pyrolysis oil fractionation. *Fluid Phase Equilib*. 2018 Jul 25;468:49–57.

18. Ille Y, Sánchez FA, Dahmen N, Pereda S. Multiphase Equilibria Modeling of Fast Pyrolysis Bio-Oils. Group Contribution Associating Equation of State Extension to Lignin Monomers and Derivatives. *Ind Eng Chem Res* [Internet]. 2019 May 1 [cited 2024 Jul 2];58(17):7318–31. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.iecr.9b00227>
19. Jalalinejad A, Seyf JY, Funke A, Dahmen N. Phase Equilibrium Calculation of Bio-Oil-Related Molecules Using Predictive Thermodynamic Models. *Energy & Fuels* [Internet]. 2024 Feb 6 [cited 2024 Feb 14]; Available from: <https://pubs.acs.org/doi/full/10.1021/acs.energyfuels.3c04395>
20. Han Y, Pinheiro Pires AP, Denson M, McDonald AG, Garcia-Perez M. Ternary phase diagram of water/bio-oil/organic solvent for bio-oil fractionation. *Energy and Fuels* [Internet]. 2020 Dec 17 [cited 2023 Oct 25];34(12):16250–64. Available from: <https://pubs.acs.org/doi/abs/10.1021/acs.energyfuels.0c03100>
21. Hempfling R, Schulten HR. Chemical characterization of the organic matter in forest soils by Curie point pyrolysis-GC/MS and pyrolysis-field ionization mass spectrometry. *Org Geochem*. 1990 Jan;15(2):131–45.
22. Fonseca FG, Funke A, Dahmen N. Aspen Plus™ modeling of Fractional Condensation schemes for production of Fast Pyrolysis bio-oil. In: 27th European Biomass Conference and Exhibition. Lisbon: ETA-Florence Renewable Energies; 2019. p. 681–3.
23. Li Q, Vlachos DG. 5-5 Lignin Linkage Cleavage over Ru: A Density Functional Theory Study. *ACS Sustain Chem Eng* [Internet]. 2021 Dec 6 [cited 2024 Jul 3];9(48):16143–52. Available from: <https://pubs.acs.org/doi/full/10.1021/acssuschemeng.1c04838>
24. Adler E, Eriksoo E, Boss E, Cagliaris A. Guaiacylglycerol and its beta-Guaiacyl Ether. *Acta Chem Scand*. 1955;9:341–2.
25. Nakatsubo F, Sato K, Higuchi T. Synthesis of guaiacylglycerol- β -guaiacyl ether. *Holzforschung* [Internet]. 1975 Jan 1 [cited 2023 Jun 16];29(5):165–8. Available from: <https://www.degruyter.com/document/doi/10.1515/hfsg.1975.29.5.165/html>
26. Manrique R, Terrell E, Kostetskyy P, Chejne F, Olarte M, Broadbelt L, et al. Elucidating Biomass-Derived Pyrolytic Lignin Structures from Demethylation Reactions through Density Functional Theory Calculations. *Energy and Fuels*. 2023 Apr 6;37(7):5189–205.

27. Kishimoto T, Uraki Y, Ubukata M. Synthesis of bromoacetophenone derivatives as starting monomers for β -O-4 type artificial lignin polymers. *Journal of Wood Chemistry and Technology*. 2008 Apr;28(2):97–105.
28. Kishimoto T, Uraki Y, Ubukata M. Synthesis of β -O-4-type artificial lignin polymers and their analysis by NMR spectroscopy. *Org Biomol Chem* [Internet]. 2008 Jul 31 [cited 2023 Jun 22];6(16):2982–7. Available from: <https://pubs.rsc.org/en/content/articlehtml/2008/ob/b805460f>
29. Ranzi E, Debiagi PEA, Frassoldati A. Mathematical Modeling of Fast Biomass Pyrolysis and Bio-Oil Formation. Note I: Kinetic Mechanism of Biomass Pyrolysis. *ACS Sustain Chem Eng* [Internet]. 2017 Apr 3 [cited 2023 Jun 19];5(4):2867–81. Available from: <https://pubs.acs.org/doi/full/10.1021/acssuschemeng.6b03096>
30. Manrique R, Terrell E, Kostetskyy P, Chejne F, Olarte M, Broadbelt L, et al. Elucidating Biomass-Derived Pyrolytic Lignin Structures from Demethylation Reactions through Density Functional Theory Calculations. *Energy and Fuels* [Internet]. 2023 Apr 6 [cited 2024 Jun 17];37(7):5189–205. Available from: <https://pubs.acs.org/doi/full/10.1021/acs.energyfuels.2c04292>
31. Denson MD, Terrell E, Kostetskyy P, Olarte M, Broadbelt L, Garcia-Perez M. Elucidation of Structure and Physical Properties of Pyrolytic Sugar Oligomers Derived from Cellulose Depolymerization/Dehydration Reactions: A Density Functional Theory Study. *Energy and Fuels* [Internet]. 2023 Jun 1 [cited 2024 Jul 10];37(11):7834–47. Available from: <https://pubs.acs.org/doi/abs/10.1021/acs.energyfuels.3c00641>
32. George KParku,, Axel Funke,. Phase equilibria aided optimization of Levoglucosan extraction during condensation of fast pyrolysis bio-oils. *Energy & Fuels*. 2024;